Vocal discrimination of African lions and its potential for collar-free tracking

Matthew Wijers[1], Paul Trethowan[1], Byron du Preez[1], Simon Chamaillé-Jammes[3,4], Andrew J. Loveridge[1], David W. Macdonald[1] and Andrew Markham[2]

[1] Wildlife Conservation Research Unit, Recanati-Kaplan Centre, Department of Zoology, University of Oxford, United Kingdom.

[2] Department of Computer Science, University of Oxford, United Kingdom

[3] CEFE, CNRS, University Montpellier, University Paul Valéry Montpellier, EPHE, IRD, Montpellier, France

[4] Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

Corresponding author:

Matthew Wijers

matthew.wijers@gmail.com

Key words: African lion, fundamental frequency, passive acoustic monitoring, vocal individuality, vocalisation.

Abstract

Previous research has shown that African lions (*Panthera leo*) have the ability to discriminate between conspecific vocalisations, but little is known about how individual identity is conveyed in the spectral structure of roars. Using acoustic - accelerometer biologgers that allow vocalisations to be reliably associated with individual identity, we test for vocal individuality in the fundamental frequency (*f*0) of roars from 5 male lions, firstly by comparing simple *f*0 summary features and secondly by modelling the temporal pattern of the *f*0 contour. We then assess the application of this method for discriminating between individuals using passive acoustic monitoring. Results indicate that *f*0 summary features only allow for vocal discrimination with 70.7% accuracy. By comparison, vocal discrimination can be achieved with an accuracy of 91.5% based on individual differences in the temporal pattern of the *f*0 sequence. We further demonstrate that passively recorded lion roars can be localised and differentiated with similar accuracy. The existence of individually unique *f*0 contours in lion roars and their relatively lower attenuation indicates a likely mechanism enabling individual lions to identify conspecifics over long distances. These differences can be exploited by researchers to track individuals across the landscape and thereby supplement conventional lion monitoring approaches.

Keywords

African lion, fundamental frequency, passive acoustic monitoring, vocalisation, vocal individuality.

Introduction

African lions (*Panthera leo*) are well known for their loud, characteristic roars, frequently audible from several kilometres away. Lions are, however, just one of many terrestrial mammalian species that are capable of communicating vocally over large distances. The basic function of long distance signalling is to aid individuals in searching for, locating and avoiding each other and therefore facilitating group cohesion and the maintenance of territorial boundaries (Marler 1967; Mitani and Nishida 1993; Ramsauer 2005). For many species, a key component of the information conveyed to conspecifics through the use of long-distance calls is individual identity. Detection and decoding of this information are crucial for influencing the response of the receiver. For example, in the context of territoriality, identifying a call from a neighbouring animal in their usual location will likely elicit a different response to that of an unfamiliar animal (McGregor 1993). Several studies on terrestrial mammals such as elephants (*Loxodonta africana*; Clemins et al. 2005), orangutans (*Pongo pygmaeua wurmbii*; Spillmann et al. 2017), tigers (*Panthera tigris*; Ji et al. 2013) wild dogs (*Lycaon pictus*; Hartwig 2005) and wolves (*Canis lupus lycaon*; Root-Gutteridge et al. 2013) have shown that certain elements of vocal signals are unique to individuals and may convey information relating to caller identity. The production of these unique call features is likely related to small variations in the morphology of the vocal apparatus between individuals (Fitch et al. 1997; Ey et al. 2007).

Despite the ability of many animals to encode individually unique features within their calls, sound signals degrade progressively as they propagate through the environment. Higher frequencies are more rapidly absorbed through air and incur greater scattering by small objects (e.g. vegetation) compared to lower frequencies, resulting in changes to the spectral structure of the original signal (Forrest 1994; Wiley and Richards 1978). This degradation can impair

59 the ability of distant receivers to extract information. Several studies on birds have shown that
60 individuals responded more strongly to undegraded calls than to degraded calls (McGregor et
61 al. 1983; Mathevon and Aubin 1997). Despite the loss of information incurred during signal
62 propagation, some species are able to learn and accurately recognise the degraded call of a
63 conspecific through experience (Aubin and Jouventin 2002; Mouterde et al. 2014).

64 Knowledge of how species transmit information relating to individual identity is important for
65 understanding species communication mechanisms. In some cases, vocal individuality can also
66 be exploited by researchers and conservationists to monitor and survey species populations
67 using individually distinct vocal features as a non-invasive marking method (Terry et al. 2005).
68 For example, Gilbert et al. (2002) investigated the survival and movements of Great Bitterns
69 (*Botaurus stellaris*) using spectrogram measures of vocalisations. Similarly, Delport et al.
70 (2002) relied on vocal identification of African wood owls (*Strix woodfordii*) to monitor
71 territory turnover. However, despite the large number of studies that have explored vocal
72 individuality, there are relatively few published examples of how acoustic monitoring of
73 individuals can be applied in practice, particularly for mammalian species.

74 In order to assess animal vocal individuality, acoustic features must first be extracted from the
75 call recordings. Although this can be done automatically using state of the art deep learning
76 techniques (Stowell et al. 2018), acoustic features are often extracted manually, whereby
77 simple summary variables (e.g. min, max and mean) associated with the fundamental
78 frequency and the harmonics are calculated (Fan et al. 2019; Ji et al. 2013). Another common
79 approach involves the calculation of Mel-frequency cepstral coefficients (MFCCs) which
80 effectively warps the frequency axis to the Mel-scale under the assumption that the species
81 perceives frequencies on a logarithmic scale. The MFCC method is based the human auditory

system and is widely used in human-speech recognition, however it is also becoming more popular for the identification of animal vocalisations (Clemins et al. 2005; Mielke and Zuberbühler 2013; Spillmann et al. 2017). In this work we adopt a different and more simple approach that uses short-time Fourier transformation to extract the fundamental frequency contour sequence which is used as the only acoustic feature. Following feature extraction, pattern recognition algorithms are implemented to 'learn' the acoustical features associated with each individual and subsequently classify 'unseen' calls. The recognition models are assessed based on their ability to classify the 'unseen' calls correctly. Several algorithms have been used for this purpose, including discriminant functions (Blumstein and Munos 2005; Fan et al. 2019), artificial neural networks (Mielke and Zuberbühler 2013; Reby et al. 1998), gaussian mixture models (Cheng et al. 2010) and hidden Markov models (Clemins et al. 2005; Ji et al. 2013). One of the main issues affecting the classification performance of vocal recognition models is the presence of background noise. Noise from any natural or anthropogenic source introduces additional sound information and thereby increases model variation and classification uncertainty (Terry et al. 2001). The majority of studies have dealt with this issue by removing samples with overlapping signals (Clemins et al. 2005; Reby et al. 1998).

Lions emit loud, low-pitched vocal signals referred to as 'roars'. A single roar is typically delivered in a bout consisting of one or two soft moans followed by several full-throated roars and a terminating sequence of short grunts (Grinnell and McComb 2001; McComb et al. 1994). The low fundamental frequency of the roar can be attributed to the long and heavy vocal folds, characteristic of the species' vocal anatomy (Weissengruber et al. 2002). Both male and female members of a pride are known to roar and will do so to either maintain contact with distant companions, or to advertise territory ownership (Grinnell and McComb 2001; McComb et al.

1994). Lions also use roars to transmit information relating to group size. A series of overlapping roars emitted by three or fewer individuals provides an honest indicator of the number of individuals present (McComb et al. 1994). Receivers are then able to assess their chances of successfully deterring potential competitors. Several studies have reported evidence of vocal recognition in lions. For example, McComb et al. (1993) found that adult females are able to distinguish between playbacks of familiar, resident males and those of unfamiliar males. Similarly, individual lions have been shown to display stronger responses to situations where there is conflicting information (e.g. a particular call does not match the individual from which it is expected), suggesting that specific calls are associated with specific individuals (Gilfillan et al. 2016). Other studies have examined the attributes of individual roars and have found differences in the temporal pattern and acoustic features between male and female roars, however, little is known about whether consistent differences exist between individuals (Pfefferle et al. 2007; Stander and Stander 1988). This knowledge gap is likely attributable to the difficulties associated with obtaining sufficient samples of roar recordings from known individuals in the wild and highlights the need for a new approach to acquire data on animal vocalisations.

Lions defend large territories, often in excess of 500 km^2 (Tumenta et al. 2013; Zehnder et al. 2018). Given that the primary functions of the roar require receivers to differentiate between individuals in order to respond appropriately, it can be reasoned that the transmission of information relating to caller identity over long distances would be advantageous. It can therefore be expected that selection would favour animal calls that transmit information efficiently by minimizing degradation of the intended signal. Despite many studies being dedicated to lion vocal behaviour, little is known about how lions convey information relating to individual identity in their call structure. In this study, we investigate whether calls from

individual animals might be differentiated based solely on the fundamental frequency of the full-throated roars; the frequency that is likely to suffer the least degradation over large distances. To achieve this, we use a novel acoustic - accelerometer biologger which facilitates the acquisition of a relatively large dataset of lion roar events. We also demonstrate the conservation implications of our findings by testing the application of acoustic recognition and localisation to lion tracking.

Materials and methods

Study site

The study took place in the Bubye Valley Conservancy (BVC), a privately-owned wildlife area located in southern Zimbabwe between latitudes 21.209 and 21.851° South, and longitudes 29.789 and 30.521° East. The BVC measures approximately 3400 km^2 in area and hosts a variety of indigenous megafauna including a high density of African lions (du Preez et al. 2015). Habitat within the Conservancy is dominated by mopane woodland savannah with riparian woodland occurring along several seasonal river lines. Annual rainfall is typically low, averaging 351 mm, and falling mostly in summer between November and March (du Preez et al. 2014). Maximum daily temperatures are high (> 40°C) in the summer months and comparatively mild in winter.

Data collection

Animal-borne biologgers

149 Given that lions vocalise predominantly during the night and can be difficult to identify and
150 observe, we used an on-animal audio recording approach to investigate lion vocal individuality.
151 In November 2014 we fitted custom-designed biologgers to five male and three female lions
152 as described in Wijers et al. (2018). Each device comprised a triaxial accelerometer and
153 magnetometer sampling at 32 Hz per axis and a mono-electret microphone sampling audio at
154 16 Hz with 8-bit resolution (frequency response ~ 30 Hz – 8 kHz; dynamic range ~ 40 dB). All
155 components were encased in an epoxy resin reinforced housing with a hydrophobic vent
156 provided for the microphone. Study animals were chemically immobilised using 75-100 mg
157 Zoletil (Virbac RSA (Pty) Ltd, Halfway House, South Africa) combined with 5 mg
158 medetomidine (Kyron Laboratories, Johannesburg, South Africa). Immobilisation drugs were
159 delivered intramuscularly by 1 cc darts (Pneudart, Williamsport, Pennsylvania, USA) projected
160 from a Dan-Inject CO_2-pressurised dartgun (Dan-Inject, Børkop, Denmark). After fitting the
161 biologger, ~ 25 mg atipamazol (Antisedan, Pfizer Animal Health, Johannesburg, South Africa)
162 was administered to reverse the effects of medetomidine allowing the animal to recover within
163 15 – 90 mins. The biologgers recorded continuous audio (8 bit, 16 kHz mono) for between 4
164 and 10 days before the batteries were depleted. Study animals were then recaptured and the
165 data downloaded for processing and analysis.

166 *Acoustic array*

167 To explore the application of acoustic recognition and localisation for lion tracking, we set up
168 a passive acoustic array system using CARACAL audio recorders (Wijers et al. 2019) within
169 the home-range of a coalition of two male lions fitted with GPS collars (neither of which had
170 been previously fitted with biologgers). Animals were captured using the same capture
171 technique as described in the preceding section. The acoustic array consisted of 8 GPS-time-

synchronised audio recording stations each separated by ~ 500 m from the nearest station. When operational, each station recorded audio continuously (32 bit, 44.1 kHz, 4 channel). The array was activated on selected nights between June and November 2018 when weather conditions were optimal (dry and low wind speed).

Vocal individuality

Roar extraction

All lion audio recordings were processed manually in Audacity 2.1.1 (Audacity Team 2015) by visually inspecting spectrograms and annotating the position of lion roar events. Labelled lion roars were then extracted and classified by individual identity. Roars could be reliably associated with the tagged individual simply by the roar amplitude as roars emitted by other individuals close by had distinctly lower amplitude. In addition, the concurrent accelerometer data provided a further means of validation as lions make consistent head and neck movements while vocalizing (Fig. 1). Roars containing interference (e.g. noise from natural or anthropogenic sources) or overlapping roars of nearby lions were excluded from the analyses. Because the aim of our study was to test for individual differences in the full-throated (FT) roars for each study animal, we further annotated and extracted audio segments containing the first three FT roar units for each roar bout (Fig. 2). As described in several other studies, the FT roars are the high amplitude signals emitted after the initial soft moans and before the sequence of short grunts (Grinnell and McComb 2001; McComb et al. 1994; Stander and Stander 1988).

Fundamental frequency extraction

For each FT roar unit, we first applied a 4th order digital Butterworth bandpass filter function to focus further analyses on the fundamental frequency (*f*0) between 40 Hz and 230 Hz. We extracted the peak frequency contour for this bandwidth (the *f*0 contour) by computing the short-time Fourier transformation (STFT) for the signal using a 2048-point moving Hann window with 68% overlap and zero-padding to four times the window length. The above extraction process was conducted in python using the 'SciPy' library (Jones et al. 2001). Plots of the extracted *f*0 contours are shown in figure S1 in the Supplementary Information.

Fundamental frequency summary features

Following the method used by Ji et al. (2013), we aimed to investigate whether four *f*0 summary features differed between individuals: maximum *f*0, minimum *f*0, mean *f*0 and roar duration. Because the data were non-normal, we used the Kruskal Wallis method to test the hypothesis that the median feature value of least one individual differed from that of at least one other individual. Post-hoc pairwise multiple comparisons were then conducted using Dunn's test with Bonferroni adjustment. In addition to the statistical analyses, we used the k-nearest neighbours (K-NN) algorithm to test the classification ability of models built using the four *f*0 features, individually and combined. The purpose of this test was to measure the discriminative ability of each *f*0 feature and to determine whether the combined feature set was sufficient to accurately discriminate between individuals. Each K-NN classifier was assessed using the overall accuracy metric (proportion of correctly classified data) produced from leave-one-out cross validation where the test sample always included FT roars from the same bout. This data splitting approach was necessary to ensure that the test set was independent of the training set

and thus to avoid any bias resulting from the temporal autocorrelation between roars in a single bout. All statistical analyses were conducted in R (R Core Team 2019).

HMM classification of f0 contour

A Hidden Markov Model (HMM) is a statistical tool for representing the probability distributions over a sequence of observations and is therefore useful for modelling time series data (Ghahramani 2001). An additional benefit of HMMs is their ability to model patterns of varying length as would be expected from variable duration lion roars. We used HMMs to model the temporal pattern of the *f*0 contour for each individual lion. The roar recognition experiment was implemented using the leave-one-out cross validation method whereby 10-state Gaussian HMMs were trained for each individual using all but one sample of the *f*0 sequences. As in the K-NN validation, a sample consisted of FT roar *f*0 sequences from the same bout. The 'unseen' *f*0 sequences were then tested against each HMM by determining which model was most likely to produce the given sequence (using the log-likelihood metric). Classification performance was assessed by calculating overall accuracy, recall (proportion of data of a particular class that is classified correctly as positive) and precision (proportion of correctly predicted positive classifications for a particular class) metrics of classified sequences (Sokolova and Lapalme 2009; Bidder et al. 2014). HMM classification was performed in python using the 'hmmlearn' library (hmmlearn development team 2019)

Lion tracking application

We used an acoustic array to record roars from two known male lions fitted with GPS tracking collars. Identified roars were manually matched across at least 4 stations with the time of arrival

(ToA) of the signal recorded for each roar at each station. The difference in the ToA of the signal between stations allowed us to localise the source of the roar with an accuracy of ~ 100 m using custom-written functions in python. Estimated roar locations were then compared with GPS collar data from the two lions to determine the identity of the caller. An example of the comparison between GPS collar positions and roar localisations is shown in figure 3. In this particular example, the second male lion was not near the array, allowing us to confidently identify the individual vocalizing. For the analysis, we excluded roars that contained noise from other sources (e.g. wind) as wells as roars for which the identity of the caller was uncertain (e.g. when lions roared while in close proximity to one another). To test roar recognition performance, we used audio recorded by the closest station to the lion. Raw 44.1 kHz audio was first down-sampled to 16 kHz, followed by FT roar and *f*0 extraction. We then trained 4-state Gaussian HMMs for each individual and carried out tests using the leave-one-out cross validation method as described previously.

Results

Vocal individuality

We recorded a total of 60 lion-days of audio from the eight lions fitted with biologgers. Manual inspection of the audio revealed a total of 296 roaring bouts from the 5 male lions. The 3 female lions did not appear to roar and only uttered soft moans. Of the total number of bouts, only 78 were useable as the majority contained overlapping roars from conspecifics. From the useable roar bouts, a total of 164 FT roars were extracted (Table 1).

Fundamental frequency summary features

Results of the Kruskal Wallis tests, produced *p* values below 0.001 for all *f*0 features indicating significant differences in max *f*0, min *f*0, mean *f*0 and duration between at least one pair of individuals. Post-hoc analyses using the Bonferroni-Dunn test revealed that *f*0 features did not differ significantly between all individuals (Table 2). Roars emitted by Lion A4, in particular, were found to have significantly lower maximum *f*0 (median = 162.1 Hz), mean *f*0 (median = 122.3 Hz) and duration (median = 1.01 s) compared to other lions. Consistent differences between lions A8, A9, A10 and A11 were not as apparent (Fig. 4).

The analysis of the classification ability of K-NN models built using each of the *f*0 features showed a performance range of between 24.4% (min *f*0) and 50.6% (max *f*0) when single features were used (Table 2) Overall classification accuracy increased to 70.7% when all four features were used together. Optimal performance was achieved using a k value of 7.

HMM classification of f0 contour

Visible differences in the shape of the f0 contour were evident in the FT roars for each lion (Fig. 5). Modelling the temporal pattern of individual *f*0 sequences resulted in high classification performance with an overall accuracy of 91.5% and average recall and precision of 91.0% and 91.7% respectively (Table 3). Four out of the five lions were classified with recall greater than 90%. Recall for lion A10 was found to be lower at 83%. We also investigated classification performance using training set sizes of between 40 (approx. 8 roars per individual) and 160 (approx. 32 roars per individual) roars. Performance was found to increase considerably from 40 to 120 roars and then more gradually beyond 120 roars (Fig. 6). Optimal

performance is therefore likely to be achieved with datasets consisting of 24 or more roars per individual.

Lion tracking application

We extracted a total of 50 FT roars from 31 roaring bouts emitted by the two collared males while they were in the vicinity of the acoustic array. The FT roars were then localised using ToA differences and identity assigned according to corresponding collar GPS data. Results of the HMM classification showed that, using passively recorded audio, we could differentiate between the two individuals with an overall accuracy of 90.0%. Recall was slightly higher for lion B1 compared to lion B2 (Table 4). Vocalisations used for identification originated from distances ranging from 45 m to 1094 m with incorrectly classified calls distributed relatively evenly across the distance spectrum (Fig. 7).

Discussion

Previous research dedicated to African lion vocal communication focused on receiver responses to experimental playbacks with few studies exploring variations in the acoustic structure of calls (Gilfillan et al. 2016; Grinnell and McComb 2001; McComb et al. 1993). While playback-response experiments have undoubtedly provided important insight into the function of vocal signalling and the possible information content of the signal, knowledge of how this information is conveyed has been lacking. In this study, we have demonstrated that the temporal pattern of the fundamental frequency of lion roars is a possible mechanism for lion vocal identification and is useful for automated digital differentiation of individuals.

We acquired our lion roar dataset using a novel acoustic - accelerometer biologging method which made it easy to obtain a sufficient number of roar recordings from each individual and to be able to reliably associate roars with lion identity. When recording animal vocalisations manually, it can be difficult to tell which individual is vocalizing (Clemins et al. 2005). This is particularly true for identifying lions in the wild where the only clues to identity are subtle differences in whisker spot patterns or unique scars (Pennycuick and Rudnai 1970). Furthermore, lions roar mostly during the night which can make it even more difficult to accurately differentiate between individuals. Manual recording approaches are also costly and time consuming and can potentially influence animal behaviour as a result of observer presence.

Results of the statistical analysis exploring differences in roar duration and maximum, minimum and mean fundamental frequency between individuals showed that these factors were not consistently different between all individuals. This suggests that simple overall summary features that do not incorporate temporal variation may not provide optimum discriminatory power to differentiate between individual lions. This was confirmed by the lower classification accuracy of the K-NN model which classified 70.7% of calls correctly, although this is still better than random chance (20%) and suggests that these factors are relevant for acoustic monitoring of the species. Similar classification performance (69.9%) was reported by Ji et al. (2013) for tigers (*Panthera tigris*), using these same four features. Other *f*0 features have been used in vocal identification studies on spotted hyaenas (*Crocuta crocuta*; Mathevon et al. 2010), monkeys (*Rhinopithecus roxellana*; Fan et al. 2019) and wolves *(Canus lupus lycaon*; Root-Gutteridge et al. 2013) with varying success (32%-88%). The roars emitted by lion A4 appeared to be significantly shorter and have lower maximum and mean fundamental frequencies compared to the other four lions. Lion A4, although resident in the

study area, was known to have originated from the Tuli Block in eastern Botswana from where it dispersed more than 60 km before breaking into the fenced Bubye Valley Conservancy. The differences observed for this lion alludes to the possible existence of geographical variation in lion roars. Stander and Stander (1988) also reported that lions from Etosha National Park have shorter roars than lions in other parts of Africa. Geographical variations in vocalisations are known to exist among other species and may occur between different lion sub-populations (Deecke et al. 2000; Martins et al. 2018; Mitani et al. 1992).

Although the use of *f*0 summary features resulted in relatively poor classification performance, when we modelled the temporal pattern of the *f*0 sequence using HMMs, classification performance improved considerably to an overall accuracy of 91.5%. The overall shape of the *f*0 contour is therefore a defining characteristic that allows for better discrimination between individual lions. This finding is consistent with studies on other species which have found that the frequency distribution of vocalisations from each individual are similar but the temporal patterns of the signals are unique to each individual (Clemins et al. 2005; Ji et al. 2013). The key novelty of our approach is in the use of the *f*0 sequence as the only feature for building individual HMM classifiers rather than the more common approach of using MFCC features derived from a range of frequencies which include the harmonics. As shown in figure 8, the higher frequencies of a full-throated lion roar degrade noticeably more than the fundamental frequency which is more consistent across space. Information contained in the fundamental frequency is therefore likely to be preserved over distance and received effectively by a listener. A study on long-distance communication between elephants also showed that the higher formant frequencies of a call are unable to carry information related to individual identity over long distance and therefore identity is more likely to be discerned from the lower frequency harmonics in the 115 Hz region (McComb et al. 2003). Our results support this theory and

thereby provide a likely explanation for the mechanism underpinning individual vocal recognition amongst African lions. We acknowledge, however, that it would be necessary to conduct playback experiments using spectrally modified calls in order to determine the extent to which lions are able to recognise conspecifics purely by the fundamental frequency of their calls. We therefore suggest this as a beneficial avenue for future research in lion communication as well as for other species. From a digital recognition perspective, the accuracy of roar classification models could also be improved by using fundamental frequency extraction algorithms that leverage cepstral transforms and autocorrelation as well as using additional predictive features such as contour slopes.

Understanding how lions are able to recognise conspecifics from their vocal signals is not only important for advancing our understanding of lion ecology, but also provides a useful feature by which lions could be identified using autonomous, acoustic recorders. Traditionally, abundance estimates for lion populations have been achieved using call-in surveys (Ferreira and Funston 2010), spoor transects (Stander 1998) and camera trap surveys (Cusack et al. 2015). Vocal identification of lions may allow for the use of passive acoustic monitoring methods which have the potential to be fully automated using detection and recognition algorithms. Automation would facilitate longer term or even continuous surveys and significantly reduce the number of man-hours required for data management; a challenge that is inherent to camera trapping when individuals need to be manually identified.

Hartwig (2005) highlighted the value of individual vocal identification of African wild dogs as a supplement to radio telemetry and visual recognition. Similarly, one of the primary benefits of acoustic monitoring of lions would be its facilitation of collar-free tracking which is likely to provide movement information for a greater number of individuals at lower cost compared

to animal-borne systems. Using an acoustic array and sound signal localisation techniques, we were able to obtain position estimates for several roars emitted by two resident male lions. The locations of the roars corresponded closely with collar GPS positions which provided reliable ground truth information for assigning identity labels to each roar in order to build individual recognition models. Models trained for each individual were able to accurately classify 'unseen' roars from varying distances. We therefore show that, with sufficient training data, it would be possible to locate and identify lion roars and thereby track individual movements across the landscape using passive acoustics as illustrated in figure 3. Furthermore, by integrating in-network data processing, real-time tracking would be achievable, particularly as the data content of the *f*0 contour is extremely small and therefore easily transmitted over the air. However, despite the roar classification performance being considerably better than chance, we acknowledge that this experiment was conducted for only two individuals, whereas in practice, discrimination between several different lions may be required including out-of-set (unknown) individuals. Obtaining a sufficient number of clean vocal samples from known lions without the use of animal-borne biologgers is a difficult task. As shown in our results, too few training samples can lead to reduced classification performance. This is especially true when attempting to distinguish between signals that have a greater degree of similarity (Trifa et al. 2008). Collecting such data is likely to present the biggest challenge to the development of large-scale, collar-free lion tracking systems. However, with the development of machine learning methods, particularly unsupervised clustering approaches, discriminating between individuals may be possible without the need for manual training (Frasier et al. 2017; Stowell et al. 2019, 2018).

Although we have shown that lion FT roars are individually unique, it is necessary to highlight that fact that lion vocal behaviour can vary across space, time and social circumstances. None

391 of the three tagged female lions in this study emitted full throated roars. This was unexpected,
392 but could be related to the presence of small cubs in the pride (we acknowledge that further
393 work would be required to confirm this assumption). Similarly, other research has shown that
394 resident male lions tend to avoid roaring when outside of their own territories while nomadic
395 males roar at greatly reduced rates or avoid roaring entirely until they have gained prides
396 (Grinnell and McComb 2001). Surveys that rely entirely on roaring may therefore be biased by
397 the fact that not all lions vocalise consistently across the landscape or through time. In addition,
398 communal roaring by several members of a pride is a common occurrence and can make it
399 particularly difficult to dissociate individual roars, resulting in reduced classification
400 performance of recognition algorithms. Another important consideration for long term surveys
401 is the temporal consistency of individual calls. Age-related variations in the calls of other
402 species have been reported (Ey et al. 2007; Ota and Soma 2014). Our study was conducted
403 over a short period and therefore it remains uncertain as to whether the structure of an
404 individual lions' roar varies with age. In addition to these biological caveats and uncertainties,
405 we acknowledge that the hyperparameters used for the HMM and K-NN classification
406 algorithms were selected to provide the best classification performance and could have led to
407 overfitting, however, the leave-one-out cross-validation approach provides an effective
408 measure of generalisation performance.

409 In this study, we have shown that the roar of an African lion contains an individually unique
410 identifier in the form of the fundamental frequency contour. Although this finding has also
411 been reported for other species (Caldwell and Caldwell 1965; Lenhardt 1977), our method of
412 modelling the extracted f0 sequences using HMMs is unprecedented in the field of animal
413 vocal recognition and is likely applicable to a number of species which emit loud, long distance
414 calls. We have also demonstrated the conservation value of our findings through the use of a

passive acoustic array to localise and identify individual lion roars. This work highlights the value of passive acoustic monitoring for African lion research, particularly as current GPS tracking methods, although reliable, are highly expensive and involve invasive capture procedures. With rapidly advancing technological innovations, especially open-source tools, monitoring vocally active wildlife species is likely to become more efficient and affordable, providing new avenues for research and conservation.

Acknowledgements

We thank the John Fell Fund and The Beit Trust for funding this research. We are also grateful to the management staff of the Bubye Valley Conservancy for giving us access to their property. Finally, we acknowledge the assistance provided by Biotrack in manufacturing the biologgers and W. Gush in maintaining the CARACAL array.

Ethical Statement

This study was carried out in accordance with the recommendations of the Use of Animals in Research, ASAB/ABS, with the research protocol approved by the University of Oxford Animal Welfare and Ethical Review Board. Project members attended Zimbabwe's Physical and Chemical Capture of Wild Animals Course and held valid drugs licenses (Dangerous Drugs License No. 2014/16 and 2017/07) permitting the capture and handling of study animals.

Declaration of Interest

No conflict of interest is reported by the authors

Data Availability Statement

The data that support the findings of this study are available from the corresponding author, MW, upon reasonable request.

References

Aubin T, Jouventin P. 2002. How to vocally identify kin in a crowd: the penguin model. Adv Study Behav 31:243-278.

Audacity Team. 2015. Audacity(R): Free Audio Editor and Recorder [Computer application] version 2.1.1. https://audacityteam.org/ [accessed 2015 Nov 20].

Bidder OR, Campbell HA, Gómez-Laich A, Urgé P, Walker J, Cai Y, Gao L, Quintana F, Wilson RP. (2014). Love thy neighbour: automatic animal behavioural classification of acceleration data using the K-nearest neighbour algorithm. PLoS ONE 9:e88609.

Blumstein DT, Munos O. 2005. Individual, age and sex-specific information is contained in yellow-bellied marmot alarm calls. Anim Behav. 69(2):353–361.

Caldwell MC, Caldwell DK. 1965. Individualized Whistle Contours in Bottle-nosed Dolphines (*Tusiops truncatus*). Nature 207(4995):434–435.

Cheng J, Sun Y, Ji L. 2010. A call-independent and automatic acoustic system for the individual recognition of animals: A novel model using four passerines. Pattern Recognit. 43(11):3846–3852.

Clemins PJ, Johnson MT, Leong KM, Savage A. 2005. Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. J Acoust Soc Am. 117(2):956–963.

455 Cusack JJ, Swanson A, Coulson T, Packer C, Carbone C, Dickman AJ, Kosmala M, Lintott
456 C, Rowcliffe JM. 2015. Applying a random encounter model to estimate lion density
457 from camera traps in Serengeti National Park, Tanzania. J Wildl Manage. 79(6):1014–
458 1021.
459 Deecke VB, Ford JKB, Spong P. 2000. Dialect change in resident killer whales: Implications
460 for vocal learning and cultural transmission. Anim Behav. 60(5):629–638.
461 Delport W, Kemp AC, Ferguson JWH. 2002. Vocal identification of individual African
462 Wood Owls *Strix woodfordii*: A technique to monitor long-term adult turnover and
463 residency. Ibis. 144(1):30–39.
464 du Preez B, Hart T, Loveridge AJ, Macdonald DW. 2015. Impact of risk on animal behaviour
465 and habitat transition probabilities. Anim Behav. 100:22–37.
466 du Preez BD, Loveridge AJ, Macdonald DW. 2014. To bait or not to bait: A comparison of
467 camera-trapping methods for estimating leopard Panthera pardus density. Biol Conserv.
468 176: 153–161.
469 Ey E, Pfefferle D, Fischer J. 2007. Do age- and sex-related variations reliably reflect body
470 size in non-human primate vocalizations? A review. Primates. 48(4):253–267.
471 Fan P, Liu R, Grueter CC, Li F, Wu F, Huang T, Yao H, Liu D, Liu X. 2019. Individuality in
472 coo calls of adult male golden snub-nosed monkeys (*Rhinopithecus roxellana*) living in
473 a multilevel society. Anim Cogn. 22(1):71–79.
474 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: Prey and disease effects
475 in Kruger National Park, South Africa Wildl Res. 37(3):194–206.
476 Fitch WT. 1997. Vocal tract length and formant frequency dispersion correlate with body size
477 in rhesus macaques. J Acoust Soc Am. 102(2):1213-1222.
478 Forrest TG. 1994. From Sender to Receiver: Propagation and Environmental Effects on
479 Acoustic Signals. Am J Zool. 36:644–654.

480 Frasier KE, Roch MA, Soldevilla MS, Wiggins SM, Garrison LP, Hildebrand JA. 2017.
481 Automated classification of dolphin echolocation click types from the Gulf of Mexico.
482 PLoS Comput Biol. 13(12):1–23.
483 Ghahramani Z. 2001. An introduction to hidden markov models and bayesian networks. Int J
484 Pattern Recognit Artif Intell. 15(1):9–42.
485 Gilbert G, Tyler GA, Smith KW. 2002. Local annual survival of booming male Great Bittern
486 Botaurus stellaris in Britain, in the period 1990-1999. Ibis 144(1):51–61.
487 Gilfillan G, Vitale J, McNutt JW, McComb K. 2016. Cross-modal individual recognition in
488 wild African lions. Biol Lett. 12(8):7–9.
489 Grinnell J, McComb K. 2001. Roaring and social communication in African lions: the
490 limitations imposed by listeners. Anim Behav. 62(1):93–98.
491 Hartwig S. 2005. Individual Acoustic Identification As a Non-Invasive Conservation Tool: an
492 Approach To the Conservation of the African Wild Dog *Lycaon Pictus* (Temminck,
493 1820). Bioacoustics 15(1):35–50.
494 hmmlearn Development Team, 2019. hmmlearn 0.2.2 library. GitHub repository,
495 https://github.com/hmmlearn/hmmlearn [accessed 2019 May 15].
496 Ji A, Johnson MT, Walsh EJ, McGee J, Armstrong DL. 2013. Discrimination of individual
497 tigers (*Panthera tigris*) from long distance roars. J Acoust Soc Am. 133(3):1762–1769.
498 Jones E, Oliphant E, Peterson P. 2001. SciPy: Open Source Scientific Tools for Python.
499 http://www.scipy.org/ [accessed 2019 May 3].
500 Lenhardt ML. 1977. Vocal contour cues in maternal recognition of goat kids. Appl Anim
501 Ethol 3:211-219.
502 Marler P. 1967. Animal communication signals: We are beginning to understand how the
503 structure of animal signals relates to the function they serve. Science. 157(3790):769–
504 774.

505 Martins BA, Rodrigues GSR, de Araújo CB. 2018. Vocal dialects and their implications for
506 bird reintroductions. Perspect Ecol Conserv. 16(2):83–89.
507 Mathevon N, Aubin T. 1997. Reaction to conespecific degraded song by the wren
508 *Troglodites troglodytes*: territorial response and choice of song post. Behav Processes.
509 39:77–84.
510 Mathevon N, Koralek A, Weldele M, Glickman SE, Theunissen FE. 2010. What the hyena's
511 laugh tells: sex, age, dominance and individual signature in the giggling call of Crocuta
512 crocuta. BMC Ecol. 10(9):1–16.
513 McComb K, Packer C, Pusey AE. 1994. Roaring and numerical assessment in contests
514 between groups of female lions, Panthera leo. Anim Behav. 47:379–387.
515 McComb K, Pusey AE, Packer C, Grinnell J. 1993. Female lions can identify potentially
516 infanticidal males from their roars. Proc R Soc London 252:59–64.
517 McComb K, Reby D, Baker L, Moss C, Sayialel S. 2003. Long-distance communication of
518 acoustic cues to social identity in African elephants. Anim Behav. 65(2):317–329.
519 McGregor PK. 1993. Signalling in territorial systems: a context for individual identification,
520 ranging and eavesdropping. Philos Trans R Soc Lond B. 340:237–244.
521 McGregor PK, Krebs JR, Ratcliffe LM. 1983. The Reaction of Great Tits (*Parus major*) to
522 Playback of Degraded and Undegraded Songs: The Effect of Familiarity with the
523 Stimulus Song Type. The Auk. 100(4):898–906.
524 Mielke A, Zuberbühler K. 2013. A method for automated individual, species and call type
525 recognition in free-ranging animals. Anim Behav. 86(2):475–482.
526 Mitani, J.C., Hasegawa, T., Gros-Louis, J., Marler, P., Byrne, R., 1992. Dialects in Wild
527 Chimpanzees. Am. J. Primatol. 27, 233–243.
528 Mitani JC, Nishida T. 1993. Contexts and social correlates of long-distance calling by male
529 chimpanzees. Anim Behav. 45:735–746.

530 Mouterde SC, Elie JE, Theunissen FE, Mathevon N. 2014. Learning to cope with degraded
531 sounds: female zebra finches can improve their expertise in discriminating between male
532 voices at long distances. J Exp Biol. 217(17):3169–3177.
533 Ota N, Soma M. 2014. Age-dependent song changes in a closed-ended vocal learner:
534 Elevation of song performance after song crystallization. J Avian Biol. 45(6):566–573.
535 Pennycuick CJ, Rudnai J. 1970. A method of identifying individual lion with an analysis of
536 the reliability of identification. J Zool London. 160:497–508.
537 Pfefferle D, West PM, Grinnell J, Packer C, Fischer J. 2007. Do acoustic features of lion,
538 Panthera leo, roars reflect sex and male condition? J Acoust Soc Am. 121(6):3947–53.
539 R Core Team. 2019. R: A Language and Environment for Statistical Computing. R
540 Foundation for Statistical Computing, Vienna. https://www.R-project.org/ [accessed
541 2019 April 26].
542 Ramsauer S. 2005. Acoustic communication in lions and its use in territoriality. Cogn Brain
543 Behav. 9(3):539–550.
544 Reby D, Joachim J, Lauga J, Lek S, Aulagnier S. 1998. Individuality in the groans of fallow
545 deer (*Dama dama*) bucks. J Zool. 245(1):79–84.
546 Root-Gutteridge H, Bencsik M, Chebli M, Gentle LK, Terrell-nield C, Bourit A, Yarnell RW.
547 2013. Identifying individual wild Eastern grey wolves (*Canis lupus lycaon*) using
548 fundamental frequency and amplitude of howls. Bioacoustics. 23(1):55–66.
549 Sokolova M, Lapalme G. 2009. A systematic analysis of performance measures for
550 classification tasks. Inf Process Manag. 45:427–437.
551 Spillmann B, van Schaik CP, Setia TM, Sadjadi SO. 2017. Who shall I say is calling?
552 Validation of a caller recognition procedure in Bornean flanged male orangutan (*Pongo*
553 *pygmaeus wurmbii*) long calls. Bioacoustics. 26(2):109–120.
554 Stander PE. 1998. Spoor counts as indices of large carnivore populations: The relationship

555 between spoor frequency, sampling effort and true density. J Appl Ecol. 35(3):378–385.
556 Stander PE, Stander J. 1988. Characteristics of lion roars in Etosha National Park. Madoqua
557 15(4):315–318.
558 Stowell D, Petrusková T, Šálek M, Linhart P. 2019. Automatic acoustic identification of
559 individuals in multiple species: Improving identification across recording conditions. J R
560 Soc Interface. 16(153):1–13.
561 Stowell D, Wood MD, Pamuła H, Stylianou Y, Glotin H. 2018. Automatic acoustic detection
562 of birds through deep learning: The first Bird Audio Detection challenge. Methods Ecol
563 Evol. 10(3):368–380.
564 Terry A, Peake TM, Mcgregor PK. 2005. The role of vocal individuality in conservation.
565 Front. Zool. 2(10):1–16.
566 Terry AMR, McGregor PK, Peake TM. 2001. A Comparison of Some Techniques Used To
567 Assess Vocal Individuality. Bioacoustics. 11(3):169–188.
568 Trifa VM, Kirschel ANG, Taylor CE, Vallejo EE. 2008. Automated species recognition of
569 antbirds in a Mexican rainforest using hidden Markov models. J Acoust Soc Am.
570 123(4):2424–2431.
571 Tumenta PN, van't Zelfde M, Croes BM, Buij R, Funston PJ, Udo de Haes HA, De Iongh
572 HH. 2013. Changes in lion (*Panthera leo*) home range size in Waza National Park,
573 Cameroon. Mamm Biol. 78(6):461–469.
574 Weissengruber GE, Forstenpointner G, Peters G, Kubber-Heiss A, Fitch WT. 2002. Hyoid
575 apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger
576 (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris* f. catus).
577 J Anat. 201(3):195–209.
578 Wijers M, Loveridge L, Macdonald DW, Markham M. 2019. CARACAL: a versatile passive
579 acoustic monitoring tool for wildlife research and conservation. Bioacoustics. Advance

online publication. doi: 10.1080/09524622.2019.1685408.

Wijers M, Trethowan P, Markham A, du Preez B, Chamaillé-Jammes S, Loveridge A, Macdonald D. 2018. Listening to Lions: Animal-Borne Acoustic Sensors Improve Bio-logger Calibration and Behaviour Classification Performance. Front Ecol Evol. 6(171):1–8.

Wiley RH, Richards DG. 1978. Physical Constraints on Acoustic Communication in the Atmosphere: Implications for the Evolution of Animal Vocalizations. Behav Ecol Sociobiol. 3:69–94.

Zehnder A, Henley S, Weibel R. 2018. Home ranges of lions in the Kalahari, Botswana exhibit vast sizes and high temporal variability. Zoology. 128:46–54.

591 Table 1. Number of bouts and FT roars for each individual lion.

Individual	**A4**	**A8**	**A9**	**A10**	**A11**
Bouts	17	18	16	13	14
FT roars	38	35	40	24	27

592

593

Table 2. Results of Kruskal Wallis tests (Chi-squared and *p* values) and overall accuracy of the K-NN classifiers for each of the four *f*0 features. Multiple comparisons were conducted using Dunn's test with Bonferroni adjusted p-values (significance threshold = 0.05).

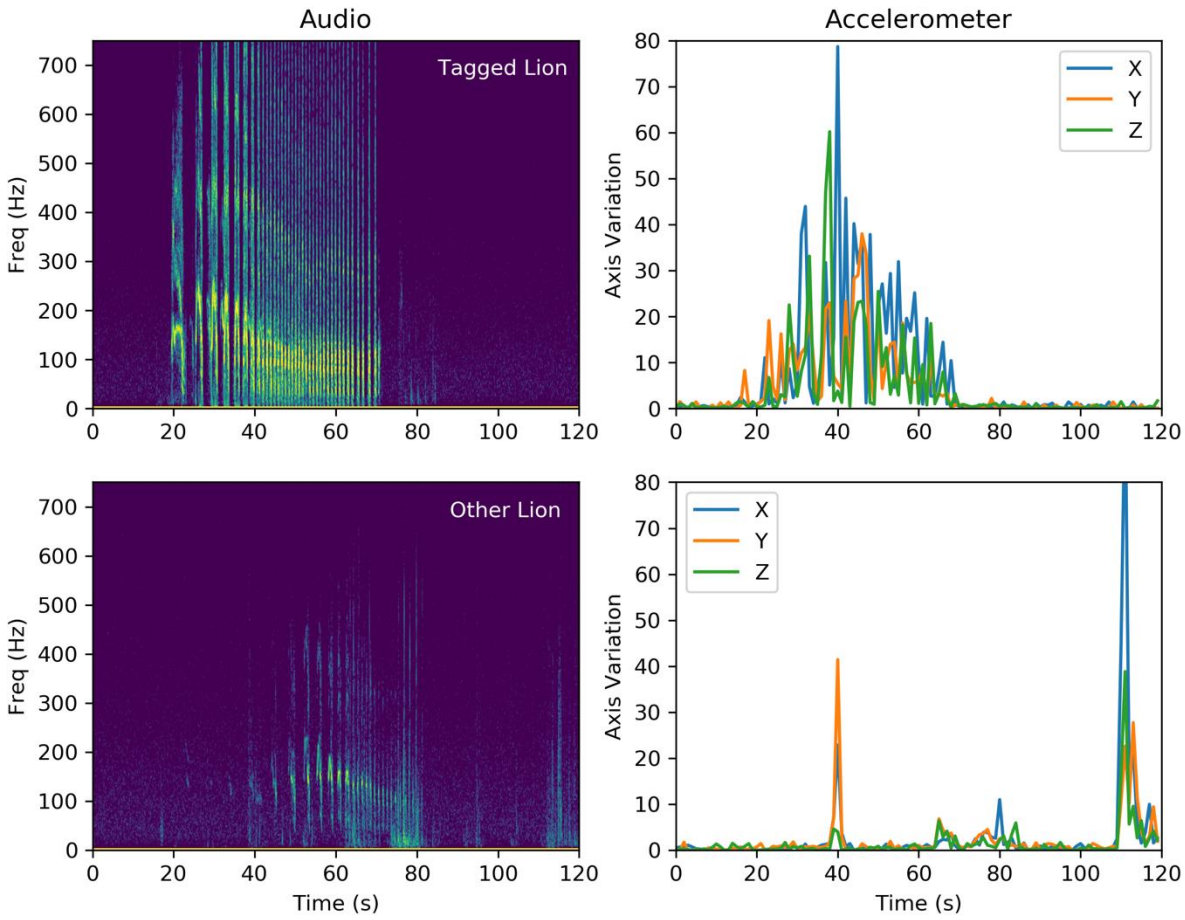
Feature	**K-NN Classifier**	**Kruskal Wallis**		**Pairs that were significantly different**
	Accuracy	χ^2	*P value*	
Min *f*0	24.4	37.6	< 0.001	A9-A10; A10-A11; A10-A4; A9-A8
Max *f*0	50.6	79.4	< 0.001	A9-A10; A9-A4; A10-A4; A11-A4; A9-A8; A8-A4
Mean *f*0	42.1	98.2	< 0.001	A9-A10; A9-A11; A9-A4; A10-A4; A11-A4; A9-A8; A4-A8
Duration	30.5	71.4	< 0.001	A9-A4; A10-A4; A11-A4; A4-A8

Table 3. Confusion matrix of actual lion IDs (rows) vs predicted IDs (columns) for biologger recorded roars using HMMs based on FT roar *f*0 contours. Model training and testing were carried out on separate roar bouts.

Individual	**A4**	**A8**	**A9**	**A10**	**A11**	**Recall %**	**Precision %**
A4	36	0	2	0	0	94.7	100.0
A8	0	33	0	0	2	94.3	84.6
A9	0	2	36	1	1	90.0	92.3
A10	0	2	1	20	1	83.3	95.2
A11	0	2	0	0	25	92.6	86.2
Average						91.0	91.7

600 Table 4. Confusion matrix of actual lion IDs (rows) vs predicted IDs (columns) for passively
601 recorded roars using HMMs based on FT roar *f*0 contours. Training and testing were carried
602 out on separate roar bouts.

Individual	**B1**	**B2**	**Recall %**	**Precision %**
B1	25	2	92.6	89.3
B2	3	20	87.0	90.9
Average			89.8	90.1

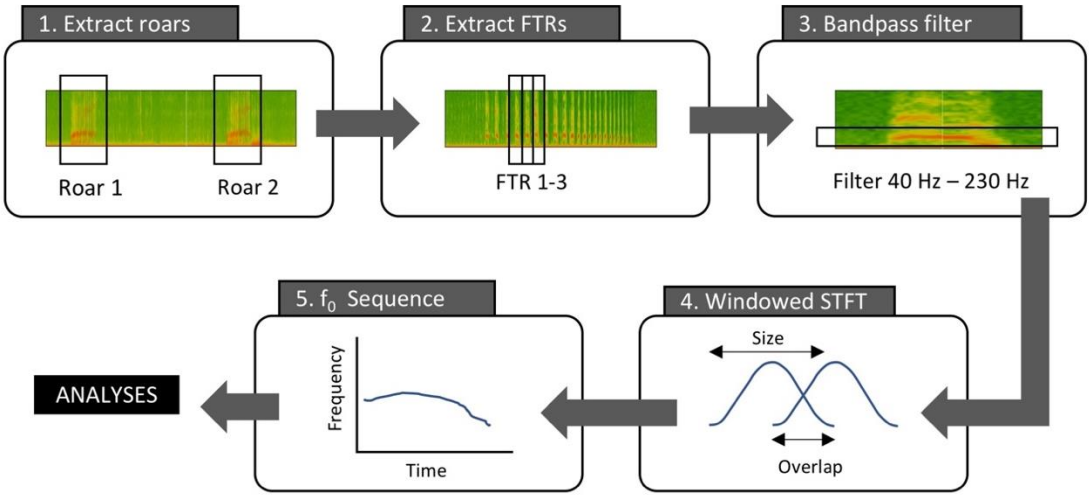


603

604 Figure 1. Concurrent audio (left) and accelerometer (right) data for two different roars recorded
605 on the same biologger, one emitted by the tagged lion (top row) and the other emitted by a
606 neighbouring lion (bottom row). The combination of the audio amplitude and the clear body
607 motion associated with roaring that is recorded by the accelerometer facilitated reliable
608 labelling of roars according to individual identity.

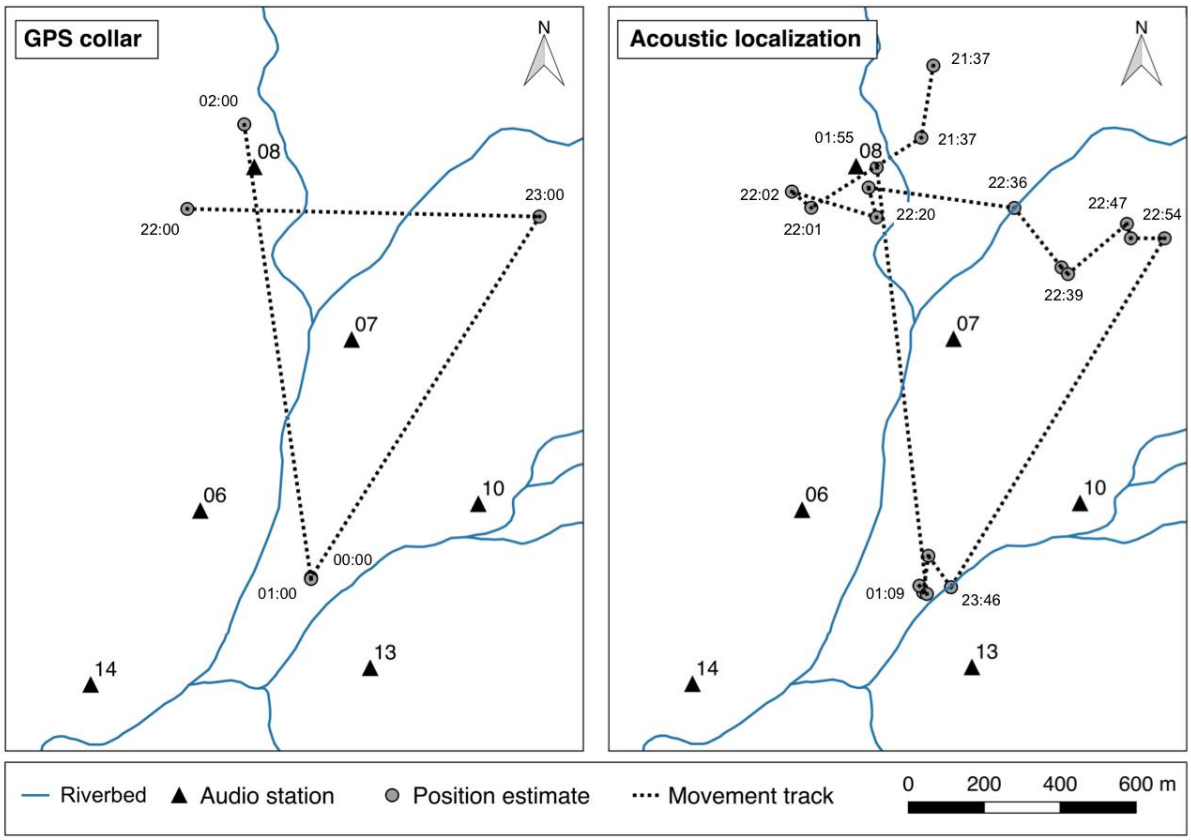
609

610



611

612 Figure 2. Fundamental frequency extraction process.



613

614 Figure 3. Comparison between collar GPS positions (left) and roar localisation (right) for a 4-

615 hour period from a single lion.

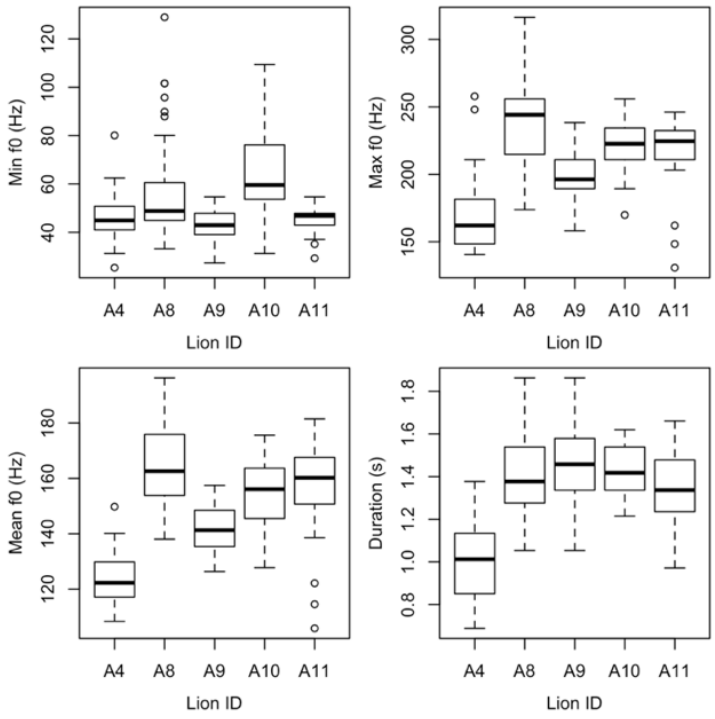
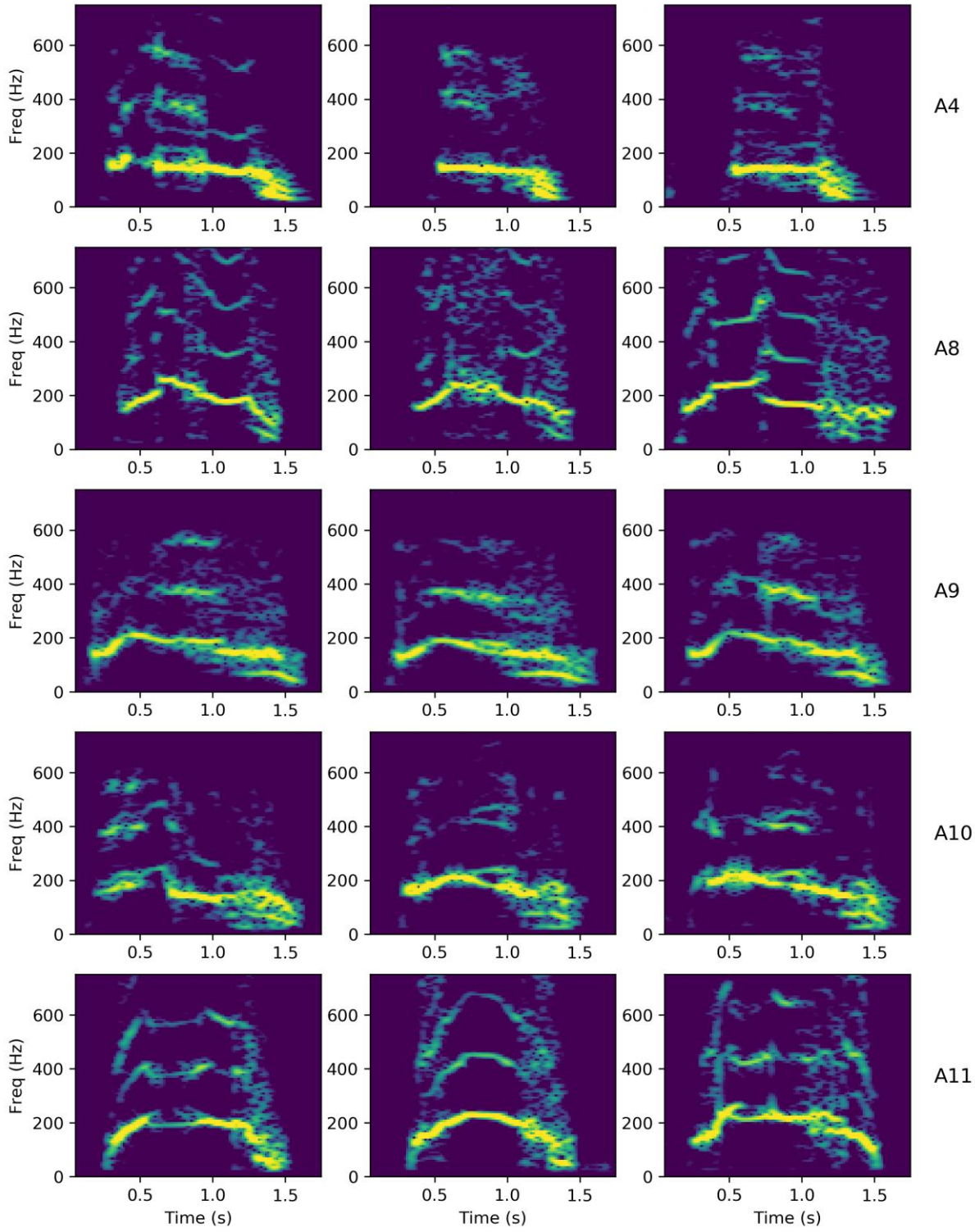


Figure 4. Boxplots of the four *f*0 summary features for each lion.



618

619 Figure 5. Examples of FT roar structure for each individual lion as recorded by the biologgers.

620 Where possible, FT roars are taken from different bouts on different days to highlight temporal

621 consistency.

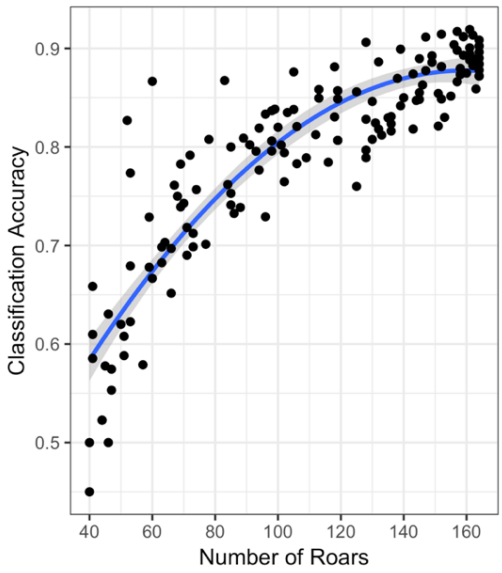
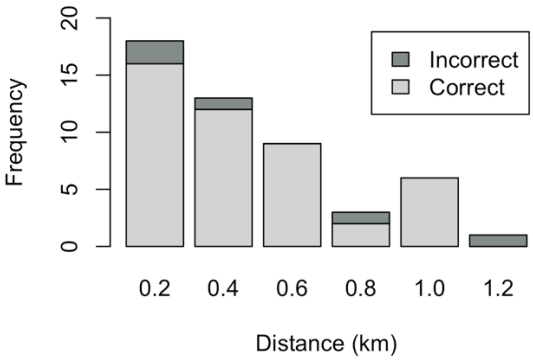
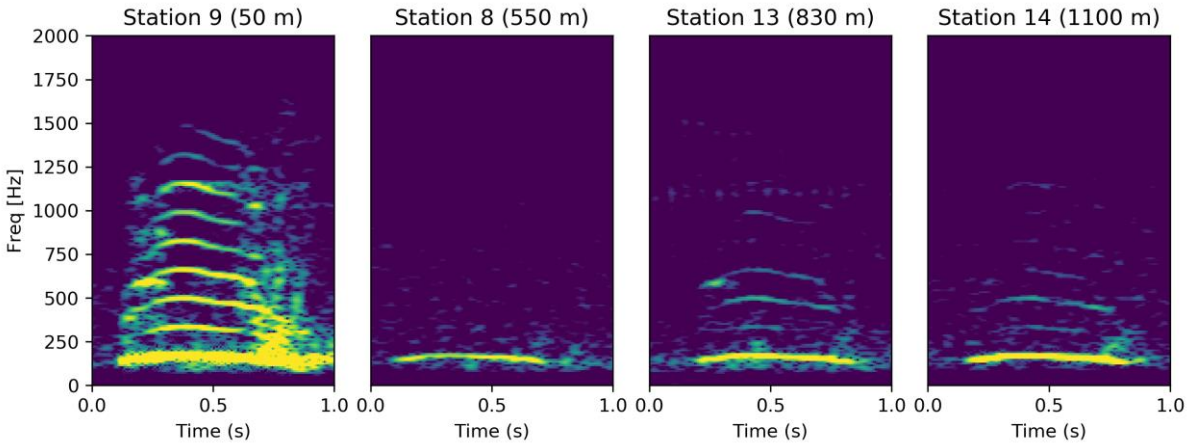


Figure 6. Classification accuracy using an increasing number of training set samples. Shaded region represents 95% confidence interval.



625

626 Figure 7. Distance distribution of correct and incorrectly classified roars



627

628 Figure 8. Spectrograms of the same FT roar recorded at varying distances from the animal.

629 **Figure captions**

630 Figure 1. Concurrent audio (left) and accelerometer (right) data for two different roars recorded
631 on the same biologger, one emitted by the tagged lion (top row) and the other emitted by a
632 neighbouring lion (bottom row). The combination of the audio amplitude and the clear body
633 motion associated with roaring that is recorded by the accelerometer facilitated reliable
634 labelling of roars according to individual identity.

635 Figure 2. Fundamental frequency extraction process.

636 Figure 3. Comparison between collar GPS positions (left) and roar localisation (right) for a 4-
637 hour period from a single lion.

638 Figure 4. Boxplots of the four *f*0 summary features for each lion.

639 Figure 5. Examples of FT roar structure for each individual lion as recorded by the biologgers.
640 Where possible, FT roars are taken from different bouts on different days to highlight temporal
641 consistency.

642 Figure 6. Classification accuracy using an increasing number of training set samples. Shaded
643 region represents 95% confidence interval.

644 Figure 7. Distance distribution of correct and incorrectly classified roars.

645 Figure 8. Spectrograms of the same FT roar recorded at varying distances from the animal.