The *Ushichka* dataset: a multi-channel, multi-sensor dataset to unravel active sensing in groups

Thejasvi Beleyur, Holger R. Goerlitz

Document updated: 2021-01-06

The collective behaviour and motion of animal groups has been intensely investigated. The behavioural heuristics underlying the impressive acts of self-organisation have been revealed through a combination of experimental tracking and modelling. More specifically, visually driven animal collectives such as fish and mammals have been the focus of studies. Visually driven collective behaviours scale well with group size given sufficient ambient light. In contrast to passive sensing animals, active sensing animals like echolocating bats emit probes of energy to detect their surroundings. Echolocating bats emit loud ultrasonic calls and listen for faint returning echoes to detect objects. When multiple bats echolocate together however, it is expected that they will not be able to detect their surroundings well due to masking. This masking means that collective behaviour in active sensing bats is not expected to scale well with group size. However, despite this expectation, bats aggregate and fly together in large groups during cave emergences and mating swarms. The sensory and behavioural heuristics behind these formations remain understudied to date. The study of group echolocation has also been impeded by technological issues in experimental tracking, analysis of recordings with high call densities, and a deficiency of parametrised computational models. To fill these gaps in the study of active sensing groups, we present the *Ushichka* dataset. *Ushichka* is a multi-channel, multi-sensor dataset consisting of audio and video data of bats echolocating as they fly in a cave. A LiDAR scan of the recording volume is also part of the dataset. *Ushichka* promises to generate a detailed spatio-temporal understanding of how bats in groups deal with the problem of masking as they echolocate together in close proximity. The dataset is, to our knowledge, one of its kind in capturing bat collective behaviour with multiple sensors, and promises novel insights into the sensorimotor heuristics of active sensing groups. While promising novel insights, *Uschichka* also presents a series of technical challenges ripe for inter-disciplinary collaborations such as LiDAR-thermal camera alignment, multi-channel call correspondence matching and automatic microphone position estimation. The processed dataset will consist of individual level trajectory and call information that are placed in a LiDAR scan of the cave. This processed dataset can be used to parametrise existing models of collective behaviour, compare behavioural heuristics with group size, and to reconstruct the auditory scene of each bat in the group.

## Introduction

Many animals across the tree of life move and live in groups, and show impressive coordinated behaviours (Sumpter 2006). Coordinated behaviours shown in these collectives are reproduced by simple heuristics in models with agents responding independently to the movement and positions of their neighbours (Couzin et al. 2002). These coordinated behaviours have been well investigated in visually dominant animal collectives such as birds, mammals and fish (Ballerini et al. 2008; Strandburg-Peshkin et al. 2013; Pita et al. 2016). Vision is a ‘passive’ sensory modality, in that the animal acts solely as a receiver of light (Nelson and MacIver 2006). Visually driven collective behaviour can thus scale well with group size, given sufficient ambient light.

In contrast to passive sensing animals, are ‘active’ sensing animals. Active sensing animals (sensu stricto), such as bats, dolphins and electric fish, emit probes of energy and monitor the modulation of the probes by the environment around them (Nelson and MacIver 2006). Echolocating bats emit loud ultrasonic calls, and listen to the returning echoes to detect objects around them (Griffin 1958). When multiple echolocating bats echolocate together, they can start to negatively affect each other’s ability to detect their own echoes (Ulanovsky and Moss 2008). Every bat emits loud calls, and is attempting to listen to its own echo. However, along with its own returning echoes, each bat hears loud ‘non-target’ sounds (calls and echoes of neighbouring bats) which hamper its own echo detection. A bat that cannot detect its own echoes is metaphorically ‘flying blind’, and risks colliding into its neighbours , and of more concern, into hard structures around it. As the number of bats in a group increases, the number of non-target sounds increases rapidly in a non-linear fashion (~). Sensory simulations show that in large groups of bats, a bat may be detecting its neighbours every few calls - and not every time it emits a call (Beleyur and Goerlitz 2019). Despite the expected deterioration of echolocation with increasing group sizes, echolocating bats in nature are social and form very large aggregations in the form of mating swarms, cave roosts and evening emergences (Ortega 2016; Erkert et al. 1982; Gillam et al. 2010). The dynamics of active sensing groups is understudied (but see Theriault et al. 2010), and perhaps likely to differ from passive sensing groups in one respect. Given sufficient light, an animal in a visually driven collective will be able to detect all neighbours in its vicinity even at large group sizes. In contrast, the number of neighbours potentially detected by a bat in a group decreases with group size (Beleyur and Goerlitz 2019), leading to a limited sensing ability. While it is parsimonius to assume that existing computational models may be able to capture the dynamics of sensorially limited active sensing groups (through models with ‘limited interactions’ sensu (Bode, Franks, and Wood 2011)), it remains to be experimentally seen in what respects the behaviour of active sensing groups differ from the visually driven animal collectives studied so far. Even if active sensing groups do indeed behave much like visually driven passove sensing groups, experimental confirmation remains lacking to date.

Here we present the conceptual motivation and technical equipment behind the *Ushichka* dataset. The word Ushichka (Ушичка, *OO-shi-ch-kaa*) in Bulgarian has a dimunitive connotation for something with multiple ears. The name is chosen to highlight the fact that *Ushichka* is a multichannel, multi-sensor dataset which aims to unobtrusively ‘listen’ (observe with technological eyes and ears) to groups of multiple bats echolocating in the wild. The *Ushichka* dataset is, to our knowledge, a unique dataset in that it comprises of multichannel audio and video datasets of bats echolocating in groups in their *natural environment* captured by a LiDAR scan of the space. The multichannel nature of the dataset allows 3D localisation and tracking with high spatio-temporal resolution. Previous experimental work on multi-bat echolocation in the lab and field to date has mostly been with upto 2 bats a time (Götze et al. 2016; Giuggioli, McKetterick, and Holderied 2015; Necknig and Zahn 2011; Fawcett and Ratcliffe 2015; Fawcett et al. 2015), or multichannel data of solely video or audio (Theriault et al. 2010; Lin, Abaid, and Müller 2016). The multi-camera, multi-mic nature of *Ushichka* allows us to explore echolocation in groups in much greater detail than previously attempted.

### The experimental study of echolocating bat groups

The study of bat echolocation has a long and grand history dating back over 200 years, first documented in Spallanzi’s experiments in the 18th century. After the formal discovery of echolocation (Griffin 1958; Dijkgraaf 1946), the field has moved leaps and bounds. A detailed understanding of the physiology, behaviour, neurobiology and acoustics behind echolocation has been uncovered in great detail (Popper and Fay 1995; Fenton et al. 2016). While a large body of knowledge is in place about how *individual bats* are able to echolocate through the use of trained animals in laboratory settings and field observations (Fenton et al. 2016), comparitively little is known on how they echolocate in groups(Ulanovsky and Moss 2008).

Previously, the challenges in echolocation fieldwork included bulky analog recording and signal analysis equipment; limiting the speed and quantity of data that could be collected, and analysed (Griffin 1958; Habersetzer 1981). The onset of the digital revolution has allowed for lightweight devices that can collect more data, which can then be analysed more easily. Current experimental approaches to study echolocation in groups include, for example, on-body tags that record audio from the focal animal and its neighbours (along with GPS and accelerometer data)(Cvikel et al. 2015), microphone-camera pairs to quantify broad echolocation characteristics with group size (Lin, Abaid, and Müller 2016) and thermal camera 3D tracking with sophisticated computer vision to track individuals in the million-strong emergence behaviour of *Tadarida brasiliensis* (Betke et al. 2007; Theriault et al. 2010). Each of these studies have provided exciting and important insights into how echolocating bats manage to echolocate in groups. For instance, Cvikel et al. (2015) showed that *Rhinopoma hardwickei* do not seem to show the ‘jamming avoidance response’ (shifting call frequencies to avoid overlap with neighbours, sensu Ulanovsky et al. (2004)) first seen by Habersetzer (1981). Lin, Abaid, and Müller (2016) find that bats echolocating in caves actually increase their pulse rate with group size, in contrast to theoretical expectations and results from another species that decreases its pulse rate (Jarvis, Jackson, and Smotherman 2013). Theriault et al. (2010) advanced the field by quantifying the geometry of a *T. brasiliensis* emergence, showing that animals in larger groups fly relatively close to each other (0.5m), and that neighbours are placed uniformly across azimuth and elevation around individuals, unlike that seen in bird flocks (Ballerini et al. 2008). In starling flocks for instance, neighbours are located non-uniformly in the elevation and azimuth in that they are less likely to be found in front or behind a focal individual.

*Ushichka* represents a natural and quantitative step forward to all the recent work in the study of group echolocation in the field. The one important respect where *Ushichka* differs from previous work is its inherent ‘multi-modality’, in that it consists of simultaneously captured multi-microphone, multi-camera data streams, along with 3D scans of the recording volume for physical context. On-body tags and single microphone recordings lack explicit information of how many or where neighbours were located. Pure video tracking with one or more cameras does not provide insights into the echolocation behaviour of animals across group sizes. *Ushichka* captures flight and echolocation data of the group, and holds promise to provide access to the auditory scene (Moss and Surlykke 2001) of each individual in the group. Reconstructing the sensory inputs of individuals in a group allows us to better understand the sensorimotor decisions animals make (Strandburg-Peshkin et al. 2013). While an established method in visually driven animal groups like fish, auditory scene reconstruction is an exciting new frontier for the field of group echolocation, that *Ushichka* promises to provide.

## Experimental methods

### Field site and study species

Recordings of wild echolocating bats were made in the Orlova Chuka cave system (Ruse, Bulgaria). Recordings were made of bat flight and echolocation in the chamber first encountered after the entry/exit corridor (Figures 1, 2). Recordings were conducted in the recording volume (*w x l x h*~5x9x3m) between 19th June to 19th August 2018 over the course of 14 nights (Suplementary Information: Table 1).

Orlova Chuka is known to host at least 9 species of bats (Ivanova 2005; “Natura 2000 Bulgarian Ministry of Environment and Water Report” 2013). During the summer, resident *Myotis myotis* and *Myotis blythii* females form the dominant bat population. These female bats arrive in spring and spend the summer in the cave system raising their young. Though rigorous mark-recapture data is missing , individual mothers are known to be consistent in returning to the cave system over multiple nights (pers comm, Laura Stidsholt, Stefan Greif). Given the dominance of *Myotis myotis* and *Myotis blythii* in the cave system, it can be fairly said that the *Ushichka* dataset consists mainly of these two species. Moreover, the frequency ranges of *M. myotis* and *M. blythii* are lower than those of the other resident frequency modulating bats in the cave (eg. *M. daubentonii, M. capacinii, Miniopterus schreibersii*) and are very easily distinguishable from the rhinilophid bats (*Rhinolophus ferrumequinum, R.euryale, R. mehelyi*).



Figure 1: The *Ushichka* setup. The large inverted T shaped array in the centre of the image is the 120cm tristar array with four SANKEN CO-100 microphones. The three blue lights are from the three thermal cameras pointing towards the array. The rest of the microphones are smaller Knowles SPU 0410 units are attached to the wall on the left portion of the image and are not visible.



Figure 2: The *Ushichka* setup with the smaller Knowles SPU microphones on the wall. In the picture is a *Myotis myotis* bats flying past the array. Photo by Stefan Greif.

### Audio, video and LiDAR systems

#### Audio

Multichannel audio data was recorded by multiple (2-3) Fireface (RME GmbH, Germany) sound cards running at 192kHz (Figure 5) connected to a laptop (DELL Latitude E6540, Windows 7 enterprise, core i5, 16GB RAM). The laptop was verified not to emit ultrasound. All sound cards were Fireface USB type, ASIO protocol based soundcards. The number of channels of audio in the *Ushichka* dataset varied between 12-22 channels over the course of the field season. The 12-22 channel data consisted of four SANKEN CO-100 microphones with the rest of the 8-18 channels coming from Knowles SPU-0410 microphones. All microphones were connected to either the inbuilt pre-amplifiers of the sound cards, or to two RME QuadMic II (RME GmbH, Germany) pre-amplfiers. Later recordings with the *Ushichka* setup also included a Focusrite Scarlett OctoPre (Focusrite Plc, UK) pre-amplifier.

All recording nights had a 4-channel 120cm tristar mounted in a fairly constant position in front of the memorial plaque in the recording volume. A ‘tristar’ array is a common planar array configuration used in previous publications (Hügel et al. 2017; Goerlitz et al. 2010; Lewanzik and Goerlitz 2018), with three microphones placed at a fixed radius from a central microphone. The three peripheral microphones are separated from each other by a 120 degree angular separation. The entire set of four microphones is accomodated on an inverted-T shaped metal array. The Knowles SPU-0410 microphones were spread around the recording volume to increase the variety of locations at which bat calls were recorded. The positions of microphone were constant within a recording night, and were often changed between nights. On a few occasions, the microphones were kept in the same position across consecutive recording nights. Microphones were not left in the cave across multiple days due to concerns of humidity affecting the electronic circuits and membranes, especially of moisture build up in the Knowles SPU 0410 recording inlet.

A subset of all possible inter-mic distances were measured every recording session using a GLM 50C (Bosch, Germany with ~1mm accuracy)laser range finder . Ultrasonic sweeps were played back from multiple points in the recording volume as part of an automatic mic position estimation workflow. Most of the microphones were calibrated for their frequency response and directionality at least once after the field season. For more information on the microphone automatic position estimation and calibrations please refer to the Supplementary Information.

#### Video

Three thermal cameras (TeAx ThermalCapture 2.0, TeAx Technology GmbH, Germany) running with FLIR Tau 640 cores (9mm focal length, 640x512 pixel image resolution) at 25Hz formed the camera array. All three cameras (Figure 3) were frame synchronised. There were a few occasions where one or two cameras recorded an extra frame, and these extra frames are known from previous tests to be at the end of the recording, and not jitter in recording initiation.

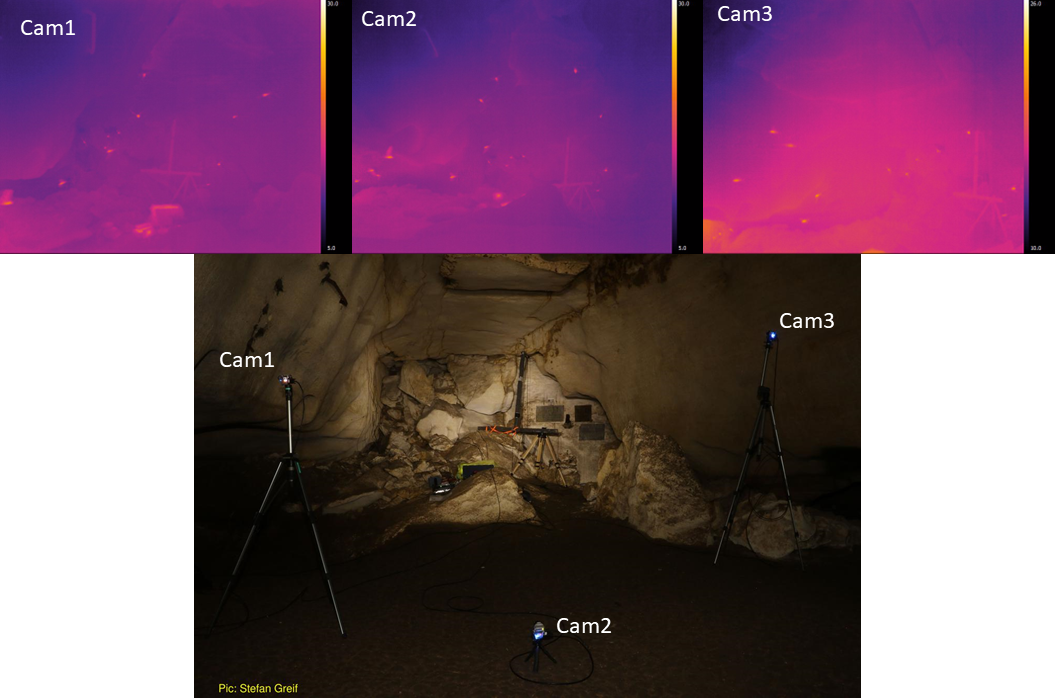


Figure 3: A multi-camera view of bats flying in the recording volume, allowing 3D tracking of flying bats (from the frame-synchronised thermal video recordings).

The microSD cards (SAMSUNG PRO+, 32 GB) were inserted into each camera at the start of each recording session. Removing the microSD card from the cameras required significant disturbance of their position, and thus the total video recording time was limited by the one-time microSD card memory capacity. The location of the microSD cards *inside* the camera housing meant that the data from the cameras could only be transferred in the morning after the end of the recording session. All cameras were calibrated using the ‘wand’ protocol described in Theriault et al. (2014). More details of the camera calibration runs are in the Supplementary Information. Further details on camera intrinsic and extrinsic parameter calibrations, wand setup, and synchronisation with audio are in the Supplementary Information.

#### LiDAR

A LiDAR scan of the recording volume provides a physical context to place the observed echolocation and flight behaviour of the bats. With only the audio and video data, we will at most be able to recreate the positions and time of call emissions of the bats themselves. While bats are likely reacting to each other’s presence, they will also be reacting to the presence or absence of physical obstructions such as cave walls. With the LiDAR scan for example, we will be able to understand if a change in trajectory or call behaviour arose from the bats’ proximity to the cave wall, or to its nearest neighbour.

A LiDAR scan of the recording volume, along with a large portion of the Orlova Chuka cave was carried out in collaboration with Dr. Asparuh Kamburov (University of Mining and Geology “St. Ivan Rilski”, Sofia). The LiDAR scan provides a high spatial resolution (<1cm) 3D map of the cave’s surface and objects in it (Figure 4). The details of the LiDAR scanning, data processing and description can be found in Kamburov, Goerlitz, and Beleyur (2018).

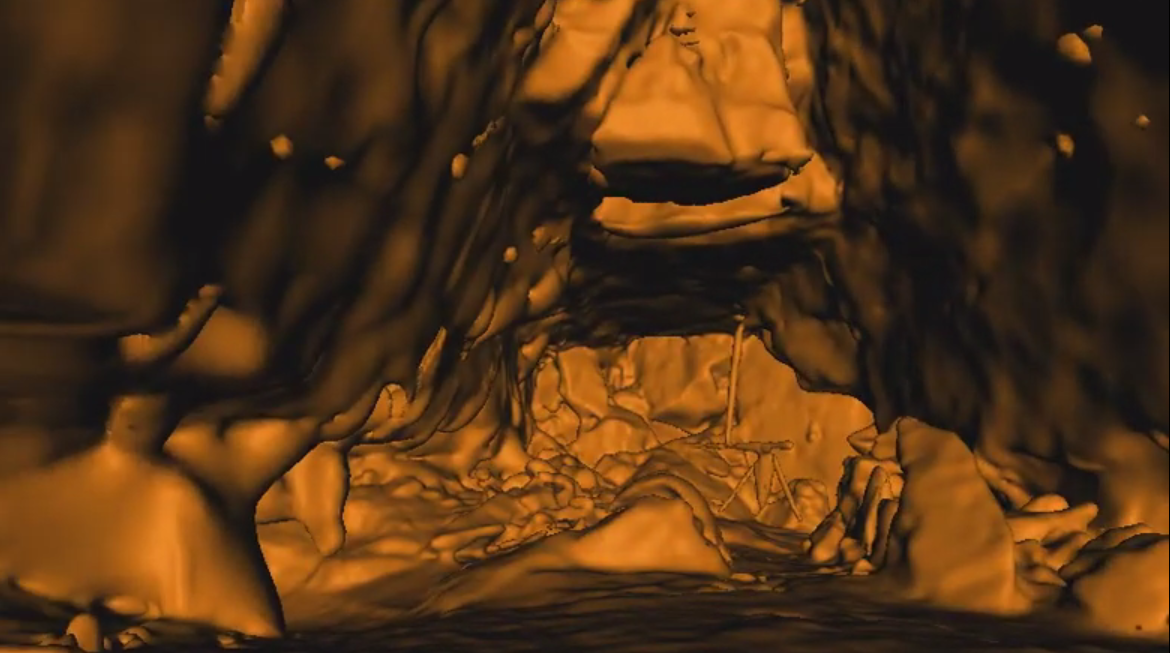


Figure 4: The LiDAR map of the recording volume showing the same view as in Figure 3, with the large tristar array.

#### Weather data

The local weather conditions (humidity, temperature and atmospheric pressure) within the cave were recorded by a Kestrel 4000 weather meter (Nielsen-Kellerman, USA) placed in the recording volume. The logged weather variables play a role in the accuracy of position and source level estimates (Goerlitz 2018; Lawrence and Simmons 1982), and were thus measured to later estimate the speed of sound and absorption of ultrasonic sound more accurately across recording sessions.

### Recording schedule

The audio and video arrays were mostly setup before sunset (~21:00 Eastern European Time) and collected data till around sunrise (~5:30 EET) with one break in between. The recording break between around 12-3 am co-incided with an observed drop in bat activity around midnight.

Recordings were triggered manually in 2 of 14 recording sessions in response to bat calls heard on a bat detector (SI Table 1). For the rest of the sessions, recordings were triggered automatically. A recording was automatically triggered by bat calls that went above a threshold rms level (between -50 to -40 dB, threshold adjusted across recording sessions) across a subset of monitor channels. The recording was set to stop after a pre-defined duration. There were no other sources of ultrasonic sound in the cave, and it can be said with great certainty that all recordings were triggered only by bat calls. Each recording in the *Ushichka* dataset consists of 10-15 second long audio-video file pairs (See SI for details of audio-video synchronisation). The variation in recording duration is due to adjustments made over the course of the field season.

Bats can emit over 10 calls per second, which would lead to an almost continuously triggered recording system. While technically feasible, continuous recording of data would lead to unmanageably large files. Moreover, the TeAX thermal cameras relied on high speed microSD cards that were the limiting factor in the amount of data (32 GB) that could be collected over the course of a recording session. The audio-video recording system was thus set to record at a fixed duty cycle of around 20% across sessions. For example, a 20% duty cycle and 10 second recording duration meant that recordings could be triggered at most every 50 seconds (10s recording + 40s inactivity).

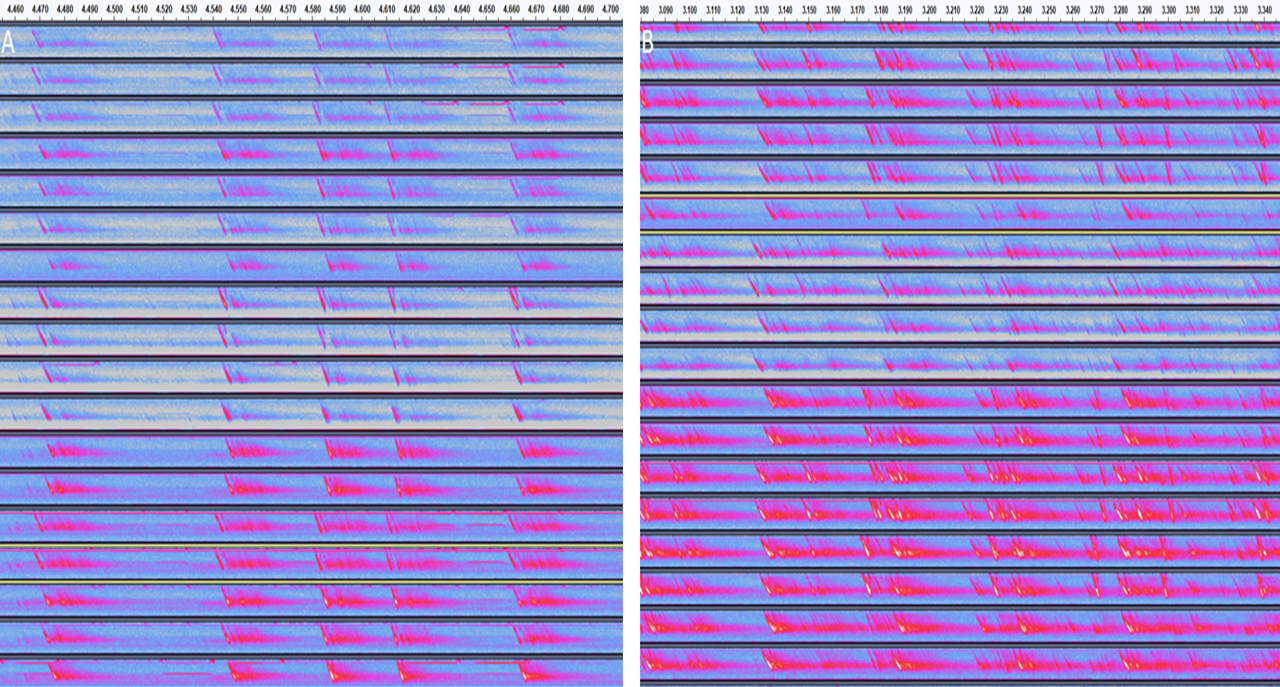


Figure 5: Spectrograms showing a subset channels from the multichannel audio data collected. The spectrograms show highlight some of the challenges of working with the audio data including reverberation, multi-path propagation, call overlaps, and correct call identity matching across channels. A) an example of a recording with a few bats in it B) recording with multiple bats in it.

There are a few occasions of skipped morning and evening recordings due to various technical and logistical issues that were encountered over the course of the field season.

### Bat activity

Bats began to fly from the interior roosting sites into the recording volume around the time of sunset. The initial activity is typically due to of small groups (2-5) of *R.euryale* and *R. mehelyi* flying rapidly and unpredictably in the volume. The activity of the rhinolophid bats seemed to reduce with time, and gave way to a slowly rising number of *M. myotis* and *M. blythii* (hereon referred to as *M. myotis/blythii*) bats. The number of bats flying in the recording volume could be as high as thirty bats at a time. Towards end June, it was noticed that the *M. myotis/blythii* bats also formed temporary roosting clusters on the walls of the recording volume, and this behaviour seemed to decrease in its frequency as the season progressed. Cloudy and rainy evenings lead to a delay or complete absence of emergence from the cave.

Peak bat activity in the recording volume was typically seen about 1-1.5 hours after sun-set. The bats then began to exit the cave (Figure 6). The exit seemed relatively rapid, after which only a few bats were seen to still fly in the recording volume. Individual bats returned in the early morning from one to two hours before sunset onwards. It seemed that *R. euryale* and *R. mehelyi* activity remained fairly constant through the night, even during the drop in activity around midnight.

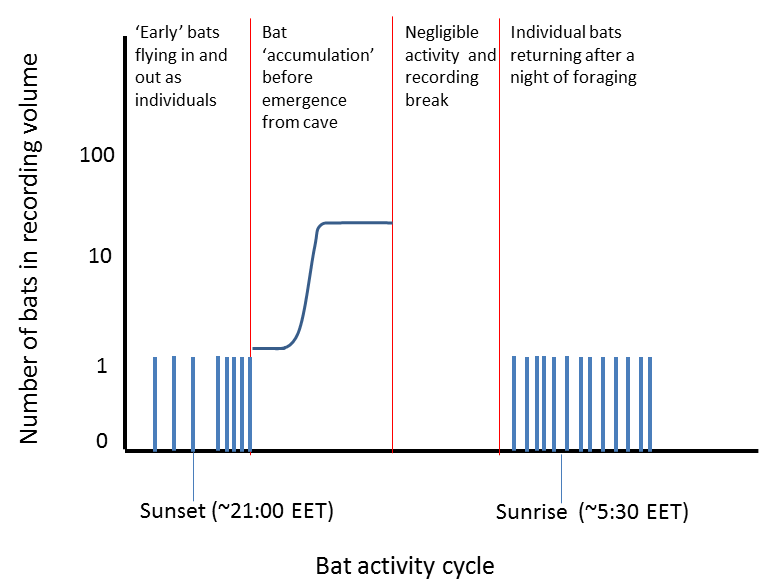


Figure 6: Schematic of typical bat activity pattern of Myotis myotis and Myotis blythii between sunset to sunrise in the *Ushichka* dataset. The multichannel recording system was typically activated before sunset and was run till after sunrise with a break in between

### Software used in this work

Automatic recording triggering implemented in the fieldrecorder\_trigger module relies on the open-source Python language (Van Rossum and Drake 2011), and the sounddevice, scipy and numpy (Geier 2015; Virtanen 2019; Oliphant 2006) packages. This manuscript was written using the knitr package (Xie 2015).

## Technical challenges ahead and potential solutions in analysing the *Ushichka* dataset

Analysing the echolocation call data in *Uschichka* will require surmounting a series of challenges, here discussed in increasing order of difficulty.

1. Reflections : The recording volume consisted of acoustically reflective surfaces, leading to significant reverberation and multi-path propagation. Dealing with reverberation is likely to be relatively straightforward while performing acoustic tracking, as cross-channel correlations should be able to filter out the noise contributed by reverberation. Multi-path propagation, where a sound arrives multiple times at each microphone due to distinct reflections is likely to be a somewhat more difficult problem to deal with, especially considering that if a single call leads to *N* multi-path reflections, then *M* calls in the volume could then lead to *NxM* reflections.
2. Call density : Each bat flying in and out of the volume emits calls 10-20 times per second (10-20 Hz call rate). With one bat, it is relatively straightforward to acoustically localise the bat as calls are separated by tens of milliseconds. With an increasing number of bats in the volume, (as happens towards the end of the night, Figure 6), the temporal separation between calls will decrease, and also lead to a higher occurence of call overlaps. Handling such high call densities will be a challenge, and require the latest techniques in call detection in the presence of overlaps (Izadi, Stevenson, and Kloepper 2020), or in general blind signal separation techniques (Brandstein and Ward 2013).
3. Correspondence matching : Even if calls can be detected reliably in single channels, the problem of cross-channel correspondence still remains to be solved,a common problem faced in any multi-channel acoustic array (Brandstein and Ward 2013). An echolocation call detected in one channel needs to be matched with itself in another channel, despite the possible presence of multiple similar calls in the expected time-window. Individual bats do have unique call characteristics (Masters, Raver, and Kazial (1995), and specifically in *M. myotis*, Yovel et al. (2009) ; but see Siemers and Kerth (2006)). The unique echolocation call ‘signatures’ are sometimes apparent from visual inspection of the spectrogram. However, manual correspondence matching is unlikely to scale well in terms of effort required with group sizes beyond 3 bats and will require robust automated algorithms to solve this problem.

One important ‘support system’ which may help significantly in solving the issues of multi-path propagation and correspondence matching is the multi-camera video data. Each call is emitted from one position by a bat, and captured on multiple audio channels. The time-synchronised 3D video tracking data contains the microphone positions and bat trajectories. With this in mind, we can combine the audio and video tracking to simplify cross-channel correspondence matching and the process of assigning a call to its source bat. The basic principle of acoustic localisation relies on the fact that with at least 4 microphones spread in space, each position of sound emission will generate a unique set of time-differences-of-arrival (TDOA). A TDOA in a microphone array is the difference in arrival times of a sound in comparison to a chosen reference channel. Measuring these TDOAs allows inferring the position at which a bat emitted a call from. The 3D video trajectories allow the generation of a series of ‘hypotheses’ of which calls could have been emitted at which point in space. By calculating the time-varying distances between a bat and each microphone across time, it will allow us to generate a set of ‘predicted’ TDOAs that calls could be arriving at. The problem thus boils down to a matching problem, where all TDOAs from all possible call correspondence matches are filtered based on which TDOAs are predictions from a given flight trajectory. The video trajectory based filtering approach could in principle elegantly solve the problems of multi-path propagation and cross-channel correspondence at one shot. However, whether the 3D video data will actually be able to deliver sufficiently accurate position estimates, and thus accurate TDOAs remains to be seen.

### Fusing multiple data streams to reconstruct individual auditory scenes

*Ushichka* consists of multiple data streams consisting of audio, video and LiDAR data. The audio and video streams are tightly linked as they describe the flight and echolocation behaviour of bats in the recording volume. However, the bats in the recording volume are primarily trying to avoid collisions with the walls and structures in the cave while the location of other bats is likely secondary. This is where the LiDAR dataset provides an extremely important contextual understanding of how and why bats may be choosing to fly or echolocate in the space they occupy. Fusing the acoustic and video data streams is relatively straightforward in that they are time-synchronised and they both track common ‘objects’ (bats). The 3D positions derived from both data streams can be overlaid to match with a simple rigid rotation of coordinate systems. To understand the bats’ echolocation and flight in their physical context however, the LiDAR scan of the recording volume must be aligned with the coordinate systems of the camera array. This will allow us to interpret the sensory decisions of bats correctly, eg. whether a sudden turn was due to proximity to the cave wall.

Recent advances in the alignment of LiDAR scans with camera views have led to the formulation of ‘targetless’ workflows that do not require explicit calibration for alignment (Kang and Doh 2020; Muñoz-Bañón, Candelas, and Torres 2020). These approaches rely on recreating synthetic camera images by projecting the 3D LiDAR scan into 2D. The best LiDAR-camera alignment is one which which achieves the highest similarity between the synthetic projection and observed camera image (as seen in Figure 3) (or some transformation of the image). The Ushichka dataset is fit to use the alignment methods mentioned above as the extrinsic (position and rotation) and intrinsic (eg. focal length, distortion) camera parameters are known for each night in the dataset. One potential issue that may arise is that the methods of Kang and Doh (2020) and Muñoz-Bañón, Candelas, and Torres (2020) have been proposed for visible light cameras. Thermal camera images often look very different from light camera images, and may require different processing steps. To our limited knowledge current thermal camera-LiDAR alignment protocols require a specifically designed calibration object (Krishnan and Saripalli 2017; Zhang et al. 2019, 2018). While it may be possible to use the target-less workflows for the *Ushichka* dataset, the lack of thermal dynamic range in images may present a challenge. In contrast to RGB images, thermal images are mono-channel and the cameras we used also have a limited thermal resolution (~1C). In addition to device limitations, the actual thermal environment in the cave may pose a challenge. The cave system maintains a relatively stable temperature of around 10C throughout the year. From a thermal perspective, the constrast between one part of the cave and another may not be as large as what is seen in many RGB images of a given view. Despite these challenges, one advantage *Uschichka* may have is the fact that each recording session has *three* (instead of one) camera views that can be used to align the LiDAR scan. In this respect, our three camera situation may result in a more constrained alignment despite the poor thermal gradient expected in the images.

### Automatic mic position estimation removes the need for bulky frames

Working with multi-microphone arrays typically includes carrying many long cables and heavy frames to hold microphones in known positions. Microphone positions need to be known for successful acoustic localisation, and precise mic positions lead to lower localisation errors (Wahlberg 1999). However, carrying large array frames into the field, especially in cluttered environments such as caves or forests is often difficult and time-consuming. In response to these conspicuous objects, bats may actually focus their calls specifically on the frames and the recorded behaviour might actually be an unwanted artifact of the setup itself. There is thus a pressing need in bioacoustics for the development of acoustic arrays that are less conspicuous, and can be setup universally and quickly. One option, as seen in *Ushichka*, is to place the microphones on existing structures such as cave walls and rocks (branches, signposts and other structures may also be useful in other field sites). Using existing structures to hold microphones does not generate new obstacles or points of interest, and is thus less likely to generate artifactual inspection behaviours from the animal.

Once frames have been abandoned in favour of existing structures, the task of measuring microphone positions still remains. The most accurate solution to directly measure microphone positions is to use a Total station theodolite. Total station’s have measurement acccuracies of a few mm, and directly output the 3D positions of microphones. A major issue with the use of a Total station is the addition of yet another piece of equipment, and the time taken to set up the device itself for measurements. Alternately, it is possible to estimate mic positions indirectly by exhaustively measuring inter-mic distances using a laser range finder or tape-measure (as has been done for *Ushichka*). However, experience has shown manual inter-mic measurements can be error-prone at larger distances (~4-5m) with a laser range finder due to jittery hands, and become impractical with tape-measures. Manual inter-mic distance measurements require direct line-of-sight between microphones and do not scale well with an increasing numbe of microphones in an array (number of measurements to be made scales rapidly with ).

Given the problems around microphone position measurements without frames, we were looking for solutions that have the ease of today’s camera calibration workflows (eg. Theriault et al. 2014), and found the ‘Structure-from-Sound’ approach of Zhayida et al. (2016). Zhayida et al. (2016) were able to infer microphone positions based only on commonly recorded playbacks, and without the need for any actual measurements. We were also able to verify the utility of the Structure-from-Sound approach for our freely-placed microphones on ground-truthed data (Batstone et al. 2019, also see Supplementary Information). Such automatic position estimation workflows have the potential to drastically reduce the effort needed to run multi-channel recordings in the field, and open up the exciting realm of inconspicuous experimental setups. We are currently continuing the collaboration and looking to formulate mic position estimation workflows that are field-friendly and widely applicable across field-site conditions.

## Investigative opportunities opened by the *Ushichka* dataset

Here we detail some of the possible lines of investigations that can be undertaken with *Ushichka* once the necessary analytical workflows are in place.

### Collective behaviour in active sensing groups

The combination of acoustic and video tracking along with a LiDAR scan of the cave provides us sufficient data to reconstruct the sensory inputs of each bat in a group. To reconstruct the sensory inputs of each bat, we would require baseline data on a series of parameters such as call and hearing directionality, auditory masking, hearing threshold and source level. These parameters can either be extracted from the literature (as done by Beleyur and Goerlitz (2019) and Mazar and Yovel (2020)) or measured using the tracking system itself (eg. source level and call directionality). Given a group of *N* bats, we can thus reconstruct all the sounds each bat would have heard in their inter-pulse intervals (the silent gap between calls). The sounds would include the focal bat’s own echoes, secondary echoes from other bats, and the calls of other bats. Previous studies have attempted to simulate such contexts with biologically parametrised values of the input sounds (Beleyur and Goerlitz 2019; Mazar and Yovel 2020). Our experimental measurements in combination with detailed sound propagation models could allow for *direct* sensory input reconstructions in a whole group of bats simultaneously. Knowing the sensory inputs animals receive over time would allow us to finely parametrise the relevant models of collective behaviour and test how well they predict group motion. For instance, Bode, Franks, and Wood (2011) model a group of agents that perform non-simultaneous and discontinuous ‘updates’ of their environment. Their model is very similar to how groups of bats get information about their surroundings, as each bat is very likely emitting calls independently (Hase 2018). Bode, Franks, and Wood (2011) moreover model ‘limited’ sensing where each agent can only detect the positions of a few neighbours at each update, very similar to what bats may be experiencing in large groups (Beleyur and Goerlitz 2019). The Bode, Franks, and Wood (2011) model can be parametrised by estimating how often bats call in a group (update rate), and how many neighbours they are typically able to detect (eg. using the framework in Beleyur and Goerlitz (2019)). The dynamics of *in silico* active sensing groups can be compared with observed bat groups. Sensory reconstruction approaches in echolocation would also open the field to fine-scale temporal analyses of echolocation and flight behaviour: do bats turn away more often in response to an echo from a wall, or from the direction of a calling neighbour? Do bats call more often when they may have ‘missed’ echo detections due to call overlaps in the past few interpulse intervals?

### Optimising echolocation and flight strategies under challenging conditions

Bats are known to alter their echolocation rapidly according to the behavioural context, and ambient soundscape (Corcoran and Moss 2017). *Ushichka* offers direct access to bat echolocation strategies as they encounter an increasing gradient of acoustical challenge as group size increases (Fig 6). Aside from call-level data such as source-level, duration, spectral properties - the presence of multiple microphones spread across the room also provides access to call direction and beam shape modulations (Figure 7A). While other call parameters have been investigated in studies with multiple animals, call directionality and beam-shape measurements have so far mainly been done with single animals in the field and lab (Surlykke, Pedersen, and Jakobsen 2012; Giuggioli, McKetterick, and Holderied 2015). Revealing where bats aim their calls, and how they spread the energy in each call will uncover the sensory priorities bats may have while flying under challenging conditions. Do bats choose to focus narrowly on fast moving conspecifics, but broadly onto bigger obstructions like walls. Does their call aim and beam shape change with increasing group size?

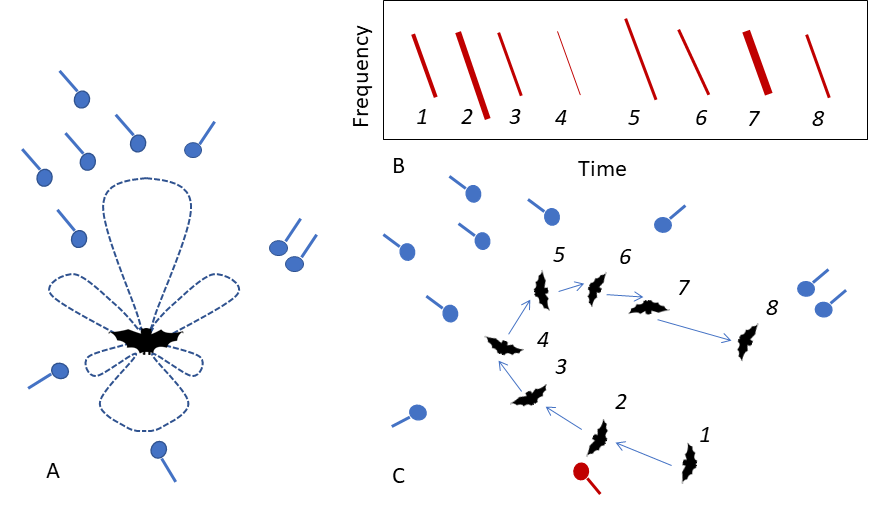


Figure 7: The benefits of multi-mic arrays for the study of echolocation and voice recognition (microphone: blue circle with line attached) A) Biosonar beam-shape can be estimated from microphones spread around the animals B,C) As the bat moves through the volume (C 1-8), the amplitude and spectral properties of the call received by the focal microphone (red) changes, as seen on the spectrogram (B). Having multiple microphones provides a multi-view picture of a call from many directions, allowing a better assessment of intra-individual call variation required for robust voice recognition

The echolocation strategies in use may also inform why certain types of apparently stereotypical flight paths are seen within the recording volume. It was observed that bats sometimes tend to fly a circular loop close to the edges of the recording volume before exiting out or heading back into the roosting site (Figure 8). This ‘looping’ flight behaviour was observed in the evening flight activity as bats began to accumulate in the recording volume, but also when bats returned alone in the morning. Bats are known to show stereotypic flight behaviours in the lab and in the wild, and this may perhaps be due to their limited attention or spatial memory (Barchi, Knowles, and Simmons 2013; Möhres and Oettingen-Spielberg 1949). However, in a cave setting, where there are no obvious obstacles, what drives the formation of these looping behaviours? Do the loops form a type of stereotypic trajectory that allows the bat to limit its attention only along that route? Why do some bats not take the shortest path and fly from the exit straight into roosting site? Initial observations of group flights suggested that even when multiple bats were flying together, not all parts of the recording volume were being ‘used’ by the bats. Some regions of the recording volume (regions close to walls) seemed to have a higher density of bats. Is this wall-clinging behaviour a behavioural adaptation to maintain a high received level of wall-echoes (in the presence of non-target sounds)? The wall-clinging behaviour is somewhat reminiscent of structure-following behaviour seen in the field with horsehoe bats, that follow hedgerows instead of taking shortcuts over open fields. Horseshoe bats however, are likely to follow structures due to the limited sensing range from their high-frequency echolocation - why would myotids also adopt this behaviour? Another interesting parameter to estimate is the dynamics of the loop’s chirality. Alignment is a standard measure of how similar animals are in their direction of movement. When there are multiple animals in the recording volume, a measure of the alignment over time and group size might reveal whether the looping behaviour is an individually driven behaviour or is modulated by the flight directions of group members.

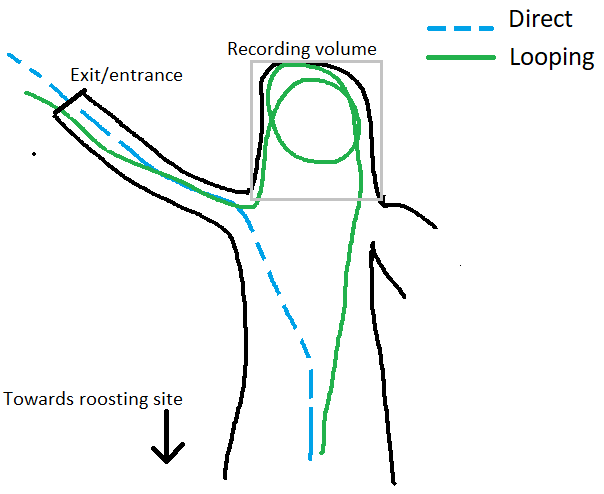


Figure 8: Schematic showing a view of the direct and looped flight paths as seen from above. The direct (broken blue line) and indirect (continuous green line) paths are flown in both directions (exit/entrance-cave interior, cave interior-exit/entrance). The grey quadrilateral outlines the recording volume.

### Voice recognition for short-term and long-term behavioural tracking

Many vocalising vertebrates across the animal kingdom have individually identifiable vocalisations, or voices (Carlson, Kelly, and Couzin 2020). The voice of each animal is a temporally stable cue, unlike experimental markings which may be lost over time. Despite the known stability of voice, one of the main challenges with developing a stable voice identifier is the variation in the vocalisation itself. Bird call repertoires can change over time, and complicate identification (Stowell et al. 2019). Approaches to date rely on machine-learning and statistical models, which require a training dataset of example vocalisations to optimise their predictive performance. Variation caused by sound directionality, distance and ambient noise may require larger training datasets, which are often hard to obtain (Stowell et al. 2019). Unlike many types of animal vocalisations (eg. bird song), the spectro-temporal structure of bat echolocation calls are stereotypical and likely to remain stable because they serve a single sensory purpose – to detect objects. The dedicated sensory functionality of echolocation calls, and the general absence of background ultrasonic noise in the wild makes bat calls a convenient system to develop voice identification workflows.

Bat echolocation calls can be assigned to individuals by their acoustic features, though this has typically been done in small lab populations. Previous work (Masters, Raver, and Kazial 1995; Yovel et al. 2009) solved the ‘closed-set’ problem where all animal identities are known. *Ushichka* provides a large dataset over multiple nights and animals that fly by [also see Figure 7B,C]. In contrast to the ‘closed-set’ problems solved in the past, *Ushichka* could be used to generate methods to solve the ‘open-set’ voice recognition task. Open-set classification problems deal with models that uniquely identify individuals that they have been trained on, but also those they have *not* been explicitly trained on. Even though the identities of bats echolocating in the recordings are unknown, it may be possible to exploit the behavioural patterns of the animals to approximate identity. Bats always return alone in the morning, and these recordings contain multiple echolocation calls from one individual. In addition, it is known that female *M. myotis/blythii* roost in the cave system over the entire summer period. Using these two facts, if successful, a voice recognition model should at least show certain patterns 1) correctly assign identity to an independent subset of calls it was not trained on (validation data) from a morning recording 2) identify at least some individuals across multiple morning returns, and in the best case 3) be able to pick up the individuals as they begin their activity in the recording volume around sunset. Of course, even if a trained model satisfies the three criteria above, it does not directly follow that the detections in scenarios 2) and 3) are not false positives. It may be necessary to systematically collect a larger sample of echolocation calls from hand-released individuals. The state of signal-processing and machine learning know-how in the field of human voice recognition is relatively mature at this point, with the use of speaker-specific Gaussian Mixture Models compared with Universal Background Models trained against multiple individuals (Hennebert 2009). *Ushichka* may provide a test-dataset to examine how well the latest voice recognition approaches perform for echolocation calls.

While seemingly challenging, if a validated voice recognition model is indeed developed successfully, it will allow revolutionary markerless short-term and long-term tracking of individual behaviours. Short-term behaviours such as flight and echolocation decisions over the course of a few seconds can provide insights into individual sensorimotor strategies and the tracking of social interactions. Long-term decisions such as the change of evening cave exit-time and morning return times can reveal the changing physiological status of the resident females as their pups mature. Voice based individual recognition brings the same advantages gained by tracking mammals with unique visible markings (eg. zebras, tigers) (Lahiri et al. 2011; Hiby et al. 2009), allowing unobtrusive observations over multiple time-points and scales.

## Supplementary code and website

The fieldrecorder\_trigger.py module used to trigger automatic recording is uploaded at the following repository: <https://github.com/thejasvibr/fieldrecorder>. Those interested in viewing more snippets of the the *Ushichka* dataset and the progress of research related to it, may wish to visit <https://thejasvibr.github.io/ushichka/>.

## Field work permits

All field work at the Orlova Chuka cave was done with permission from the relevant authorities. *Ushichka* is a purely observational dataset, no animals were handled or subjected to experimental treatments of any sort over the course of data collection.

## Author Contributions

TB: conception, experiment design, field data collection. HRG: conception, experiment and recording-system design.

## Acknowledgements

We would like to thank Antoniya Hubancheva, for her constant help and support in the collection of the *Ushichka* dataset (also for christening the dataset with its Bulgarian name) and the Tabachka field crew of 2018. The electronics (Markus Abels, Hannes Sagunsky, Reinhard Biller) and Feinmechanik (Erich Koch, Felix Hartl, Klaus Pichler) teams were of invaluable help and support all through the process of experimental setup design and troubleshooting. We would like to thank Pranav Khandelwal for extensive help setting up the thermal camera calibration workflow and Hedrick Tyson for providing helpful feedback and estimating camera parameters. Fieldwork-wise, we are grateful to Joanna Furmankiewicz for supporting initial recordings at the Jaskinia Niedźwiedzia and Diane Theriault for providing advice on thermal cameras. Special thanks and acknowledgement to the field assistance of Aditya Krishna and Neetash MR, without whom the data collection would not have been possible. T.B. was funded by a doctoral fellowship from the German Academic Exchange Service (DAAD) and the International Max Planck Research School for Organismal Biology. H.R.G. was funded by the Emmy Noether program of the German Research Foundation (DFG, GO 2091/2-1, GO 2091/2-2). We would like to thank the members of the Acoustic and Functional Ecology groups and the Max Planck Institute for Ornithology, Seewiesen for its support and infrastructure.

## References

Ballerini, Michele, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Vivien Lecomte, et al. 2008. “Interaction Ruling Animal Collective Behavior Depends on Topological Rather Than Metric Distance: Evidence from a Field Study.” *Proceedings of the National Academy of Sciences* 105 (4): 1232–7.

Barchi, Jonathan R, Jeffrey M Knowles, and James A Simmons. 2013. “Spatial Memory and Stereotypy of Flight Paths by Big Brown Bats in Cluttered Surroundings.” *Journal of Experimental Biology* 216 (6): 1053–63.

Batstone, K., G. Flood, T. Beleyur, V. Larsson, H. R. Goerlitz, M. Oskarsson, and K. Åström. 2019. “Robust Self-Calibration of Constant Offset Time-Difference-of-Arrival.” In *ICASSP 2019 - 2019 Ieee International Conference on Acoustics, Speech and Signal Processing (Icassp)*, 4410–4.

Beleyur, Thejasvi, and Holger R Goerlitz. 2019. “Modeling Active Sensing Reveals Echo Detection Even in Large Groups of Bats.” *Proceedings of the National Academy of Sciences* 116 (52): 26662–8.

Betke, Margrit, Diane E Hirsh, Angshuman Bagchi, Nickolay I Hristov, Nicholas C Makris, and Thomas H Kunz. 2007. “Tracking Large Variable Numbers of Objects in Clutter.” In *2007 Ieee Conference on Computer Vision and Pattern Recognition*, 1–8. IEEE.

Bode, N. W. F. F., D. W. Franks, and A. J. Wood. 2011. “Limited Interactions in Flocks: Relating Model Simulations to Empirical Data.” *J. R. Soc. Interface* 8: 301–4.

Brandstein, Michael, and Darren (Eds.) Ward. 2013. *Microphone Arrays: Signal Processing Techniques and Applications*. Springer Science & Business Media.

Carlson, Nora V, E McKenna Kelly, and Iain Couzin. 2020. “Individual Vocal Recognition Across Taxa: A Review of the Literature and a Look into the Future.” *Philosophical Transactions of the Royal Society B* 375 (1802): 20190479.

Corcoran, Aaron J, and Cynthia F Moss. 2017. “Sensing in a Noisy World: Lessons from Auditory Specialists, Echolocating Bats.” *Journal of Experimental Biology* 220 (24): 4554–66.

Couzin, I. D., J. Krause, R. James, G. D. Ruxton, and N. R. Franks. 2002. “Collective Memory and Spatial Sorting in Animal Groups.” *J. Theor. Biol* 218: 1–11.

Cvikel, Noam, Eran Levin, Edward Hurme, Ivailo Borissov, Arjan Boonman, Eran Amichai, and Yossi Yovel. 2015. “On-Board Recordings Reveal No Jamming Avoidance in Wild Bats.” *Proceedings of the Royal Society B: Biological Sciences* 282 (1798): 20142274.

Dijkgraaf, Sven. 1946. “Die Sinneswelt der Fledermäuse.” *Experientia* 2 (11): 438–48. <https://doi.org/10.1007/BF02153592>.

Erkert, Hans G., M Brock Fenton, James S. Findley, Theodore H. Fleming, Raymond E. Heithaus, Thomas H. Kunz, Adrian G. Marshall, et al. 1982. *Ecology of Bats*. <https://doi.org/10.1007/978-1-4613-3421-7>.

Fawcett, Kayleigh, David S Jacobs, Annemarie Surlykke, and John M Ratcliffe. 2015. “Echolocation in the Bat, Rhinolophus Capensis: The Influence of Clutter, Conspecifics and Prey on Call Design and Intensity.” *Biology Open* 4 (6): 693–701.

Fawcett, Kayleigh, and John M Ratcliffe. 2015. “Clutter and Conspecifics: A Comparison of Their Influence on Echolocation and Flight Behaviour in Daubenton’s Bat, Myotis Daubentonii.” *Journal of Comparative Physiology A* 201 (3): 295–304.

Fenton, M Brock, Alan D Grinnell, Arthur N Popper, and Richard R Fay Editors. 2016. *Bat Bioacoustics*. Edited by Brock M Fenton, Alan D Grinnell, Arthur N Popper, and Richard R Fay. Vol. 54. New York: Springer Science+Business Media k. <https://doi.org/10.1007/978-1-4939-3527-7>.

Geier, M. 2015. <https://python-sounddevice.readthedocs.io/en/0.3.5/_modules/sounddevice.html.>

Gillam, Erin H, Nickolay I Hristov, Thomas H Kunz, and Gary F McCracken. 2010. “Echolocation Behavior of Brazilian Free-Tailed Bats During Dense Emergence Flights.” *Journal of Mammalogy* 91 (4): 967–75.

Giuggioli, L., T. J. McKetterick, and M. Holderied. 2015. “Delayed Response and Biosonar Perception Explain Movement Coordination in Trawling Bats.” *PLoS Comput Biol* 11 (3).

Goerlitz, Holger R. 2018. “Weather Conditions Determine Attenuation and Speed of Sound: Environmental Limitations for Monitoring and Analyzing Bat Echolocation.” *Ecology and Evolution* 8 (10): 5090–5100.

Goerlitz, Holger R., Hannah M. Ter Hofstede, Matt R. K. Zeale, Gareth Jones, and Marc W. Holderied. 2010. “An aerial-hawking bat uses stealth echolocation to counter moth hearing.” *Current Biology* 20 (17): 1568–72. <https://doi.org/10.1016/j.cub.2010.07.046>.

Götze, S., J. C. Koblitz, A. Denzinger, and H. U. Schnitzler. 2016. “No Evidence for Spectral Jamming Avoidance in Echolocation Behavior of Foraging Pipistrelle Bats.” *Sci Rep* 6 (July).

Griffin, Donald R. 1958. “Listening in the dark: the acoustic orientation of bats and men.”

Habersetzer, J. 1981. “Adaptive Echolocation Sounds in the Bat Rhinopoma Hardwickei.” *J. Comp. Physiol* 144: 559–66.

Hase, K. 2018. “Bats Enhance Their Call Identities to Solve the Cocktail Party Problem.” *Commun Biol* 1: 39.

Hennebert, Jean. 2009. “Speaker Recognition, Overview.” In *Encyclopedia of Biometrics*, edited by Stan Z. Li and Anil Jain, 1262–70. Boston, MA: Springer US. <https://doi.org/10.1007/978-0-387-73003-5_198>.

Hiby, Lex, Phil Lovell, Narendra Patil, N Samba Kumar, Arjun M Gopalaswamy, and K Ullas Karanth. 2009. “A Tiger Cannot Change Its Stripes: Using a Three-Dimensional Model to Match Images of Living Tigers and Tiger Skins.” *Biology Letters* 5 (3): 383–86.

Hügel, Theresa, Vincent van Meir, Amanda Muñoz-Meneses, B. Markus Clarin, Björn M. Siemers, and Holger R. Goerlitz. 2017. “Does similarity in call structure or foraging ecology explain interspecific information transfer in wild Myotis bats?” *Behavioral Ecology and Sociobiology* 71 (11). <https://doi.org/10.1007/s00265-017-2398-x>.

Ivanova, Teodora. 2005. “Important Bat Underground Habitats (IBUH) in Bulgaria.” *Acta Zoologica Bulgarica* 57 (2): 197–206.

Izadi, Mohammad Rasool, Robert Stevenson, and Laura N Kloepper. 2020. “Separation of Overlapping Sources in Bioacoustic Mixtures.” *The Journal of the Acoustical Society of America* 147 (3): 1688–96.

Jarvis, Jenna, William Jackson, and Michael Smotherman. 2013. “Groups of Bats Improve Sonar Efficiency Through Mutual Suppression of Pulse Emissions.” *Frontiers in Physiology* 4: 140.

Kamburov, Asparuh, Holger R Goerlitz, and Thejasvi Beleyur. 2018. “Geospatial Modelling Inside the "Orlova Chuka" Cave in Bulgaria.” *XXVIII International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*. <http://symp2018.geodesy-union.org/wp-content/uploads/2018/11/20.pdf>.

Kang, Jaehyeon, and Nakju L Doh. 2020. “Automatic Targetless Camera–Lidar Calibration by Aligning Edge with Gaussian Mixture Model.” *Journal of Field Robotics* 37 (1): 158–79.

Krishnan, Aravindhan K, and Srikanth Saripalli. 2017. “Cross-Calibration of Rgb and Thermal Cameras with a Lidar for Rgb-Depth-Thermal Mapping.” *Unmanned Systems* 5 (02): 59–78.

Lahiri, M., C. Tantipathananandh, R. Warungu, D. I. Rubenstein, and T. Y. Berger-Wolf. 2011. “Biometric Animal Databases from Field Photographs: Identification of Individual Zebra in the Wild.” In *Proceedings of the Acm International Conference on Multimedia Retrieval (Icmr 2011)*. ACM Press.

Lawrence, Beatrice D, and James A Simmons. 1982. “Measurements of Atmospheric Attenuation at Ultrasonic Frequencies and the Significance for Echolocation by Bats.” *The Journal of the Acoustical Society of America* 71 (3): 585–90.

Lewanzik, Daniel, and Holger R. Goerlitz. 2018. “Continued source level reduction during attack in the low-amplitude bat Barbastella barbastellus prevents moth evasive flight.” *Functional Ecology* 32 (5): 1251–61. <https://doi.org/10.1111/1365-2435.13073>.

Lin, Yuan, Nicole Abaid, and Rolf Müller. 2016. “Bats Adjust Their Pulse Emission Rates with Swarm Size in the Field.” *The Journal of the Acoustical Society of America* 140 (6): 4318–25.

Masters, W Mitch, Kelley AS Raver, and Karry A Kazial. 1995. “Sonar Signals of Big Brown Bats, Eptesicus Fuscus, Contain Information About Individual Identity, Age and Family Affiliation.” *Animal Behaviour* 50 (5): 1243–60.

Mazar, Omer, and Yossi Yovel. 2020. “A Sensorimotor Model Shows Why a Spectral Jamming Avoidance Response Does Not Help Bats Deal with Jamming.” *Elife* 9: e55539.

Moss, Cynthia F., and Annemarie Surlykke. 2001. “Auditory scene analysis by echolocation in bats.” *The Journal of the Acoustical Society of America* 110 (4): 2207–26. <https://doi.org/10.1121/1.1398051>.

Möhres, FP, and Th Oettingen-Spielberg. 1949. “Versuche über Die Nahorientierung Und Das Heimfindevermögen Der Fledermäuse.” *Verhandlungen Der Deutschen Zoologen in Mainz*, 248–52.

Muñoz-Bañón, M. Á., F. A. Candelas, and F. Torres. 2020. “Targetless Camera-Lidar Calibration in Unstructured Environments.” *IEEE Access* 8: 143692–143705. <https://doi.org/10.1109/ACCESS.2020.3014121>.

“Natura 2000 Bulgarian Ministry of Environment and Water Report.” 2013. <http://natura2000.moew.government.bg/Home/ProtectedSite?code=BG0000608&siteType=HabitatDirective>.

Necknig, Veronika, and Andreas Zahn. 2011. “Between-Species Jamming Avoidance in Pipistrelles?” *Journal of Comparative Physiology A* 197 (5): 469–73.

Nelson, M. E., and M. A. MacIver. 2006. “Sensory Acquisition in Active Sensing Systems.” *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol* 192: 573–86.

Oliphant, T. E. 2006. *A Guide to Numpy*. Trelgol Publishing USA.

Ortega, Jorge. 2016. *Sociality in bats*. Edited by Jorge Ortega. Springer International Publishing Switzerland. <https://doi.org/10.1007/978-3-319-38953-0>.

Pita, Diana, Bertrand Collignon, José Halloy, and Esteban Fernández-Juricic. 2016. “Collective Behaviour in Vertebrates: A Sensory Perspective.” *Royal Society Open Science* 3 (11): 160377.

Popper, Arthur N, and Richard R Fay. 1995. *Hearing by Bats*. Springer.

Siemers, Björn M, and Gerald Kerth. 2006. “Do Echolocation Calls of Wild Colony-Living Bechstein’s Bats (Myotis Bechsteinii) Provide Individual-Specific Signatures?” *Behavioral Ecology and Sociobiology* 59 (3): 443–54.

Stowell, Dan, Tereza Petrusková, Martin Šálek, and Pavel Linhart. 2019. “Automatic Acoustic Identification of Individuals in Multiple Species: Improving Identification Across Recording Conditions.” *Journal of the Royal Society Interface* 16 (153): 20180940.

Strandburg-Peshkin, Ariana, Colin R Twomey, Nikolai WF Bode, Albert B Kao, Yael Katz, Christos C Ioannou, Sara B Rosenthal, et al. 2013. “Visual Sensory Networks and Effective Information Transfer in Animal Groups.” *Current Biology* 23 (17): R709–R711.

Sumpter, David JT. 2006. “The Principles of Collective Animal Behaviour.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 361 (1465): 5–22.

Surlykke, A., S. B. Pedersen, and L. Jakobsen. 2012. “Echolocating Bats Emit a Highly Directional Sonar Sound Beam in the Field.” *Proc R Soc B Biol Sci* 276 (1658).

Theriault, D. H., Z Wu, N. I. Hristov, S. M. Swartz, K. S. Breuer, T. H. Kunz, and M. Betke. 2010. *Reconstruction and Analysis of 3D Trajectories of Brazilian Free-Tailed Bats in Flight*. Bost Univ Comput Sci Technical Rep.

Theriault, Diane H., Nathan W. Fuller, Brandon E. Jackson, Evan Bluhm, Dennis Evangelista, Zheng Wu, Margrit Betke, and Tyson L. Hedrick. 2014. “A protocol and calibration method for accurate multi-camera field videography.” *Journal of Experimental Biology* 217 (11): 1843–8. <https://doi.org/10.1242/jeb.100529>.

Ulanovsky, Nachum, M Brock Fenton, Asaf Tsoar, and Carmi Korine. 2004. “Dynamics of Jamming Avoidance in Echolocating Bats.” *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271 (1547): 1467–75.

Ulanovsky, Nachum, and Cynthia F Moss. 2008. “What the Bat’s Voice Tells the Bat’s Brain.” *Proceedings of the National Academy of Sciences* 105 (25): 8491–8.

Van Rossum, Guido, and Fred L Drake. 2011. *The Python Language Reference Manual*. Network Theory Ltd.

Virtanen, P. 2019. “SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python.”

Wahlberg, Magnus. 1999. “Positioning accuracy of a large‐aperture hydrophone array for sperm whale research.” *The Journal of the Acoustical Society of America* 105 (2): 1318–8. <https://doi.org/10.1121/1.424536>.

Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.org/knitr/>.

Yovel, Yossi, Mariana Laura Melcon, Matthias O Franz, Annette Denzinger, and Hans-Ulrich Schnitzler. 2009. “The Voice of Bats: How Greater Mouse-Eared Bats Recognize Individuals Based on Their Echolocation Calls.” *PLoS Comput Biol* 5 (6): e1000400.

Zhang, J., P. Siritanawan, Y. Yue, C. Yang, M. Wen, and D. Wang. 2018. “A Two-Step Method for Extrinsic Calibration Between a Sparse 3D Lidar and a Thermal Camera.” In *2018 15th International Conference on Control, Automation, Robotics and Vision (Icarcv)*, 1039–44. <https://doi.org/10.1109/ICARCV.2018.8581170>.

Zhang, J., R. Zhang, Y. Yue, C. Yang, M. Wen, and D. Wang. 2019. “SLAT-Calib: Extrinsic Calibration Between a Sparse 3D Lidar and a Limited-Fov Low-Resolution Thermal Camera.” In *2019 Ieee International Conference on Robotics and Biomimetics (Robio)*, 648–53. <https://doi.org/10.1109/ROBIO49542.2019.8961462>.

Zhayida, Simayijiang, Simon Segerblom Rex, Yubin Kuang, Fredrik Andersson, and Kalle Åström. 2016. “An Automatic System for Acoustic Microphone Geometry Calibration Based on Minimal Solvers.” *CoRR* abs/1610.02392. <http://arxiv.org/abs/1610.02392>.