



The potential for acoustic individual identification in mammals

Pavel Linhart¹ · Mathieu Mahamoud-Issa² · Dan Stowell^{3,4} · Daniel T. Blumstein⁵

Received: 30 November 2020 / Accepted: 27 December 2021

© The Author(s) under exclusive licence to Deutsche Gesellschaft für Säugetierkunde 2022

Abstract

Many studies have revealed that animal vocalizations, including those from mammals, are individually distinctive. Therefore, acoustic identification of individuals (AIID) has been repeatedly suggested as a non-invasive and labor efficient alternative to mark-recapture identification methods. We present a pipeline of steps for successful AIID in a given species. By conducting such work, we will also improve our understanding of identity signals in general. Strong and stable acoustic signatures are necessary for successful AIID. We reviewed studies of individual variation in mammalian vocalizations as well as pilot studies using acoustic identification to census mammals and birds. We found the greatest potential for AIID (characterized by strong and stable acoustic signatures) was in Cetacea and Primates (including humans). In species with weaker acoustic signatures, AIID could still be a valuable tool once its limitations are fully acknowledged. A major obstacle for widespread utilization of AIID is the absence of tools integrating all AIID subtasks within a single package. Automation of AIID could be achieved with the use of advanced machine learning techniques inspired by those used in human speaker recognition or tailored to specific challenges of animal AIID. Unfortunately, further progress in this area is currently hindered by the lack of appropriate publicly available datasets. However, we believe that after overcoming the issues outlined above, AIID can quickly become a widespread and valuable tool in field research and conservation of mammals and other animals.

Keywords Acoustic communication · Acoustic individual identification (AIID) · Acoustic signature · Individual identity · Individual variation · Machine learning · Mammal · Vocalizations

Handling editors: Leszek Karczmarski and Daniel I. Rubenstein.

This article is a contribution to the special issue on “Individual Identification and Photographic Techniques in Mammalian Ecological and Behavioural Research – Part 1: Methods and Concepts” — Editors: Leszek Karczmarski, Stephen C.Y. Chan, Daniel I. Rubenstein, Scott Y.S. Chui and Elissa Z. Cameron.

✉ Pavel Linhart
pavel.linhart83@gmail.com

¹ Zoology Department, Faculty of Science, University of South Bohemia, Ceske Budejovice, Czechia

² Department of Behavioural Ecology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

³ Tilburg University, Tilburg, The Netherlands

⁴ Naturalis Biodiversity Center, Leiden, The Netherlands

⁵ Department of Ecology and Evolutionary Biology, University of California, Los Angeles, USA

Introduction

For many studies, it is important to be able to identify individual animals, because by doing so, we can gain detailed insights about their ecology and behavior. Animals themselves often need to recognize conspecifics individually as well. Such individual recognition is a necessary requirement for complex social interactions and we see it being used in a variety of contexts including territorial behavior, parent–offspring interactions, mate choice, the allocation of potentially altruistic behavior, etc. For wildlife conservationists, knowledge of individual identity can be used to help calculate precise population estimates (which are essential to provide robust estimates of population sustainability), to collect data about survival (which can be used to estimate trends and identify causes of mortality), and to quantify individual movement (which represents ranging and resource use). Thus, for all these reasons, identifying individuals and following their fate and behavior may assist in the conservation and management of species and populations (Terry et al. 2005; Pimm et al. 2015).

Among free-ranging mammals, a certain proportion of individuals are visually distinctive and can be identified based solely on morphological characteristics (e.g., Würsig and Würsig 1977; Karanth and Nichols 1998) and successfully re-identified across years (Karczmarski et al. 2022a, b). In other cases, however, physical marking of individuals is needed and has traditionally been achieved by capturing and marking them with tags, radio tags, rings, pelage marks, PIT tags, etc. Marking facilitates the ability to identify individuals over time, either when re-sighted or re-captured. However, field techniques that involve physical capture of animals are invasive and pose a risk to the animals captured, depending on the method of trapping and species involved (Powell and Proulx 2003; Iossa et al. 2007; Cunningham et al. 2015; Soulsbury et al. 2020), and for threatened or endangered species, such risks may be considered extreme. Furthermore, in cases where capturing individuals might change their behavior or otherwise negatively impact research results (Linhart et al. 2012; Kukulová et al. 2013; Byers et al. 2019), other field techniques such as photographic (Karczmarski et al. 2022a, b) or acoustic (discussed further) should be considered.

Many species produce individually distinctive vocalizations and these vocalizations can be used to identify, census, and track individuals (Terry and McGregor 2002; Terry et al. 2005). Acoustic individual identification (AIID) is a non-invasive technique that identifies individuals using their vocalizations. It is, therefore, especially suitable for individual monitoring of endangered species or those vulnerable to capturing and handling to minimize risks associated with capture–recapture identification techniques.

AIID requires species that are vocally active. Many cryptic species (e.g., canopy-dwelling birds and mammals, nocturnal animals, and species that are hard to visually detect) should likely rely on communication through acoustic signals. While acoustic signaling evolved in nocturnal vertebrates (Chen and Wiens 2020), it is not limited to them and it is a widely used modality for signaling in both terrestrial and aquatic habitats, because it has the ability for signals to be localizable, carry relevant information, and do so over a large area.

AIID requires a species to produce individually distinctive vocalizations. Mammalian vocal production can be modeled using source–filter theory whereby air is pushed through vocal folds, which causes them to vibrate and produce a sound (the ‘source’). The structure of the sound is then modified as it travels through the rest of the vocal tract (the ‘filter’). Individual morphological differences of the vocal tract translate into individually distinctive sounds (Taylor and Reby 2010). In addition to vocal tract-driven individual variation, individual differences can also be represented by any spectro-temporal patterns that an individual can reliably produce. For instance, dolphin signature whistles (Janik and

Sayigh 2013; Kershenbaum et al. 2013), and the individually specific “rhythm” of elephant seals grunts (Mathevon et al. 2017) both illustrate acoustic individuality not driven by vocal tract morphological variation. Such diverse acoustic signatures might result from cultural processes such as innovation learning as seen with dolphin signature whistles (Tyack 1997; Janik and Slater 2000). Once produced, adaptive processes shape the individual identity information in vocalizations (Wilkinson 2003; Pollard and Blumstein 2011; Charrier 2020). Many studies have documented individually specific vocalizations in mammals suggesting that these species could potentially be candidate species for individual acoustic identification.

Several pilot studies investigating the potential of AIID show that AIID can be effectively used to follow individuals over time and provide information about survival, territorial dynamics and replacements, and migration (Laiolo et al. 2007; Adi et al. 2010; Kirschel et al. 2011; Petrusková et al. 2016). AIID may provide a cost-effective solution for a particular task or permit the collection of unique data unavailable with more traditional mark-recapture identification methods (Laiolo et al. 2007; Vögeli et al. 2008; Petrusková et al. 2016). However, and importantly, these previously mentioned case studies of AIID were conducted with birds. We are aware of just a single case where AIID has been used to acquire information about population characteristics of wild non-human mammals (Longden et al. 2020).

Despite many mammals having individually distinctive vocalizations (Supplementary Material 1), there is remarkably limited work using acoustic individual signatures to study mammalian behavior and population biology. Terry et al. (2005) previously noted that most previous applications have focused on birds. The fact that this has not changed in a decade and a half is notable because, researchers frequently use various acoustic monitoring techniques to collect non-individualized data about mammals to study their behavioral activities in time and space (Heinicke et al. 2015; Spillmann et al. 2015; Kalan et al. 2015, 2016; Wrege et al. 2017; Kershenbaum et al. 2019). Matching such data with information about individual identity of callers would clearly be an important step forward. There are several potential reasons why previous studies of mammals have not capitalized on the widespread individual variation contained in their vocalizations.

Terry et al. (2005) suggested that the amount of within-individual variation in mammalian vocalizations might be greater than in birds making mammals potentially less suitable for individual acoustic monitoring which requires low within-individual variation for optimum results. Acoustic communication evolved independently in different vertebrate lineages as documented by comparative analyses on acoustic signals, vocal tracts and auditory systems, and their underlying neural systems (Clack 2002; Senter 2008;

Albersheim-Carter et al. 2016; Kingsley et al. 2018; Riede et al. 2019; Barkan and Zornik 2020; Chen and Wiens 2020), and, therefore, it is possible there are yet unrecognized fundamental differences in the acoustic communication systems of birds and mammals making acoustic identification in mammals difficult.

However, our recent work has shown that vocalizations of birds and mammals exhibit comparable amounts of individual identity information in their vocalizations (Linhart et al. unpublished data). Indeed, AIID has been used with great success in a mammalian species with very complex vocalizations—humans (Hansen and Hasan 2015). Accordingly, human listeners can identify individuals based on their vocalizations with high accuracy (Lavan et al. 2019), which has been shown for many other mammalian species as well (Wiley 2013). Vocal production is based on similar principles across the whole class of mammals (Taylor and Reby 2010) and, therefore, with the right approach, AIID should be feasible in mammals as well.

We review the potential for acoustic individual identification in mammals based on studies reporting individual variation in mammalian vocalizations. We also list a few pilot studies that successfully used AIID to collect new data about a population in a setting resembling a traditional mark-recapture study. In such studies, AIID has been put to the ultimate test, and from these studies, we can learn about the realistic conditions (signature strength, signature stability) required for the successful application of AIID. To be comprehensive, we sought analogous information from the avian literature where concepts and applications regarding AIID are relatively more developed. Many aspects of the process of AIID were nicely reviewed in Terry et al. (2005). Here, we offer a slightly different perspective and we map progress since the publication of their seminal paper. We provide a pipeline of the steps necessary for using AIID for a given species. We discuss practical and theoretical issues associated with each step. We also describe some potential limitations of AIID along with drawbacks currently hindering its more widespread use. Finally, we discuss the potential for its future development.

Acoustic individual identification pipeline

In general, there are three basic steps forming the process of individual identification: collection of a signal (recording), analysis of the signal (extraction of individually distinct features, feature extraction), and evaluation of the signal (classification of the signal to an individual emitting the signal based on the extracted features). These steps are repeated at least twice when AIID is used to follow individuals within the population in a way similar to traditional mark-recapture study. We also add three additional steps to our

pipeline. These two steps—assessment of signature strength and assessment of its stability—can inform potential AIID users early on during pilot studies about whether their species and vocalization under study are suitable for AIID or not. Finally, we must validate AIID whereby the effect of mistakes in AIID process on the population estimates are investigated (Fig. 1).

Step 1: Recording

Technically, the equipment and recording techniques used for recording vocalizations may differ in aquatic and terrestrial environments (Blumstein et al. 2011; Marques et al. 2013). We do not aim to provide a review on technical issues associated with sound recording. Rather, we focus on conceptual aspects of recording mammal vocalizations that have important implications for AIID.

From the perspective of AIID, vocalizations can be recorded in two conceptually different ways. Traditionally, animals are recorded using handheld microphones and (these days) digital recorders. In this situation, the person who is recording can also collect rich information about the vocalizing animal (e.g., its position, distance from recorder, context of the vocalization, and the number of and identities of other vocalizing individuals) that may inform further identification process. This is called focal recording (it is also referred to as manual, or target recording).

Alternatively, passive recording, where single or multiple autonomous recording units (ARUs) are placed within the animal's environment, gained popularity and rapid development in recent years (Blumstein et al. 2011; Marques et al. 2013; Gibb et al. 2019). Nowadays, users may select from a variety of ARUs ranging from fully featured commercial units to relatively inexpensive open-source solutions (Gibb et al. 2019; Rhinehart et al. 2020). The identification of individuals from passive recordings is a relatively more challenging task, because of the lack of background information about the context of the recordings (except, of course, information about the location and time of a recording). Nevertheless, the possibility of identifying individuals in passive recordings would open valuable new opportunities for applications of passive recording in ecology and conservation.

Focal recording typically provides higher quality audio samples due to targeted effort of the recording person and better equipment (a directional shotgun or parabolic microphone). This is important, because fine acoustic structural details may be required for individual identification. Yet, some individually distinct acoustic features have been found to be robust in signal degradation studies (Blumstein and Munos 2005; Lameira and Wich 2008; Mouterde et al. 2014). It is also possible to obtain high-quality recordings with ARUs. For example, recorders can be deployed close to favorite spots animals use for vocalizing (nests, burrows,

Acoustic individual identification (AIID) in a capture-recapture-like study

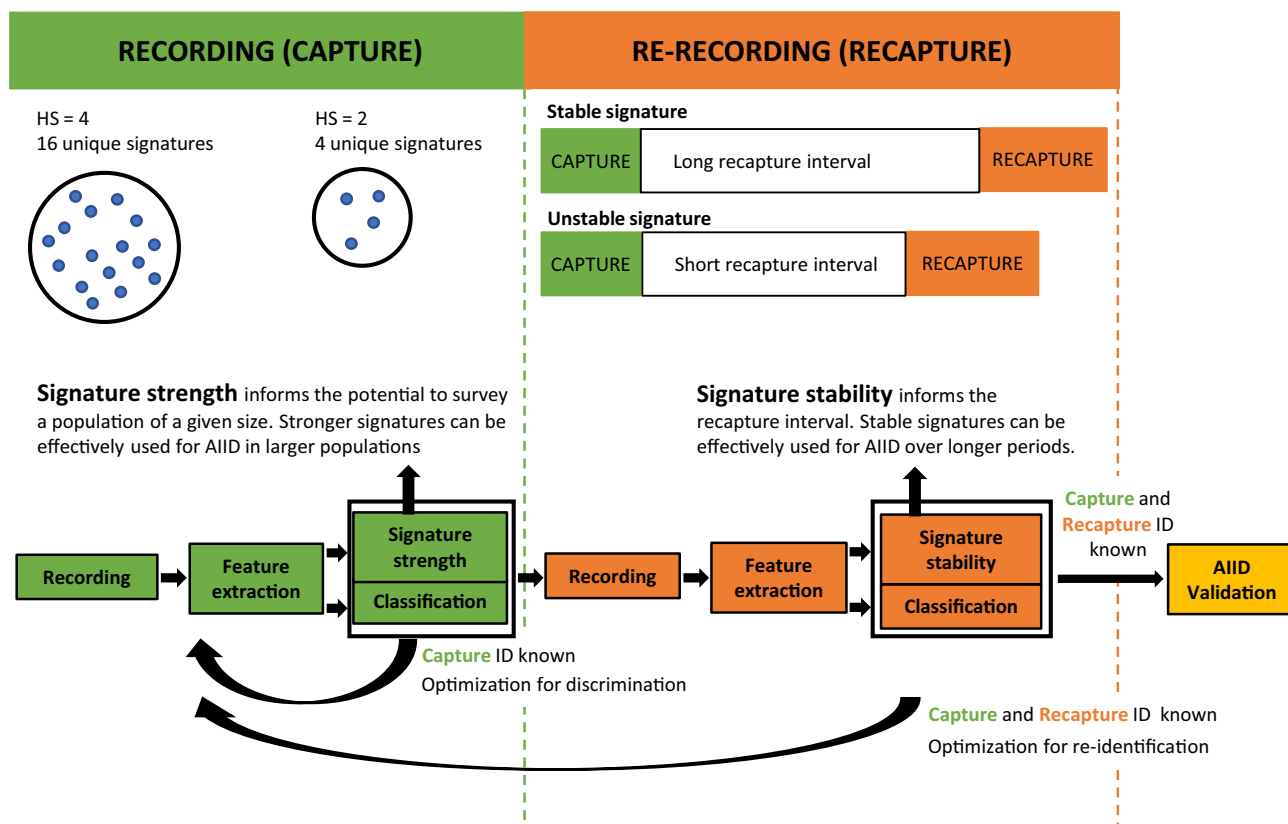


Fig. 1 Flowchart of AIID steps in a capture–recapture like study. The first recording of an individual represents CAPTURE and re-recording of that individual represents RECAPTURE in the terminology of capture–recapture studies. Evaluation of the signature strength allows estimating the size of a population that is surveyable with AIID. Stronger signatures permit more reliable surveys in larger populations. Assessment of signature stability is important to estimate the appropriate recapture interval and helps determine the temporal limits of AIID. Stable signatures allow for reliable AIID of individuals with longer intervals between recording and re-recording (capture and recapture) of the individual than unstable signatures. Knowing

the ID of the individuals permits the inclusion of optimization loops to identify the best sets of features and classification approaches for discrimination between individuals and long-term re-identification of individuals. AIID validation (i.e., how mistakes in AIID influence population estimates based on capture–recapture data) is only possible if real and estimated ID can be verified in both the capture and recapture stage. Optimized methods from pilot studies could be later applied to larger scale studies or to new conditions (new individuals appearing within the population, new populations, another species) where individual identities do not need to be always known

etc.). High-quality recordings could also be obtained by deployment of large numbers of ARUs making the grid of ARUs thicker. And, for ARUs with multiple microphones, beam-forming (Chen et al. 2003) can be used to increase the signal to noise ratio of vocalizations and therefore improve their quality for subsequent analyses. Playback can be used to elicit vocalizations and increase the quantity and quality of recorded material for both focal as well as passive recording (Plumptre et al. 2013; Hansen et al. 2015). In that case, it is necessary to ensure that potential behavioral changes caused by the playback do not affect expression of individual identity.

The crucial aspect of the AIID pipeline is whether the identity of recorded animals is known or unknown in the

recording stage. Ultimately, we want AIID to be able to be used in a population of completely unknown individuals. This, however, is difficult. Knowing the identity of recorded animals in capture as well as recapture phases of the AIID process increases the ability to better optimize and ultimately better evaluate AIID (Fig. 1) and this knowledge can help design studies where identity of individuals is entirely unknown. Hence, we recommend initial pilot studies are conducted on a sample of known individuals (captive animals, marked individuals). It is also important to include any background knowledge about the recordings that can help estimate individuality or provide at least partial validation of AIID results (territory, number of individuals within the

surveyed area or group, movement patterns, temporal scale of recordings, etc.)

Step 2: Feature extraction

Commercial and open-source software alternatives are available for the analysis of the recorded sounds (https://en.wikipedia.org/wiki/List_of_Bioacoustics_Software). However, there is no dedicated, off-the-shelf software solution for acoustic individual identification. Such special software solutions already exist for visual individual identification based on individual visual patterns (Tienhoven et al. 2007; Bolger et al. 2012; Crall et al. 2013), but they still need to be developed for AIID. The ideal software must integrate subtasks that are already available and achievable within various existing software packages: detection and segmentation of species calls in the recording, extraction of relevant features from the calls, and assigning calls or their sequence to a particular individual (Blumstein et al. 2011; Marques et al. 2013; McIoughlin et al. 2019). However, an integrated solution of these steps is currently lacking, which means that users must develop their own analysis pipelines from scratch. This will ultimately lead to a variety of unique and non-compatible solutions. In many of the existing software tools, clustering/classification can be used for species detection or classification, and these could be likely re-used to achieve AIID, at least in cases of highly pronounced individual distinctiveness, using individuals, instead of species, as the classification classes.

Sometimes, animals may have dedicated calls to signal their individuality, or individuality may be expressed across different calls from the animal's repertoire. Individuality can be found in many different vocal traits (features) within the spectral, temporal, or amplitude domains, depending on the likely origin of acoustic signature in a given species. Some of these individually distinct features may have their origin in vocal tract morphology involving both vocal source and/or vocal filter. Such morphology-related vocal traits are expected to be present across different vocalization types. Vocal source-related individuality traits often include the fundamental frequency F0 (e.g., Root-Gutteridge et al. 2014a). Vocal filter-related traits translate into variation in the frequency spectrum profile of vocalizations and the occurrence of formants (e.g., Briefer and McElligott 2012). Temporal patterns of vocalizations including duration, intervals, or spacing of vocal elements (rhythm) can also be individually specific (e.g., Mathevon et al. 2017). The amplitude domain is often ignored as a potential source of individual distinctiveness, because amplitude can be heavily degraded as a sound propagates through the environment. Nevertheless, amplitude-related traits are also individually specific in some cases (e.g., Root-Gutteridge et al. 2014b).

Spectro-temporal modulation patterns, especially specific modulation of F0 over time, are easily observable individual acoustic signatures in many mammalian species (e.g., Kershenbaum et al. 2013; Feng et al. 2014; Clink et al. 2017). Also, sequences of sound units can convey complex information including individual identity (e.g., Kershenbaum et al. 2016; Petrusková et al. 2016).

Many different acoustic features and their combinations may contribute to the degree of individual specificity in a vocalization. Pilot studies should first focus on the species where acoustic signatures can be well described by a limited number of identifiable acoustic features. If this is not possible and relevant vocal features cannot be identified based on the signal's structure, or, based on past studies concerning the focal species or its relatives, as many as possible features should be measured and relevant features selected based on their performance in AIID. Alternatively, approaches based on features learnt automatically by classification algorithms could be used (Stowell and Plumbley 2014; Stowell et al. 2019). However, it should be noted that including many non-specific traits or using black-box machine learning algorithms can also cause overfitting of the identification or obscure relevant individually specific features (e.g., Linhart et al. 2019; Stowell et al. 2019) and hinder proper evaluation of the AIID methods. Therefore, non-informative vocal features should be identified and removed from further analysis. Furthermore, it is likely that vocal features that evolved to enhance individual recognition between conspecifics would be more robust on both spatial scales (they will not degrade easily with increasing distance from the sound source) as well as on temporal scales (they will be more stable in time) than features that differ between individuals solely due to stochastic variation. Hence, datasets including signals that have been re-recorded various distances from a speaker, or individuals recorded repeatedly at various time intervals can be used to help find an optimal feature set.

Step 3: Assessing the strength of acoustic signature

Measuring the strength of acoustic signature—The stronger the individual acoustic signature of a species (i.e., the more individual identity information is contained in its vocalizations), the greater the chance for efficient and unambiguous individual acoustic identification. Therefore, quantifying the amount of individual identity information in vocalizations is the first step to develop effective AIID. Proper quantification of individual identity information in vocalizations will also permit a better theoretical understanding of the mechanisms of individual identity signaling.

Discriminant function analysis (DFA) remains the most often used computational method to investigate and quantify individuality in vocalizations in different species. The discrimination score from a DFA summarizes the percentage

of calls that were correctly assigned to correct individual. Other supervised and unsupervised classification methods can be used instead of DFA and these provide analogues to the discrimination score that summarizes the probability of correct assignments to individuals emitting the call. That said, these sorts of classification methods are unsuitable to quantitatively compare individuality due to associated biases resulting from sampling and because they lack biological meaning. Instead, Beecher's information statistic HS (Beecher 1989) has been recently recommended as the best currently available individuality metric (Linhart et al. 2019). Linhart et al. (2019) describe the details to calculate HS and the other commonly used identity metrics. HS is based on information theory and is measured in bits. 2^{HS} provides an estimate of the maximum number of individuals that can be unambiguously acoustically discriminated based on the provided set of measurements (Beecher 1989). HS can be used to quantify individuality of the entire signal or to compare individuality within a single call feature (or in different subsets of call features), which can help identify call features (and subsets of call features) that can be used to maximize the efficacy of AIID.

Strength of acoustic signatures in mammals—We reviewed studies reporting individual variation in vocalizations of mammalian species between 1973 and 2017. We found 130 mammalian species in which individual variation was quantified using either Discriminant Function Analysis (DFA) or Beecher's statistic (HS) (Linhart et al. unpublished data; Supplementary Material 1). We then converted discrimination scores (percentage of correct classifications) from DFA into the HS using the IDmeasurer package (Linhart et al. 2019) to allow comparisons across studies. In these species, HS ranged between 10.5 bits (variation possibly accommodating 1448 unique individual identity signatures) and 0.1 bit (variation possibly accommodating 1 unique individual identity signature (i.e., no individual variation is present)). Our ongoing work indicates that individuality and the potential for individual acoustic identification is higher in adults and in vocalizations that are used in contact (establish and maintain contact) or advertisement (territorial, mate attraction) contexts (Linhart et al. unpublished data). In contrast, individuality and the potential for individual acoustic identification seems to be low in vocalizations used in agonistic, distress, and affiliative contexts. Additionally, bat echolocation calls are generally not suitable for acoustic identification due to their low individuality (but see e.g., Jones and Siemers 2011 for further discussion).

Strength of acoustic signature required for successful AIID and potential for AIID across mammals—Many studies report on different aspects of AIID. For instance,

studies report: the degree of individual variation in vocalizations; identify which vocal features could most likely convey the information about individual identity; describe the computational methods that are optimal to distinguish individuals; discuss whether individual signatures are stable in time, etc. However, this knowledge has only rarely been transformed into subsequent applied use of AIID for any given species (i.e., to use AIID as a complementary or alternative means to mark-capture methods to collect yet unknown information about individuals and population). Only one such study has been recently done on bottlenose dolphins (Longden et al. 2020). Dolphin signature whistles represent a unique identity signalling system, and, to our knowledge, the HS value of bottlenose dolphin signature whistles ($\text{HS} = 13.72$; c.a. 13,500 unique signatures) (Sayigh et al. 2021), is the highest individuality value ever reported for any mammalian or avian vocalization. Furthermore, we reviewed studies quantifying individuality in humans where AIID was routinely applied for various tasks and we also included literature on birds to better estimate the individuality necessary for successful AIID (Table 1). It is apparent that in mammals and birds, AIID was successfully applied in species with relatively high individuality, i.e., with $\text{HS} \geq 7$ bits equivalent to 128 unique signatures or higher. In this perspective, individuality seems surprisingly low in humans (6.9), where we would expect larger individuality. However, this may reflect how individuality was measured in humans; it focused on individuality in a single phoneme (/ε/ phoneme from the 'test n test' phrase) (Bachorowski and Owren 1999).

However, individuality $\text{HS} \geq 7$ neither guarantees successful AIID nor does it mean that AIID is impossible in species with less pronounced individuality. Rather, it should be viewed as the only currently available empirical threshold indicating a good potential for AIID. In mammals, $\text{HS} > 7$ has been found in 23 out of 130 (c.a. 18%) species with the species belonging to Rodentia (9), Primates (8), Carnivora (3), Chiroptera (2), and Cetacea (1) orders. Furthermore, there is a good potential for AIID in vocalizations used in contact (8), advertisement (5), and, in mammals, also in alarm behavioral context (8). Note that the majority of results on mammalian alarm calls have been on studies of ground squirrels. No information is currently available on failed attempts of AIID and the appropriate individuality level in calls of these species. We note again the human example: humans have in general strong vocal individuality, and yet AIID fails in specific situations (Farrús 2018; Lavan et al. 2018).

Table 1 Individuality across mammalian and avian species in which AIID has been applied

Species	Species (scientific)	HS (bits)	References
Mammals			
Bottlenose dolphin	<i>Tursiops truncatus</i>	13.72	AIID—Longden et al (2020); HS—(Sayigh et al. (2021)
Human	<i>Homo sapiens</i>	6.9	AIID—different studies, HS—Bachorowski and Owren (1999)
Birds			
Dupont's lark	<i>Chersophilus duponti</i>	11.1	Laiolo et al. (2007)
European nightjar	<i>Caprimulgus europaeus</i>	9.5	Rebbeck et al. (2001)
African wood owl	<i>Strix woodfordii</i>	8.4	Delpont et al. (2002)
Eurasian bittern	<i>Botaurus stellaris</i>	8.5	AIID—Gilbert et al. (1994); HS—Gilbert et al. (2002)
Mayan antthrush	<i>Formicarius moniliger</i>	7.6	Kirschel et al. (2011)
Scops owl	<i>Otus scops</i>	7.0	Galeotti and Sacchi (2001)

If multiple individuality values were reported, we report the maximum value in the table. Individuality could not be extracted directly from the AIID study on bottlenose dolphins, but HS in bottlenose dolphin whistle signatures was reported by Sayigh et al. (2021) in a recent study. Furthermore, individuality could not be extracted directly from the AIID study on Eurasian bitterns (Gilbert et al. 2002), and, therefore, we report individuality values from the related study by the same authors. In humans, AIID was not applied in an equivalent of mark-recapture study, but it is applied in many different identification tasks, therefore, only a study quantifying individuality is listed. The individuality values (discrimination scores) were converted into HS using IDmeasurer R package (Linhart et al. 2019)

Beyond HS—Individual signatures with larger HS values can be used for AIID in larger populations. In an ideal situation, researchers would conduct a pilot study to estimate HS in a well-sampled subset to estimate the potential of the AIID for a given species/population, and, to roughly estimate the population size that could be reliably monitored (Pollard et al. 2010; Linhart and Šálek 2017) to judge if the potential for AIID matches desired outcomes. While HS provides information about the potential for AIID, it is not the only factor determining AIID success. There are some other dimensions of individuality that may also be important. Overlooking these other dimensions will likely reduce the precision in estimates of population size that can be monitored with AIID (Linhart and Šálek 2017).

For instance, HS provides an upper limit of unique individual signatures. However, it assumes that individuals both partition and completely fill the available acoustic space. In nature, calls from different individuals often cluster together within the acoustic space with considerable overlap between individuals in their acoustic parameters. Similarities between individuals may result from genetic relatedness (Blumstein et al. 2013; Torti et al. 2017), or vocal convergence (Tyack 2008; Briefer and McElligott 2012). In such cases, HS could be overestimated, because acoustic space will not be completely filled and it would be advisable to consider the potential for AIID within, as well as across social units depending on the specific purpose of AIID task (i.e., identification of group members vs. identification of all individuals in a population). We speculate that partitioning of acoustic space might be a poorly described but interesting phenomenon associated with the remarkable vocal plasticity of certain species. Only a few mammals have been

documented to possess the ability to substantially modify their vocalizations (Garcia and Ravignani 2020; Martins and Boeckx 2020; Tyack 2020). For example, dolphins can develop unique identity signatures as opposed to just copying existing signatures (Sayigh et al. 1990), but similar partitioning of acoustic space has not been properly investigated in any other species.

Finally, species with temporally stable vocalizations will be good candidates for AIID (see Step 4). In the future, other dimensions of individuality beyond HS must be considered to better estimate both potential and real limitations for AIID in a given species.

Step 4: Select identification method/algorithm

HS provides information about the strength of acoustic signatures and about the potential for AIID. On the other hand, the actual process of AIID requires vocalizations to be assigned to individuals who most likely produced the vocalization. This can be done by human observers or various classification algorithms. While partial attempts have been made to automate AIID, automatic AIID has not yet been fully achieved and routinely applied in any non-human mammal or bird species.

In principle, AIID represents an open-set classification task, where many new classes (individuals) can appear during the ongoing survey. However, open-set classification algorithms are not trivial to apply and may require extensive material for training and optimization; most supervised classification algorithms (even in deep learning) are designed for closed-set tasks.

Unsupervised classification methods have the advantage that they can be used without the prior knowledge about the individual identity and on the open set of individuals (Clink and Klinck 2021; Sadhukhan et al. 2021). On the other hand, their results still require an external validation as the internal validation metrics were shown to be unreliable (Clink and Klinck 2021). Furthermore, without external validation, it is difficult to include optimization loops that fine-tune the features and classification methods to an AIID task (Fig. 1).

Discriminant function analysis (DFA) represents a supervised classification method that has most often been associated with individual identification, even though it is known that DFA is not suitable for the task, because it is applicable only within a closed set of individuals (Terry et al. 2005; Vögeli et al. 2008). To overcome the closed-set limitation, some AIID studies relied on human judgment (Longden et al. 2020) to assign identities to vocalizations. Alternatively, simple similarity thresholds can be used to decide whether the recorded call belongs to an already known individual (and which one) or whether it represents a new individual (Peake et al. 1998; Laiolo et al. 2007). For surveys of large areas and large populations, it may be useful to integrate information on precise geographic location where samples were recorded, with data on home range size, location and individual movements, to improve classification (Budka et al. 2014). In other words, similar vocal structures recorded in distant locations or times are likely two different individuals.

Ideally, AIID would employ both a supervised phase, to optimize the best feature set for a given task, as well as an optimization of classification criteria (e.g., under which condition a new individual should be established) and an unsupervised phase where the fine-tuned methods would be applied in a different or in a larger population (Sadhukhan et al. 2021). Nevertheless, it would still be wise to continue to check the performance of classification on a set of known individuals.

There is an extensive literature on AIID for human voices, recognised by ear or by machine (for a good overview, see (Hansen and Hasan 2015)). The set of target voices may be small and constrained (e.g., speaker diarization for a meeting) but in many cases, it is large and open-set. The term used for automatic AIID of humans is “speaker recognition” and the analysis is rather different from that of automatic speech recognition of the linguistic content. Preprocessing of acoustic signals into features is well-standardized—Mel-frequency cepstral coefficients are the dominant features that enable speaker recognition and early approaches to speaker recognition would apply generic classification algorithms to such data (Atal 1976). The most successful modern paradigm for speaker recognition is a tailored approach based on constructing a Gaussian mixture model (GMM) representing a “universal background model” (UBM) of human voices,

and then measuring the difference between the voice in question and the UBM. This measured difference (summarized in features such as *Gaussian supervector* or *i-vector*) is then treated somewhat like a fingerprint for the voice in question. GMM-UBM speaker recognition approach is suitable for closed-set as well as open-set recognition tasks and has also been used in pilot studies on animals (Ptacek et al. 2016; Spillmann et al. 2017).

Could human speaker recognition be adopted into mammalian AIID, or are there important differences? Frequency ranges aside, many terrestrial mammals have vocal production similar enough to humans that the same analysis may apply. There is an important hurdle to overcome, however. The majority of work on human speaker recognition concerns single-source voices recorded clearly, in low-noise conditions, on a close microphone (e.g., telephone calls). The dominant paradigm is relatively robust to noise through the development of normalization techniques but retains a strong dependence on the single-source assumption. The approach may thus work for focal recordings, but it is an open question whether it can usefully be applied to passive recordings. The alternative would be to apply more general methods, perhaps based on deep learning, and/or to apply pre-processing sound source separation to recover single-source audio clips. We are not aware of any work applying such modern speaker recognition algorithms to non-human mammal AIID.

However, ecological tools (including AIID) do not necessarily require the most advanced methods (Pimm et al. 2015). For instance, consider a situation that resembles a mark-recapture population study. AIID would likely employ previous recordings and something as simple as asking humans to determine if the vocalizations come from the same or different individuals could be appropriate (Janik 1999; Kershenbaum et al. 2013; Longden et al. 2020). Humans, after all, are astute classifiers and can consider (implicitly) many potentially confounding aspects (e.g., changes in acoustic structure caused by recording distance, motivation, emotional state, etc.). Of course, humans may be somewhat subjective so reviewing the results by several people (Longden et al. 2020) and quantification of classification reliability (Sayigh et al. 2007) should be an essential part of studies relying on human judgments for AIID. Additionally, humans are limited by their working memory and can struggle with AIID involving hundreds or thousands of individuals.

Nevertheless, at least, a partial automation of the AIID process is desirable, because it reduces the analysis burden for users. Potential users of AIID includes conservation managers, field workers, non-bioacoustic researchers, amateur naturalists or community scientists, to name a few, who may already be overwhelmed by other tasks and will only adopt new methods if these methods do not

require steep learning curves. More automated AIID is essential to facilitate widespread adoption of AIID, as well as to yield repeatable outcomes in situations (e.g., many different individuals) where relying on human judgment could be limited.

Advanced machine learning methods could help automate the process. An important resource for the development of advanced machine learning classification methods is the existence of appropriate datasets. In the human speaker recognition field, such a dataset is called corpus and the availability and variability of different corpora is crucial for ongoing development of speaker recognition algorithms (Campbell and Reynolds 1999; Hansen and Hasan 2015).

While there are huge repositories of bird song recordings that can be used to train machine learning algorithms, there are relatively fewer mammalian vocalizations in sound archives. For instance, the British museum lists 140,085 recordings of birds and 7237 recordings of mammals in their 'Wildlife sounds' catalogue. These recordings might be used to study potential for AIID in particular species (e.g., Root-Gutteridge et al. 2014a), but their value for the development of automatic AIID may be limited. Therefore, dedicated AIID datasets need to be published, ideally, within a single online platform similar to, for example, platforms publishing human speech corpora (e.g., <https://catalog.ldc.upenn.edu/>). With no such central platform at hand, authors should still publish their datasets along with appropriate metadata used in their manuscripts using platforms such as Dryad, Figshare, Zenodo, etc. Authors can also advertise their datasets through personal initiatives, for example, Bioacoustic Datasets web page by Justin Salamon (<https://bioacousticdatasets.weebly.com/>). Loosely coordinated data sharing is enhanced by the use of metadata standards such as Darwin Core (Wieczorek et al. 2012), which make it easy for future researchers to combine datasets drawn from multiple sources.

It is likely that AIID in animals, as opposed to human speaker recognition, will face specific challenges due to limitations specific to animal corpora. For example, many animals cannot be recorded under fully controlled conditions or without background noise requiring the use of clever data augmentation methods (Stowell et al. 2019). Both focal recording and passive recording methods have shortcomings when creating a sufficiently large and variable corpus for AIID. Focal recording yields high-quality audio, but may have limitations on the numbers of recordings per individual. Obtaining recordings in a number of different acoustic conditions (various distances, various acoustic backgrounds) might be difficult and time consuming. Passive recording's main shortcoming is that the identity of the vocalizing animal may not be known with certainty. As a result, very few datasets presently exist for testing and fine-tuning the newest classification algorithms for AIID purposes.

Creating and cultivating appropriate sound archives will be essential to support widespread studies and applications of AIID in mammals and animals in general. Missing bioacoustic datasets actually represent a pressing issue across multiple bioacoustic research domains including use of vocalizations to assess animal welfare (McLoughlin et al. 2019), environmental assessment and detection of species (Baker and Vincent 2019; Gibb et al. 2019; Morfi et al. 2019), and AIID (this review). Storage space might pose a problem for sharing entire bioacoustic datasets from large passive acoustic monitoring programs. However, development of AIID would also benefit from sharing limited datasets including only the individuals' call specimens, significantly reducing storage issues. Curation and consistency of metadata may represent a significant obstacle for creation of a single platform for dataset sharing. However, this problem also need not be detrimental in the case of AIID, because at the most extreme case, the only metadata really needed for this purpose is the confirmed identity of the caller. Nevertheless, more details about the context of the recording (call type, age, sex, behavioral context, distance or recording quality) may help to fine tune the methods and explain any arising inconsistencies.

Step 5: Evaluate acoustic signature stability

The existence of individually specific vocalizations permits discrimination between individuals at a given time. However, vocalizations may change over time and this may create challenges for individual discrimination and AIID purposes. Some animals are capable of long-term individual identification (Insley 2000; Briefer et al. 2012). Thus, understanding the temporal stability of acoustic signatures is another important dimension when designing an AIID project. To be able to identify individuals over time, individual identity signatures need to be stable over that time frame. See Terry and McGregor (2002) for distinction between discrimination and identification of individuals.

Vocalizations are dynamic and often state-dependent signals and their acoustic structure may undergo rapid short-term as well as gradual changes over longer time periods depending on various physiological, environmental or social conditions (Newman 1989; Moss et al. 1997; Fischer et al. 2002; Parks et al. 2007; Briefer 2012; Prat et al. 2016; Hradec et al. 2017; Cheney and Seyfarth 2018). Short-term and long-term changes also concern acoustic individual signatures. Such changes can be systematic or stochastic. For example, juveniles go through a maturation process which shapes the final form of their acoustic signatures (Briefer and McElligott 2012; Blumstein et al. 2013; Syrová et al. 2017; Casey et al. 2020). It is still an open question whether and how any systematic (e.g., maturation and senescence effects) and stochastic changes in vocalizations of a single individual

(e.g., health issues, behavioral context) could be compensated for or overcome in AIID and longitudinal studies on known individuals are needed to answer these questions.

To assess temporal stability, subsets of individuals must be systematically recorded over time and their vocalizations compared with those vocalizations recorded in the past. Non-significant changes and high repeatability of vocalization features would indicate a good potential for AIID over a given time period. Nevertheless, it also must be shown that the real emitter of new calls can be accurately predicted by the AIID methods devised from the first set of vocalizations.

Very few studies tested for stability in mammalian acoustic individual signatures. Those few that did recorded individuals during two or more different occasions separated by days, weeks, months, or years. The vocalizations from the first recording were used to build the identification model (typically, discriminant function analysis) and this model has been used to classify vocalizations recorded during the following recording events. Maintenance of a high percentage of correct identifications has been interpreted as evidence of a stable acoustic signature. In contrast, a substantial decrease in correct identifications has been interpreted as an indication of an unstable and changing acoustic signature.

Long-term stability of acoustic signatures (over several years) has been found, for example, in blue monkeys (Butynski et al. 1992), gibbons (Feng et al. 2014), and dolphins (Sayigh et al. 1990). Within-year acoustic stability of identity signals has been found in two species of marmosets (Jones et al. 1993; Jorgensen and French 1998), but it may be modified over the longer time periods of several years, hindering reliable identification on a longer time scale (Jorgensen and French 1998). Similarly, calls of individual cheetahs were stable within a single year but changed over years (Smirnova et al. 2016). Short-term acoustic signatures (lasting from days to weeks) were reported in two different ground squirrel species (Matrosova et al. 2009, 2010). Surprisingly, even in these species, some individuals were able to maintain unique and stable acoustic signatures over long periods spanning two different years (Matrosova et al. 2010). Unfortunately, factors like age and sex were not useful to explain why some individuals retained their signatures and others did not.

AIID in species with mid-term and short-term individuality would likely require more frequent re-recording of animals and classification algorithms capable of incorporating this additional temporal variation. Some studies have followed individuals over long time periods, but they did not divide their data by recording period. Rather, they pooled the vocalizations from different time periods to build the classification model encompassing the total variation within the dataset. A subsample of data from across

the entire dataset is then left out and used for classification and to calculate classification accuracy (Spillmann et al. 2015; Prat et al. 2016). While models from such data are likely better at accommodating many sources of additional variation (e.g., calls from different contexts, social environment, ontogeny, etc.), these studies say little about the stability of signatures or about the performance of AIID in sets of newly collected samples. Such studies may compensate for partial changes in acoustic signatures at the expense of slightly lower overall classification accuracy. Ultimately, the lesson from this is that AIID models should be gradually updated to involve older as well as newly collected samples to ensure robust performance.

Step 6: Validation of the AIID performance

The previous step is associated with the validation of AIID performance. Validation of AIID also needs to be conducted on newly collected samples. Validation consists of two phases. First, we need to compare AIID to another reliable identification method to test reliability of the AIID in a sample of individuals with known identities. In the case of mammals, using artificially marked subjects and/or radio/GPS tagging them should enable vocalizations to be reliably matched to individuals providing a gold standard for AIID performance validation. Some studies may make use of captive animals that are housed individually or in small groups. However, it is not known whether and how individuality could be affected by captive conditions.

The second phase of validation comes into place if the identification is not perfect. It is likely that AIID will suffer different types of misidentification, which is also the case of, for example, visual methods of identification which also rely on the naturally occurring individual phenotypic variation. There, it is valuable to know the amount and quality of the misidentification to minimize errors in estimates of population parameters.

Each of several types of misidentification (Johansson et al. 2020) has specific impacts on the quality of estimates of various population parameters.

- misidentification of a known individual as a new individual (splitting error)—leads to overestimation of population size and underestimation of survival
- misidentification between two known individuals (shifting error)—may affect survival and distort movement patterns
- misidentification of a new individual as a known individual (combination error)—leads to underestimation of population size and overestimation of survival
- sample cannot be classified (exclusion error)

While these identification mistakes are not specific to AIID, specific attributes of acoustic signals could shift their relative prevalence and importance of different misidentification types in AIID compared to other identification methods. For instance, the most individually distinctive acoustic signals are often long-distance signals. They have a large active space and orientation of the animal typically does not influence identification. This should result in a higher total number of detection events than what might be seen in, for example, camera trap studies which require individuals to be in a specific location and position for a good chance of identification. This aspect of AIID should lead to an overall higher number of detections and a reduction of exclusion errors.

On the other hand, acoustic signals are comparatively dynamic and simple signals, and such increased variation could make reliable identification more difficult. This might increase the prevalence of splitting, shifting, and combination errors in case that AIID would be applied within inappropriately large populations with low individuality in vocalizations.

It can be very difficult to comprehensively address each or both of the validation phases of AIID in any species. However, this should not preclude their use. For example, Longden et al. (2020), compared recaptures based on signature whistles with recaptures based on photo-identification, but without a complete knowledge about the true state of population. They found the results of both methods comparable. While neither method is perfect, they both can bring valuable and complementary insights about the species and population, especially in cases when they are the only possible source of any information.

Prospects for future

While various acoustic monitoring schemes have become a common tool to detect a species' presence and activity, monitoring programs routinely integrating acoustic identification of individuals are missing. Our review indicates that Primates and Cetacea are likely good candidates for successful application of AIID, because they involve species with both strong and stable acoustic signatures. It is telling that we are only aware of one study applying AIID in non-human mammals and this has been done with dolphins (Longden et al. 2020) which are known to possess an exceptional identity signaling system—signature whistles. However, information about stability of individual signatures is largely missing in Carnivora and Chiroptera, taxonomic groups which also possess strong identity signatures. Also, more studies should focus on rodents to find out if and under which conditions their individual acoustic signatures remain stable over time. Such studies are crucial for estimating feasibility of AIID

and require repeated recordings of known individuals. At the same time, such studies could also be used to validate AIID and understand how to avoid or account for different types of misidentifications which are currently completely unknown in the case of acoustic signatures.

Despite successful pilot studies using AIID, and despite all the single steps of AIID being able to be managed with existing knowledge, routine application of AIID remains challenging (Table 2). For example, there are no readily available software tools that integrate all necessary steps of AIID. Therefore, interested users must make huge time investments to develop their own analysis pipelines for their species of interest. Furthermore, it is possible that AIID could be limited to certain vocally active animals, certain sexes, age classes, or to certain behavioral contexts. Also, potential users of AIID would likely have to accept some degree of identification uncertainty, a common problem with other methods that exploit naturally occurring cues to individual identity like coat patterns, etc., and handle this uncertainty in further analyses.

Because these issues are, in principle, addressable, we have an optimistic vision for the future of AIID in mammals. Automatic or semiautomatic pipelines for AIID must be developed to integrate species detection in recordings, feature extraction, and identification of the likely emitter. This is likely the currently most pressing bottleneck limitation hindering the spread of AIID. Only such an integrated solution would be adopted by a broad community of potential users who may lack the time and resources to develop their own AIID pipelines.

Initial applications might focus either on model species (e.g., domesticated species), or on 'umbrella species', a species of high conservation importance or common charismatic species. Regardless, it also should be species that are easy to work with (tagging, observing) to allow thorough validation. Such species should also have a high potential for AIID regarding the amount of individual identity information and stability which would guarantee reliable identification results. Unfortunately, none of the domesticated mammalian species seem to possess such a strong individual acoustic signature (pig, HS = 2.4–5.2 bits; cow, HS = 0.5–2.0 bits; goat, HS = 1.9–6.4 bits; dog, HS = 1.1–2.2 bits; cat, HS = 2.6–3.2 bits, Supplementary Material 1), but, on the other hand, studies on domesticated species could be valuable in development and testing of AIID methods for species with suboptimal individuality in their calls. Data necessary for validation of AIID are generally difficult to collect in marine mammals so 'AIID umbrella species' could be likely found among primates and, eventually, among canids, bats, or rodents. A deeper understanding of AIID and its pitfalls within and across different model

Table 2 Summary of challenges and solutions for each AIID step

AIID step	Challenge	Solution
Recording	ID not known	- Pilot studies on a sample of marked individuals - Pilot studies on captive animals - Proxies for identity that can help validate classification (location—territorial animals, time—call sequences come from a single individual, number of individuals in a group/area, etc.)
	Low quality of recordings	- Use focal recording
	Too few vocalizations	- Playback can increase vocal output - Extended deploys of passive recorders in a subject's proximity can increase the amount of recorded material
Feature extraction	Selecting right set of features	- Check studies on similar species or having similar vocalizations - Measure many features and optimize the set of features depending on the task (short-term vs. long-term identification, validation) - Use classification methods based on automatically learnt features
Estimating signature strength	Weak signature	- Optimize extracted features - Eliminate sources of within-individual variation (vocalization type, behavioral context, sex, age, distance, etc.) - Select another type of vocalization - AIID can still be usable in small populations - Larger samples per individual may improve classification results
Classification	Classification is not reliable	- See Weak signature
	ID not known	- Use unsupervised classification methods and compare their results to other known metadata of the calls - Use the ID model parameters from studies on known individuals
Estimating signature stability	Signature not stable	- Use shorter recapture interval and update ID models in each recapture
Validation	ID not known	- See Recording/ID not known
	Splitting errors	- Find better feature sets maximizing signature stability
	Shifting errors	- Find better feature sets maximizing signature strength and stability. Check how classifications match other metadata (location, time) and how they fit to the animal life history to identify nonsense shifting errors
	Combination errors	- Find better feature sets maximizing signature strength
	Exclusion errors	- Collect better recordings

species would eventually lead to more confident use of AIID in situations where individual identity cannot be confirmed by other means and, therefore, AIID would be highly desirable. For example, detailed knowledge of AIID could help us to develop internal validation guidelines for unsupervised or self-supervised classification methods, or, decide on whether confirmation of identity is still required in novel conditions (e.g., new individuals in population, new regions and populations, or even new species). With routine AIID in hand, this method would bring more attention and interest in AIID, more users, and consequently, more data for further improvement of AIID that could solve some of the issues mentioned in this paper.

Appendix 1 Published datasets and other online resources

Bioacoustic datasets including information about individual identity of vocalizing animals. Many of these datasets were originally listed in Sainburg et al. (2020).

Mammals

- Chimpanzee (*Pan troglodytes*) (Plooij et al. 2015)
- Egyptian fruit bat (*Rousettus aegyptiacus*) (Prat et al. 2017)
- White-beaked dolphins (*Lagenorhynchus albirostris*) (Trotter et al. 2019)

- Macaque (Fukushima et al. 2015)

Birds

- Little owl, *Athene noctua* (Stowell et al. 2019)
- Tree pipit, *Anthus trivialis* (Stowell et al. 2019)
- Chiffchaff, *Phylloscopus collybita* (Stowell et al. 2019)
- Bengalese finches (*Lonchura striata* var. *domestica*) (Katahira et al. 2013; Koumura and Okanoya 2016; Nicholson et al. 2017)
- White-rumped munia, *Lonchura striata* (Katahira et al. 2013)
- California thrasher, *Toxostoma redivivum* (Cody et al. 2016)
- Cassin's Vireo, *Vireo cassinii* (Hedley 2016)
- European starling, *Sturnus vulgaris* (Arneodo et al. 2019)
- Swamp sparrow, *Melospiza georgiana* (Lachlan et al. 2018)
- Zebra finches, *Taeniopygia guttata* (Pearre et al. 2017; Elie and Theunissen 2018)

Repositories

- Overview of mammal sound archives—<https://www.mammalwatching.com/resources/mammal-vocalisations-audio-library/>
- MouseTube <https://mousetube.pasteur.fr/>
- Justin Salamon's collection of Bioacoustic datasets and repositories—<https://bioacousticsdatasets.weebly.com>
- BioAcoustica repository (Baker et al. 2015)
- Bird-DB repository (Arriaga et al. 2015)

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42991-021-00222-2>.

Acknowledgements Laela Sayigh, Frants Jensen, and Vincent Janik kindly discussed and shared yet unpublished individual identity content in bottlenose dolphin signature whistles. Justin Salamon, Tim Sainburg advised on collections of bioacoustic datasets and repositories. Alex Průchová, Ema Hrouzková, Siddharth Khopkar, Lucie Hornátová, Eliška Kovářová, and Jan Riegert commented on the manuscript within the Bioacoustic Journal Club at the University of South Bohemia. Work of PL and MMI was partially funded from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 665778 administered by the National Science Centre, Poland (UMO-2015/19/P/NZ8/02507). PL was also funded by the Czech Science Foundation (21-04023K). DTB is funded by the US National Science Foundation.

Author contributions PL and DTB conceived the idea for the article. PL and MMI performed literature search and data analysis. PL, MMI, DS, and DTB drafted and critically revised the manuscript.

Funding The work of PL was partially funded by the Czech Science Foundation (GACR 21-04023 K). Furthermore, the work of PL and also MMI was partially funded from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 665778 administered by the National Science Centre, Poland (UMO-2015/19/P/NZ8/02507). DTB is funded by the US National Science Foundation.

Availability of data and materials Data are provided within the main text of the manuscript and in the supplementary information files.

Code availability R and ID measurer packages were used for mutual conversions of DS and HS identity metrics and are freely available.

Declarations

Conflict of interest Authors declare no conflict/competing interests.

Ethics approval This study did not require ethics approval.

Consent to participate This study did not require consent to participate.

Consent for publication All co-authors reviewed the manuscript and approved its publication.

References

- Arneodo Z, Sainburg T, Jeanne J, Gentner T (2019) An acoustically isolated European starling song library. <https://doi.org/10.5281/zenodo.3237218>
- Adi K, Johnson MT, Osiejuk TS (2010) Acoustic censusing using automatic vocalization classification and identity recognition. *J Acoust Soc Am* 127:874–883. <https://doi.org/10.1121/1.3273887>
- Albersheim-Carter J, Blubaum A, Ballagh IH, Missaghi K, Siuda ER, McMurray G, Bass AH, Dubuc R, Kelley DB, Schmidt MF, Wilson RJA, Gray PA (2016) Testing the evolutionary conservation of vocal motoneurons in vertebrates. *Respir. Physiol. Neurobiol. Physiology of Respiratory Networks of Non-Mammalian Vertebrates* 224:2–10. <https://doi.org/10.1016/j.resp.2015.06.010>
- Arriaga JG, Cody ML, Vallejo EE, Taylor CE (2015) Bird-DB: a database for annotated bird song sequences. *Ecol Inform* 27:21–25. <https://doi.org/10.1016/j.ecoinf.2015.01.007>
- Atal BS (1976) Automatic recognition of speakers from their voices. *Proc IEEE* 64:460–475. <https://doi.org/10.1109/PROC.1976.10155>
- Bachorowski JA, Owren MJ (1999) Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *J Acoust Soc Am* 106:1054–1063. <https://doi.org/10.1121/1.427115>
- Baker E, Vincent S (2019) A deafening silence: a lack of data and reproducibility in published bioacoustics research? *Biodivers Data J* 7:e36783. <https://doi.org/10.3897/BDJ.7.e36783>
- Baker E, Price BW, Rycroft SD, Hill J, Smith VS (2015) BioAcoustica: a free and open repository and analysis platform for bioacoustics. Database. <https://doi.org/10.1093/database/bav054>

- Barkan CL, Zornik E (2020) Inspiring song: the role of respiratory circuitry in the evolution of vertebrate vocal behavior. *Dev Neurobiol* 80:31–41. <https://doi.org/10.1002/dneu.22752>
- Beecher MD (1989) Signalling systems for individual recognition: an information theory approach. *Anim Behav* 38:248–261. [https://doi.org/10.1016/S0003-3472\(89\)80087-9](https://doi.org/10.1016/S0003-3472(89)80087-9)
- Blumstein DT, Munos O (2005) Individual, age and sex-specific information is contained in yellow-bellied marmot alarm calls. *Anim Behav* 69:353–361. <https://doi.org/10.1016/j.anbehav.2004.10.001>
- Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark C, Cortopassi KA, Hanser SF, McCowan B, Ali AM, Kirschel ANG (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *J Appl Ecol* 48:758–767. <https://doi.org/10.1111/j.1365-2664.2011.01993.x>
- Blumstein DT, Nguyen KT, Martin JGA (2013) Ontogenetic variation of heritability and maternal effects in yellow-bellied marmot alarm calls. *Proc R Soc B Biol Sci* 280:20130176. <https://doi.org/10.1098/rspb.2013.0176>
- Bolger DT, Morrison TA, Vance B, Lee D, Farid H (2012) A computer-assisted system for photographic mark–recapture analysis. *Methods Ecol Evol* 3:813–822. <https://doi.org/10.1111/j.2041-210X.2012.00212.x>
- Briefer EF (2012) Vocal expression of emotions in mammals: mechanisms of production and evidence. *J Zool* 288:1–20. <https://doi.org/10.1111/j.1469-7998.2012.00920.x>
- Briefer EF, McElligott AG (2012) Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Anim Behav* 83:991–1000. <https://doi.org/10.1016/j.anbehav.2012.01.020>
- Briefer EF, Padilla de la Torre M, McElligott AG (2012) Mother goats do not forget their kids' calls. *Proc R Soc B Biol Sci* 279:3749–3755. <https://doi.org/10.1098/rspb.2012.0986>
- Budka M, Wojas L, Osiejuk TS (2014) Is it possible to acoustically identify individuals within a population? *J Ornithol* 156:481–488. <https://doi.org/10.1007/s10336-014-1149-2>
- Butynski TM, Chapman CA, Chapman LJ, Weary DM (1992) Use of male blue monkey “Pyow” calls for long-term individual identification. *Am J Primatol* 28:183–189. <https://doi.org/10.1002/ajp.1350280303>
- Byers KA, Lee MJ, Bidulka JJ, Patrick DM, Himsforth CG (2019) Rat in a cage: Trappability of urban Norway rats (*Rattus norvegicus*). *Front Ecol Evol*. <https://doi.org/10.3389/fevo.2019.00068>
- Campbell J, Reynolds D (1999) Corpora for the evaluation of speaker recognition systems. 1999 IEEE Int. Conf. Acoust. Speech Signal Process. Proc. ICASSP99 Cat No99CH36258. <https://doi.org/10.1109/ICASSP.1999.759799>
- Casey C, Charrier I, Mathevon N, Nasr C, Forman P, Reichmuth C (2020) The genesis of giants: behavioural ontogeny of male northern elephant seals. *Anim Behav* 166:247–259. <https://doi.org/10.1016/j.anbehav.2020.06.014>
- Charrier I (2020) Mother–offspring vocal recognition and social system in pinnipeds. In: Aubin T, Mathevon N (eds) Coding strategies in vertebrate acoustic communication, animal signals and communication. Springer International Publishing, Cham, pp 231–246. https://doi.org/10.1007/978-3-030-39200-0_9
- Chen Z, Wiens JJ (2020) The origins of acoustic communication in vertebrates. *Nat Commun* 11:1–8. <https://doi.org/10.1038/s41467-020-14356-3>
- Chen JC, Yip L, Elson J, Wang H, Maniezzo D, Hudson RE, Yao K, Estrin D (2003) Coherent acoustic array processing and localization on wireless sensor networks. *Proc IEEE* 91:1154–1162. <https://doi.org/10.1109/JPROC.2003.814924>
- Cheney DL, Seyfarth RM (2018) Flexible usage and social function in primate vocalizations. *Proc Natl Acad Sci USA* 115:1974–1979. <https://doi.org/10.1073/pnas.1717572115>
- Clack JA (2002) Patterns and processes in the early evolution of the tetrapod ear. *J Neurobiol* 53:251–264. <https://doi.org/10.1002/neu.10129>
- Clink DJ, Klinck H (2021) Unsupervised acoustic classification of individual gibbon females and the implications for passive acoustic monitoring. *Methods Ecol Evol* 12:328–341. <https://doi.org/10.1111/2041-210X.13520>
- Clink DJ, Bernard H, Crofoot MC, Marshall AJ (2017) Investigating individual vocal signatures and small-scale patterns of geographic variation in female Bornean gibbon (*Hylobates muelleri*) Great Calls. *Int J Primatol* 38:656–671. <https://doi.org/10.1007/s10764-017-9972-y>
- Cody ML, Stabler E, Castellanos HMS, Taylor CE (2016) Structure, syntax and “small-world” organization in the complex songs of California thrashers (*Toxostoma redivivum*). *Bioacoustics* 25:41–54. <https://doi.org/10.1080/09524622.2015.1089418>
- Crall JP, Stewart CV, Berger-Wolf TY, Rubenstein DI, Sundaresan SR (2013) HotSpotter-Patterned species instance recognition. In: 2013 IEEE Workshop Appl. Comput. Vis. WACV 2013, Proceedings of IEEE Workshop on Applications of Computer Vision 230–237. <https://doi.org/10.1109/WACV.2013.6475023>
- Cunningham EP, Unwin S, Setchell JM (2015) Darting primates in the field: a review of reporting trends and a survey of practices and their effect on the primates involved. *Int J Primatol* 36:911–932. <https://doi.org/10.1007/s10764-015-9862-0>
- Delport W, Kemp AC, Ferguson JWH (2002) Vocal identification of individual African Wood Owls *Strix woodfordii*: a technique to monitor long-term adult turnover and residency. *Ibis* 144:30–39. <https://doi.org/10.1046/j.0019-1019.2001.00019.x>
- Elie JE, Theunissen FE (2018) Zebra finches identify individuals using vocal signatures unique to each call type. *Nat Commun* 9:4026. <https://doi.org/10.1038/s41467-018-06394-9>
- Farrús M (2018) Voice disguise in automatic speaker recognition. *ACM Comput Surv* 51 68(1–68):22. <https://doi.org/10.1145/3195832>
- Feng J-J, Cui L-W, Ma C-Y, Fei H-L, Fan P-F (2014) Individuality and stability in male songs of Cao-vit gibbons (*Nomascus nasutus*) with potential to monitor population dynamics. *PLoS ONE* 9:e96317. <https://doi.org/10.1371/journal.pone.0096317>
- Fischer J, Hammerschmidt K, Cheney DL, Seyfarth RM (2002) Acoustic features of male baboon loud calls: influences of context, age, and individuality. *J Acoust Soc Am* 111:1465–1474. <https://doi.org/10.1121/1.1433807>
- Fukushima M, Doyle AM, Mullarkey MP, Mishkin M, Averbeck BB (2015) Distributed acoustic cues for caller identity in macaque vocalization. *R Soc Open Sci*. <https://doi.org/10.1098/rsos.150432>
- Galeotti P, Sacchi R (2001) Turnover of territorial Scops Owls *Otus scops* as estimated by spectrographic analyses of male hoots. *J Avian Biol* 32:256–262. <https://doi.org/10.1111/j.0908-8857.2001.320308.x>
- Garcia M, Ravignani A (2020) Acoustic allometry and vocal learning in mammals. *Biol Lett* 16:20200081. <https://doi.org/10.1098/rsbl.2020.0081>
- Gibb R, Browning E, Glover-Kapfer P, Jones KE (2019) Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods Ecol Evol* 10:169–185. <https://doi.org/10.1111/2041-210X.13101>
- Gilbert G, McGregor PK, Tyler G (1994) Vocal individuality as a census tool - Practical considerations illustrated by a study of two rare species. *J Field Ornithol* 65:335–348
- Gilbert G, Tyler GA, Smith KW (2002) Local annual survival of booming male Great Bittern *Botaurus stellaris* in Britain, in the period 1990–1999. *Ibis* 144:51–61. <https://doi.org/10.1046/j.0019-1019.2001.00012.x>

- Hansen JHL, Hasan T (2015) Speaker recognition by machines and humans: a tutorial review. *IEEE Signal Process Mag* 32:74–99. <https://doi.org/10.1109/MSP.2015.2462851>
- Hedley RW (2016) Complexity, predictability and time homogeneity of syntax in the songs of Cassin's vireo (*Vireo cassinii*). *PLoS ONE* 11:e0150822. <https://doi.org/10.1371/journal.pone.0150822>
- Heinicke S, Kalan AK, Wagner OJJ, Mundry R, Lukashevich H, Kühl HS (2015) Assessing the performance of a semi-automated acoustic monitoring system for primates. *Methods Ecol Evol* 6:753–763. <https://doi.org/10.1111/2041-210X.12384>
- Hradec M, Linhart P, Bartoš L, Bolechová P (2017) The traits of the great calls in the juvenile and adolescent gibbon males *Nomascus gabriellae*. *PLoS ONE* 12:e0173959. <https://doi.org/10.1371/journal.pone.0173959>
- Inslay SJ (2000) Long-term vocal recognition in the northern fur seal. *Nature* 406:404–405. <https://doi.org/10.1038/35019064>
- Iossa, G., Soulsbury, C.D., Harris, S., 2007. Mammal trapping: A review of animal welfare standards of killing and restraining traps. *Anim. Welf.*
- Janik VM (1999) Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. *Anim Behav* 57:133–143. <https://doi.org/10.1006/anbe.1998.0923>
- Janik VM, Sayigh LS (2013) Communication in bottlenose dolphins: 50 years of signature whistle research. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 199:479–489. <https://doi.org/10.1007/s00359-013-0817-7>
- Janik VM, Slater PJB (2000) The different roles of social learning in vocal communication. *Anim Behav* 60:1–11. <https://doi.org/10.1006/anbe.2000.1410>
- Johansson Ö, Samelius G, Wikberg E, Chapron G, Mishra C, Low M (2020) Identification errors in camera-trap studies result in systematic population overestimation. *Sci Rep* 10:6393. <https://doi.org/10.1038/s41598-020-63367-z>
- Jones G, Siemers BM (2011) The communicative potential of bat echolocation pulses. *J Comp Physiol A* 197:447–457. <https://doi.org/10.1007/s00359-010-0565-x>
- Jones B, Harris D, Catchpole C (1993) The stability of the vocal signature in phoe calls of the common marmoset, *Callithrix jacchus*. *Am J Primatol* 31:67–75. <https://doi.org/10.1002/ajp.1350310107>
- Jorgensen DD, French JA (1998) Individuality but not stability in marmoset long calls. *Ethology* 104:729–742. <https://doi.org/10.1111/j.1439-0310.1998.tb00107.x>
- Kalan AK, Mundry R, Wagner OJJ, Heinicke S, Boesch C, Kühl HS (2015) Towards the automated detection and occupancy estimation of primates using passive acoustic monitoring. *Ecol Indic* 54:217–226. <https://doi.org/10.1016/j.ecolind.2015.02.023>
- Kalan AK, Piel AK, Mundry R, Wittig RM, Boesch C, Kühl HS (2016) Passive acoustic monitoring reveals group ranging and territory use: a case study of wild chimpanzees (*Pan troglodytes*). *Front Zool* 13:34. <https://doi.org/10.1186/s12983-016-0167-8>
- Karanth KU, Nichols JD (1998) Estimation of tiger densities in India using photographic captures and recaptures. *Ecology* 79:2852–2862. [https://doi.org/10.1890/0012-9658\(1998\)079\[2852:EOT-DII\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1998)079[2852:EOT-DII]2.0.CO;2)
- Katahira K, Suzuki K, Kagawa H, Okanoya K (2013) A simple explanation for the evolution of complex song syntax in Bengalese finches. *Biol Lett* 9:20130842. <https://doi.org/10.1098/rsbl.2013.0842>
- Karczmarski L, Chan SCY, Rubenstein DI, Chui SYS, Cameron EZ (2022a) Individual identification and photographic techniques in mammalian ecological and behavioural research – Part 1: Methods and concepts. *Mamm Biol (Special Issue)* 102(2). <https://link.springer.com/journal/42991/volumes-and-issues/102-2>
- Karczmarski L, Chan SCY, Chui SYS, Cameron EZ (2022b) Individual identification and photographic techniques in mammalian ecological and behavioural research – Part 2: Field studies and applications. *Mamm Biol (Special Issue)* 102(3). <https://link.springer.com/journal/42991/volumes-and-issues/102-3>
- Kershenbaum A, Sayigh LS, Janik VM (2013) The encoding of individual identity in dolphin signature whistles: how much information is needed? *PLoS ONE* 8:e77671. <https://doi.org/10.1371/journal.pone.0077671>
- Kershenbaum A, Blumstein DT, Roch MA, Akçay Ç, Backus G, Bee MA, Bohn K, Cao Y, Carter G, Cäsar C, Coen M, DeRuiter SL, Doyle L, Edelman S, Ferrer-i-Cancho R, Freeberg TM, Garland EC, Gustison M, Harley HE, Huetz C, Hughes M, Bruno JH, Ilany A, Jin DZ, Johnson M, Ju C, Karnowski J, Lohr B, Manser MB, McCowan B, Mercado E, Narins PM, Piel A, Rice M, Salmi R, Sasahara K, Sayigh L, Shiu Y, Taylor C, Vallejo EE, Waller S, Zamora-Gutierrez V (2016) Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol Rev* 91:13–52. <https://doi.org/10.1111/brv.12160>
- Kershenbaum A, Owens JL, Waller S (2019) Tracking cryptic animals using acoustic multilateration: a system for long-range wolf detection. *J Acoust Soc Am* 145:1619. <https://doi.org/10.1121/1.5092973>
- Kingsley EP, Eliason CM, Riede T, Li Z, Hiscock TW, Farnsworth M, Thomson SL, Goller F, Tabin CJ, Clarke JA (2018) Identity and novelty in the avian syrinx. *Proc Natl Acad Sci USA* 115:10209–10217. <https://doi.org/10.1073/pnas.1804586115>
- Kirschel ANG, Cody ML, Harlow ZT, Promponas VJ, Vallejo EE, Taylor CE (2011) Territorial dynamics of Mexican ant-thrushes *Formicarius moniliger* revealed by individual recognition of their songs. *Ibis* 153:255–268. <https://doi.org/10.1111/j.1474-919X.2011.01102.x>
- Koumura T, Okanoya K (2016) Automatic recognition of element classes and boundaries in the birdsong with variable sequences. *PLoS ONE* 11:e0159188. <https://doi.org/10.1371/journal.pone.0159188>
- Kukalová M, Gazárková A, Adamík P (2013) Should I stay or should I go? The influence of handling by researchers on den use in an arboreal nocturnal rodent. *Ethology* 119:848–859. <https://doi.org/10.1111/eth.12126>
- Lachlan RF, Ratmann O, Nowicki S (2018) Cultural conformity generates extremely stable traditions in bird song. *Nat Commun* 9:2417. <https://doi.org/10.1038/s41467-018-04728-1>
- Laiolo P, Vogeli M, Serrano D, Tella J (2007) Testing acoustic versus physical marking: two complementary methods for individual-based monitoring of elusive species. *J Avian Biol* 38:672–681. <https://doi.org/10.1111/j.2007.0908-8857.04006.x>
- Lameira AR, Wich SA (2008) Orangutan long call degradation and individuality over distance: a playback approach. *Int J Primatol* 29:615–625. <https://doi.org/10.1007/s10764-008-9253-x>
- Lavan N, Short B, Wilding A, McGettigan C (2018) Impoverished encoding of speaker identity in spontaneous laughter. *Evol Hum Behav* 39:139–145. <https://doi.org/10.1016/j.evolhumbehav.2017.11.002>
- Lavan N, Knight S, McGettigan C (2019) Listeners form average-based representations of individual voice identities. *Nat Commun* 10:2404. <https://doi.org/10.1038/s41467-019-10295-w>
- Linhart P, Šálek M (2017) The assessment of biases in the acoustic discrimination of individuals. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0177206>
- Linhart P, Fuchs R, Poláková S, Slabbekoorn H (2012) Once bitten twice shy: long-term behavioural changes caused by trapping experience in willow warblers *Phylloscopus trochilus*. *J Avian Biol* 43:186–192. <https://doi.org/10.1111/j.1600-048X.2012.05580.x>
- Linhart P, Osiejuk TS, Budka M, Šálek M, Špinková M, Policht R, Syrová M, Blumstein DT (2019) Measuring individual identity information in animal signals: overview and performance of available

- identity metrics. *Methods Ecol Evol* 10:1558–1570. <https://doi.org/10.1111/2041-210X.13238>
- Longden EG, Elwen SH, McGovern B, James BS, Embling CB, Gridley T (2020) Mark–recapture of individually distinctive calls—a case study with signature whistles of bottlenose dolphins (*Tursiops truncatus*). *J Mammal* 101:1289–1301. <https://doi.org/10.1093/jmammal/gyaa081>
- Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Moretti DJ, Harris D, Tyack PL (2013) Estimating animal population density using passive acoustics. *Biol Rev* 88:287–309. <https://doi.org/10.1111/brv.12001>
- Martins PT, Boeckx C (2020) Vocal learning: Beyond the continuum. *PLOS Biol* 18:e3000672. <https://doi.org/10.1371/journal.pbio.3000672>
- Mathevon N, Casey C, Reichmuth C, Charrier I (2017) Northern elephant seals memorize the rhythm and timbre of their rivals' Voices. *Curr Biol* 27:2352–2356.e2. <https://doi.org/10.1016/j.cub.2017.06.035>
- Matrosova V, Volodin I, Volodina E (2009) Short-term and long-term individuality in speckled ground squirrel alarm calls. *J Mammal* 90:158–166. <https://doi.org/10.1644/08-MAMM-A-032.1>
- Matrosova V, Volodin I, Volodina E, Vasilieva N (2010) Stability of acoustic individuality in the alarm calls of wild yellow ground squirrels *Spermophilus fulvus* and contrasting calls from trapped and free-ranging callers. *Naturwissenschaften* 97:707–715. <https://doi.org/10.1007/s00114-010-0686-7>
- McLoughlin MP, Stewart R, McElligott AG (2019) Automated bioacoustics: methods in ecology and conservation and their potential for animal welfare monitoring. *J R Soc Interface* 16:20190225. <https://doi.org/10.1098/rsif.2019.0225>
- Morfi V, Bas Y, Pamuła H, Glotin H, Stowell D (2019) NIPS4Bplus: a richly annotated birdsong audio dataset. *PeerJ Comput Sci* 5:e223. <https://doi.org/10.7717/peerj-cs.223>
- Moss CF, Redish D, Gounden C, Kunz TH (1997) Ontogeny of vocal signals in the little brown bat, *Myotis lucifugus*. *Anim Behav* 54:131–141. <https://doi.org/10.1006/anbe.1996.0410>
- Mouterde SC, Theunissen FE, Elie JE, Vignal C, Mathevon N (2014) Acoustic communication and sound degradation: How do the individual signatures of male and female zebra finch calls transmit over distance? *PLoS ONE* 9:e102842. <https://doi.org/10.1371/journal.pone.0102842>
- Newman JD (1989) The physiological control of mammalian vocalization, 1st edn. Springer, New York
- Nicholson D, Queen JE, Sober SJ (2017) Bengalese finch song repository. <https://doi.org/10.6084/m9.figshare.4805749.v5>
- Parks SE, Clark CW, Tyack PL (2007) Short- and long-term changes in right whale calling behavior: the potential effects of noise on acoustic communication. *J Acoust Soc Am* 122:3725–3731. <https://doi.org/10.1121/1.2799904>
- Peake TM, McGregor PK, Smith KW, Tyler G, Gilbert G, Green RE (1998) Individuality in corncrake *Crex crex* vocalizations. *Ibis* 140:120–127. <https://doi.org/10.1111/j.1474-919X.1998.tb04548.x>
- Pearre B, Perkins LN, Markowitz JE, Gardner TJ (2017) A fast and accurate zebra finch syllable detector. *PLoS ONE* 12:e0181992. <https://doi.org/10.1371/journal.pone.0181992>
- Petrusková T, Pišvejcová I, Kinštová A, Brinke T, Petrušek A (2016) Repertoire-based individual acoustic monitoring of a migratory passerine bird with complex song as an efficient tool for tracking territorial dynamics and annual return rates. *Methods Ecol Evol* 7:274–284. <https://doi.org/10.1111/2041-210X.12496>
- Pimm SL, Alibhai S, Bergl R, Dehgan A, Giri C, Jewell Z, Joppa L, Kays R, Loarie S (2015) Emerging technologies to conserve biodiversity. *Trends Ecol Evol* 30:685–696. <https://doi.org/10.1016/j.tree.2015.08.008>
- Plooij FX, van de Rijt-Plooij H, Fischer M, Wilson ML, Pusey A (2015) An archive of longitudinal recordings of the vocalizations of adult Gombe chimpanzees. *Sci Data* 2:150027. <https://doi.org/10.1038/sdata.2015.27>
- Plumptre AJ, Sterling EJ, Buckland ST (2013) Primate census and survey techniques. In: *Primate ecology and conservation*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199659449.003.0002>
- Pollard KA, Blumstein DT (2011) Social group size predicts the evolution of individuality. *Curr Biol* 21:413–417. <https://doi.org/10.1016/j.cub.2011.01.051>
- Pollard KA, Blumstein DT, Griffin SC (2010) Pre-screening acoustic and other natural signatures for use in noninvasive individual identification. *J Appl Ecol* 47:1103–1109. <https://doi.org/10.1111/j.1365-2664.2010.01851.x>
- Powell RA, Proulx G (2003) Trapping and marking terrestrial mammals for research: integrating ethics, performance criteria, techniques, and common sense. *ILAR J* 44:259–276. <https://doi.org/10.1093/ilar.44.4.259>
- Prat Y, Taub M, Yovel Y (2016) Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Sci Rep*. <https://doi.org/10.1038/srep39419>
- Prat Y, Taub M, Pratt E, Yovel Y (2017) An annotated dataset of egyptian fruit bat vocalizations across varying contexts and during vocal ontogeny. *Sci Data* 4:170143. <https://doi.org/10.1038/sdata.2017.143>
- Ptacek L, Machlica L, Linhart P, Jaska P, Muller L (2016) Automatic recognition of bird individuals on an open set using as-is recordings. *Bioacoustics Int J Anim Sound Its Rec* 25:55–73. <https://doi.org/10.1080/09524622.2015.1089524>
- Rebbeck M, Corrick R, Eaglestone B, Stainton C (2001) Recognition of individual European Nightjars *Caprimulgus europaeus* from their song. *Ibis* 143:468–475. <https://doi.org/10.1111/j.1474-919X.2001.tb04948.x>
- Rhinehart TA, Chronister LM, Devlin T, Kitzes J (2020) Acoustic localization of terrestrial wildlife: current practices and future opportunities. *Ecol Evol* 10:6794–6818. <https://doi.org/10.1002/ece3.6216>
- Riede T, Thomson SL, Titze IR, Goller F (2019) The evolution of the syrinx: an acoustic theory. *PLOS Biol* 17:e2006507. <https://doi.org/10.1371/journal.pbio.2006507>
- Root-Gutteridge H, Bencsik M, Chebli M, Gentle LK, Terrell-Nield C, Bourit A, Yarnell RW (2014a) Identifying individual wild eastern grey wolves (*Canis lupus lycaon*) using fundamental frequency and amplitude of howls. *Bioacoustics Int J Anim Sound Its Rec* 23:55–66. <https://doi.org/10.1080/09524622.2013.817317>
- Root-Gutteridge H, Bencsik M, Chebli M, Gentle LK, Terrell-Nield C, Bourit A, Yarnell RW (2014b) Improving individual identification in captive eastern grey wolves (*Canis lupus lycaon*) using the time course of howl amplitudes. *Bioacoustics Int J Anim Sound Its Rec* 23:39–53. <https://doi.org/10.1080/09524622.2013.817318>
- Sadhukhan S, Root-Gutteridge H, Habib B (2021) Identifying unknown Indian wolves by their distinctive howls: its potential as a non-invasive survey method. *Sci Rep* 11:7309. <https://doi.org/10.1038/s41598-021-86718-w>
- Sainburg T, Thielk M, Gentner TQ (2020) Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLOS Comput Biol* 16:e1008228. <https://doi.org/10.1371/journal.pcbi.1008228>
- Sayigh LS, Tyack PL, Wells RS, Scott MD (1990) Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*:

- stability and mother-offspring comparisons. *Behav Ecol Sociobiol* 26:247–260. <https://doi.org/10.1007/BF00178318>
- Sayigh LS, Esch HC, Wells RS, Janik VM (2007) Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Anim Behav* 74:1631–1642. <https://doi.org/10.1016/j.anbehav.2007.02.018>
- Sayigh LS, Janik VM, Jensen FH, Scott MD, Tyack PL, Wells RS (2021) The Sarasota dolphin whistle database: A unique long-term resource for understanding dolphin communication. *Frontiers in Marine Science, Marine Conservation and Sustainability, Special Topic: The dolphins of Sarasota Bay: Lessons from 50 years of research and conservation (in press)*
- Senter P (2008) Voices of the past: a review of Paleozoic and Mesozoic animal sounds. *Hist Biol* 20:255–287. <https://doi.org/10.1080/08912960903303327>
- Smirnova DS, Volodin IA, Demina TS, Volodina EV (2016) Acoustic structure and contextual use of calls by captive male and female cheetahs (*Acinonyx jubatus*). *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0158546>
- Soulsbury CD, Gray HE, Smith LM, Braithwaite V, Cotter SC, Elwood RW, Wilkinson A, Collins LM (2020) The welfare and ethics of research involving wild animals: a primer. *Methods Ecol Evol* 11:1164–1181. <https://doi.org/10.1111/2041-210X.13435>
- Spillmann B, van Noordwijk MA, Willems EP, Mitra Setia T, Wipfli U, van Schaik CP (2015) Validation of an acoustic location system to monitor Bornean orangutan (*Pongo pygmaeus wurmbii*) long calls. *Am J Primatol* 77:767–776. <https://doi.org/10.1002/ajp.22398>
- Spillmann B, van Schaik CP, Setia TM, Sadjadi SO (2017) Who shall I say is calling? Validation of a caller recognition procedure in Bornean flanged male orangutan (*Pongo pygmaeus wurmbii*) long calls. *Bioacoustics Int J Anim Sound Its Rec* 26:109–120. <https://doi.org/10.1080/09524622.2016.1216802>
- Stowell D, Plumbley MD (2014) Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ* 2:e488. <https://doi.org/10.7717/peerj.488>
- Stowell D, Tereza P, Martin Š, Pavel L (2019) Automatic acoustic identification of individuals in multiple species: improving identification across recording conditions. *J R Soc Interface* 16:20180940. <https://doi.org/10.1098/rsif.2018.0940>
- Syrová M, Policht R, Linhart P, Špinka M (2017) Ontogeny of individual and litter identity signaling in grunts of piglets. *J Acoust Soc Am* 142:3116–3121. <https://doi.org/10.1121/1.5010330>
- Taylor AM, Reby D (2010) The contribution of source-filter theory to mammal vocal communication research. *J Zool* 280:221–236. <https://doi.org/10.1111/j.1469-7998.2009.00661.x>
- Terry AMR, McGregor PK (2002) Census and monitoring based on individually identifiable vocalizations: the role of neural networks. *Anim Conserv* 5:103–111. <https://doi.org/10.1017/S1367943002002147>
- Terry AM, Peake TM, McGregor PK (2005) The role of vocal individuality in conservation. *Front Zool* 2:10. <https://doi.org/10.1186/1742-9994-2-10>
- Tienhoven AMV, Hartog JED, Reijns RA, Peddemors VM (2007) A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. *J Appl Ecol* 44:273–280. <https://doi.org/10.1111/j.1365-2664.2006.01273.x>
- Torti V, Bonadonna G, De Gregorio C, Valente D, Randrianarison RM, Friard O, Pozzi L, Gamba M, Giacoma C (2017) An intra-population analysis of the indris' song dissimilarity in the light of genetic distance. *Sci Rep* 7:10140. <https://doi.org/10.1038/s41598-017-10656-9>
- Trotter C, Atkinson G, Sharpe M, McGough AS, Wright N, Berggren P (2019) The Northumberland dolphin dataset: A multimedia individual cetacean dataset for fine-grained categorisation. *arXiv190802669 Cs*
- Tyack PL (1997) Development and social functions of signature whistles in bottlenose dolphins *Tursiops truncatus*. *Bioacoustics* 8:21–46. <https://doi.org/10.1080/09524622.1997.9753352>
- Tyack PL (2008) Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *J Comp Psychol* 122:319–331. <https://doi.org/10.1037/a0013087>
- Tyack PL (2020) A taxonomy for vocal learning. *Trans R Soc B Biol Sci Philos*. <https://doi.org/10.1098/rstb.2018.0406>
- Vögeli M, Laiolo P, Serrano D, Tella JL (2008) Who are we sampling? Apparent survival differs between methods in a secretive species. *Oikos* 117:1816–1823. <https://doi.org/10.1111/j.1600-0706.2008.17225.x>
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin core: an evolving community-developed biodiversity data standard. *PLoS ONE* 7:e29715. <https://doi.org/10.1371/journal.pone.0029715>
- Wiley RH (2013) Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biol Rev* 88:179–195. <https://doi.org/10.1111/j.1469-185X.2012.00246.x>
- Wilkinson GS (2003) Social and vocal complexity in bats. In: *Animal social complexity: Intelligence, culture, and individualized societies*. Harvard University Press, Cambridge, MA, US, pp. 322–341. <https://doi.org/10.4159/harvard.9780674419131.c24>
- Wrege PH, Rowland ED, Keen S, Shiu Y (2017) Acoustic monitoring for conservation in tropical forests: examples from forest elephants. *Methods Ecol Evol* 8:1292–1301. <https://doi.org/10.1111/2041-210X.12730>
- Würsig B, Würsig M (1977) The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*). *Science* 198:755–756. <https://doi.org/10.1126/science.198.4318.755>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.