Knowing a fellow by their bellow: acoustic individuality in the bellows of the American alligator

Thomas Rejsenhus Jensen[*] , Andrey Anikin , Mathias Osvath , Stephan A. Reber

The Cognitive Zoology Group, Department of Philosophy and Cognitive Science, Lund University, Sweden

Identity cues in animal calls are essential for conspecific vocal individual recognition. Some acoustically active species mainly show reliable identity cues in their vocalizations because of variation in anatomy and life history. Long and strenuous-to-produce vocalizations may be particularly effective for showing identity cues because sustaining such calls may reveal individual anatomical differences in sound production. It is largely unknown whether reptiles possess acoustic individuality despite some groups being vocal. We analysed 814 bellows from 47 American alligators, *Alligator mississippiensis*, extracting spectral characteristics and manually corrected contours of the fundamental frequency. Recognition was up to 66% correct with a supervised classifier (random forest) and 61% with unsupervised clustering (chance = 2.1%), indicating that individual alligators have highly distinct bellows. Alligators were distinguished primarily based on the call spectrum, fundamental frequency contour and amplitude modulation, which also provided information about the animal's size. Neither manual supervision of acoustic analyses nor supervised training on labelled data was necessary to achieve reasonable accuracy, which has promising potential for identification of individuals via passive acoustic monitoring for research and conservation purposes. Additionally, our results highlight the importance of studying the utilization of acoustic individuality in the social lives of crocodylians.

© 2023 The Author(s). Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Being individually recognizable carries fitness benefits for many animals (Carlson et al., 2020; Tibbetts & Dale, 2007). Possessing cues to individuality is particularly adaptive in environments with repeated social interactions (Tibbetts & Dale, 2007), in agonistic and affiliative contexts alike (Carlson et al., 2020; Temeles, 1994; Tibbetts, 2004). Recognizing individuals by their vocalizations requires sufficient variability in acoustic parameters of a call type between individuals, as well as consistency in call production within individuals (Lengagne et al., 1997; Mitani et al., 1996; Yorzinski, 2017). It has been shown that the calls of many mammals, birds and amphibians harbour vocal identity cues (Bee et al., 2001; Carlson et al., 2020; Hambalkova et al., 2021; Sèbe et al., 2018; Thomsen et al., 2019; Wierucka et al., 2021). However, it is largely unknown whether reptiles also possess acoustic cues to individual identity, apart from three individuals of a single species of gecko (Rohtla et al., 2019).

A better understanding of vocal identity cues across taxa is particularly desirable because of their value for conservation efforts

(Terry et al., 2005). If such cues are objectively measurable and do not require direct visual observation, they can be used for remote acoustic monitoring, that is, recording the vocalizations in a given area over a longer period of time without direct human oversight, and subsequently estimating the number of callers (Marques et al., 2013). However, to construct such monitoring systems, one must first establish what makes a vocalization individually specific, and thereafter test which features can be used by automated recognition software (Linhart et al., 2022). Various statistical methods and machine learning algorithms already aid in identifying calls of specific species (Gradisek et al., 2017; Parsons & Jones, 2000) as well as individuals within species (Brown et al., 2010; Clemins et al., 2005; Thomsen et al., 2013). Additionally, deep learning artificial neural networks have been shown to be adept at identifying and differentiating between species and/or individuals (Ferreira et al., 2020; Martin et al., 2022; Parsons & Jones, 2000; Trapanotto et al., 2022), and they may prove especially useful for analysing large amounts of acoustic data with varying levels of noise and other disturbances.

Crocodylia is an order of reptiles whose socioecology might have favoured the evolution of acoustic individual recognition, and which could also benefit from conservation efforts utilizing vocal

* Corresponding author.
E-mail address: thomas_rejsenhus.jensen@lucs.lu.se (T. R. Jensen).

identity cues for acoustic monitoring. Crocodylia includes the families Alligatoridae (alligators and caimans), Crocodylidae and Gavialidae. They are all highly vocal for reptiles (Reber, 2018) but are considered to have limited capacities to fine-tune their vocalizations (Riede et al., 2015). Wild crocodylians repeatedly interact with many different rivals, mates and affiliates over decades (Grendeus & Reber, 2020; Grigg & Kirshner, 2015). Hence, vocal cues to identity are arguably valuable to them. Despite this, only one study has searched for individuality cues in crocodylian vocalizations (Vergne et al., 2007): the short contact calls of hatchling Nile crocodiles, *Crocodylus niloticus*, showed a weak individual signature, which the authors deemed unsuitable for individual recognition. However, it is possible that such weak signatures may become more pronounced as individuals grow older. Additionally, the call repertoire of crocodylians is not fully developed until close to the age of sexual maturity when 'bellows' (Alligatoridae), or 'roars' (Crocodylidae), emerge (Grigg & Kirshner, 2015). These two vocalizations have different names between the two families but are functionally equivalent. These are the most promising vocalizations for harbouring individual signatures.

The bellow is a call type produced by all alligators of either sex, and is thought to mainly serve as an advertisement call announcing an individual's presence (Reber, 2018). This signal may contain individuality cues in its frequency contour as well as in other acoustic parameters due to its advertising role and because it is an extremely loud vocalization, which drives the vocal production system to its limits. Bellows might therefore highlight individual differences in vocal apparatus anatomy. To our knowledge, no studies have looked at acoustic cues to identity in adult crocodylians. The possibility of bellows harbouring individually distinct acoustic cues was suggested by Garrick et al. (1978), who could identify several individuals by their calls, without visual access, and saw consistent differences between spectrograms of different individuals, albeit without any further analysis. Bellowing has been studied most extensively in the American alligator, *Alligator mississippiensis* (Reber, 2018) and contains information on the size of the caller (Reber et al., 2017). The vocal production mechanism of this display is inferred from observations and secondary analyses. For instance, alligators appear to actively pull their larynx to the sternum during the bellow to maximize the length of their vocal tract (Reber et al., 2017). This results in lower vocal tract resonances (Titze, 1994b), a behaviour also documented in some mammals, such as koalas, *Phascolarctos cinereus*, and red deer, *Cervus elaphus* (Charlton et al., 2012; Reby et al., 2005). It leads to an exaggeration of one's body size because the length of the vocal tract correlates with size within species. However, due to anatomical constraints, such a signal remains honest (Maynard Smith & Harper, 2003; Reby & McComb, 2003). Formants in bellows are good predictors of body size in American alligators of both sexes (Reber et al., 2017) in contrast to the dominant frequency, which seems to be primarily caused by tissue vibrations (Reber et al., 2015). The bellowing display of American alligators encompasses a series of behaviours (see Fig. 1 for a detailed description; Garrick & Lang, 1977; Vliet, 1989). The display is commonly performed in shallow water but may also be produced on land. It is thus likely that the strain imposed on a bellowing alligator may differ between land and water because of different bellowing poses and balancing needs. This may lead to differences in acoustic structure between land and water bellows. Additionally, it is likely that bellows contain information about sex, as male displays contain an additional component, the so-called water dance (Fig. 1; Garrick et al., 1978). This makes water an important part of the male display, at least for the visual component.

In addition to parameters indicating sex and size, we expect bellows to possess several characteristics that could showcase cues

to identity. The acoustic parameters that make a vocalization individually distinct differ between species, but are in most cases frequency-related features, such as the fundamental (F0, the rate of vocal fold vibrations) and dominant frequency (the frequency band carrying the most energy). Examples of such individuality cues have been found in all vocally active terrestrial vertebrate lineages, in amphibians, mammals, birds and a single reptile species (Bee et al., 2001; Carlson et al., 2020; Hambalkova et al., 2021; Rohtla et al., 2019; Sèbe et al., 2018; Thomsen et al., 2019). Bottlenose dolphins, *Tursiops truncatus*, even acquire their identity signals through learning and encode this information in the time−frequency modulation of the fundamental frequency of their signature whistles (Janik et al., 2006). In addition to frequency parameters, features such as spectral energy distribution and amplitude modulation provide cues to identity in a variety of species including rooks, *Corvus frugilegus*, peach-fronted parakeets, *Eupsittula aurea*, and giant pandas, *Ailuropoda melanoleuca* (Benti et al., 2019; Charlton et al., 2009; Thomsen et al., 2013). For some species, however, temporal cues, rather than frequency, make a call individually specific: call elements in a sequence can have individually distinct onset intervals. This may also include varying the
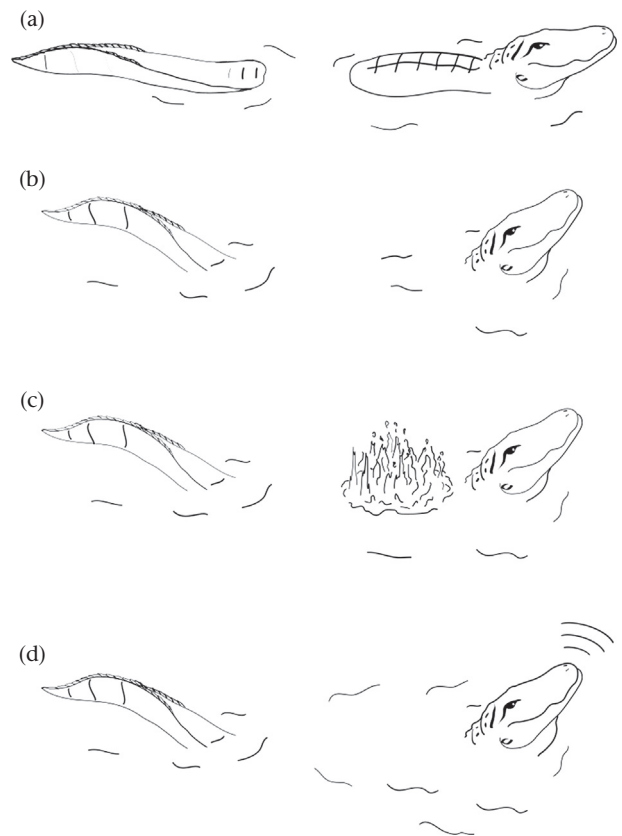
Figure 1. The bellowing display of a male American alligator. (a) When an alligator bellows in shallow water, it usually chooses a spot where it can submerge its back, but still touch the ground with its feet. (b) The alligator will then adopt the so-called HOTA position (head oblique, tail arched), which raises the tip of the snout as high out of the water as possible; the tail is simultaneously raised to stabilize the body (on land, the head is raised similarly but the tail remains on the ground). Then, the animal audibly inhales and pumps large volumes of air into its lungs. (c) Next, the alligator lowers itself into the water while the head stays raised until its back is submerged (at this stage on land, the front limbs stop pushing up the chest and the animal slightly pivots forward). Then, strong muscle contractions run through the flanks (can be visually observed, but most prominent in large males), which cause subaudible vibrations (SAVs). In males, this leads to the 'water dance', a forceful and prolonged displacement of water. (d) The SAVs now transition into the vocal part of the display: the bellow, a low-pitched, loud, and extended rumbling sound.

sequences of short elements within the same call type. This form of individuality encoding is more likely to occur in species with elaborate control over vocal production, particularly passerine birds. For example, male black-capped chickadees, *Poecile atricapillus*, have an individually distinct structure in their 'fee-bee' song (Christie et al., 2004). Because there is no evidence for advanced vocal control in American alligators, we would expect acoustic cues to individuality to be more related to anatomy than to the control of vocalizations. As such, frequency-related features, as well as amplitude modulation, may be good candidates as these are found in other species with similar vocal anatomy (Charlton et al., 2009; Riede et al., 2015; Sèbe et al., 2018).

In nonavian tetrapods, the frequency contour of a vocalization is affected by multiple factors such as an animal's subglottal pressure, lung volume, muscular strength acting on the arytenoids (which are embedded in the vocal folds), consistency of the vocal folds, the endurance to maintain the calling body posture and many more (Fant, 1960; Taylor & Reby, 2010; Titze, 1994a). It is a reasonable assumption that these factors of sound production could provide cues to identity due to 'by-product distinctiveness' between individuals based on genetics and differing growth curves, as well as diseases and injuries sustained throughout life (Boughmann & Moss, 2003; Wyman et al., 2022). To more clearly emphasize a large number of such individuality cues, a vocalization would ideally be long in duration, strenuous to produce and incorporate larger movements of the entire body (McComb & Reby, 2009). Finally, a species would be expected to have more accentuated individual cues if they are long lived and have a prolonged ontogeny as well as allometric growth, which would allow for many chances for their vocal anatomy to be affected individually (Basken et al., 2012; Rendall et al., 2005; Sataloff, 2017). The crocodylians fulfil all these criteria (Grigg & Kirshner, 2015).

In this study we investigated 814 bellows from 47 American alligators for cues to individuality. We used unsupervised (UMAP with hierarchical clustering) and supervised (random forest) classifications to test whether bellows are individually distinct. We investigated which aspects of a bellow carry the strongest cues to individuality, whether bellows indicate the sex of the caller, and whether bellows performed in water versus on land differ in their acoustic structure. Finally, we tested whether our approach could be used for passive acoustic monitoring by comparing classifications that were either informed by, or naïve to, the identity of the call providers.

METHODS

Data Collection

The American alligators were kept at two facilities (49 individuals in total). We recorded 636 bellows from 43 alligators in 2013 at the St Augustine Alligator Farm Zoological Park (Florida, U.S.A.). Another 178 bellows were collected in 2021 from six alligators at the Lund University Alligator Research Station (Sweden). In St Augustine, unprovoked bellows were recorded between 0730 and 1100 hours from March to May, while in Sweden alligators could be reliably encouraged to bellow on demand (by knocking against the wooden walls of the facility) and were thus recorded up to twice a day in the late morning and later afternoon across 7 days.

The alligators were recorded at 2—10 m from the microphone. The identity of the recorded individuals was established both prior to and during the recording sessions by at least two independent observers and was also video recorded. Calls from both locations were recorded at 44.1 kHz sampling frequency and 16-bit amplitude resolution using a Sennheiser ME66 shotgun microphone (frequency response 40 Hz—20 kHz ± 2.5 dB) with a K6 powering

module connected to a Zoom H4n Pro Audio Recorder with an XLR cable. Recordings were saved in .wav format. Absolute amplitude ('loudness') of the bellows could not be documented as this would have required a standardized recording distance and direction (including calibrations with a sound pressure level meter).

The body size of the alligators (total length and head length), in both localities, was known from measurements done during veterinary check-ups (for detailed methods see Reber et al., 2017).

Ethical Note

The procedures performed in this study were all in accordance with the ASAB/ABS guidelines (ASAB Ethical Committee/ABS Animal Care Committee, 2023). The present study was observational and noninvasive, involving recording of vocalizations only. The procedures complied with the guidelines of the Swedish Board of Agriculture (observational research is exempted from ethical approval, SJVFS 2019:9 §22) and the data collection was approved by the St Augustine Alligator Farm research committee in April 2013. The body size of the alligators was measured during veterinary check-ups unrelated to the present study and during routine keeper interactions with the subjects. No animal was restrained for the purpose of this study.

General Acoustic Analysis

A total of 814 calls from 47 individuals (16 females, 31 males) were extracted for acoustic analysis. For two alligators (of the originally 49), we recorded fewer than six bellows and decided to exclude them. Individual bellows were cut out from the recordings into separate audio files for acoustic analysis. The individual bellows were subsequently resampled to 10 kHz, to remove irrelevant high-frequency background noise, and had their peak amplitude normalized to the maximum possible value.

The fundamental frequency (F0) contours of each call was manually corrected using the 'pitch_app()' function in the soundgen package (Anikin, 2019) in R 4.2 (R Core Team, 2020), which performs pitch tracking with manual supervision. The ability to manually correct errors considerably improved the accuracy of pitch tracking because alligator bellows are low pitched, often noisy, and contain a variety of nonlinear phenomena that complicate automatic pitch tracking. The acoustic analysis itself was done using the 'analyze()' function in the soundgen package. For the analysis, we used a window length of 400 ms, step 100 ms, and the manually corrected F0 contours. Amplitude modulation was analysed in the range of frequencies from 5 to 20 Hz, and frequency modulation 1 to 5 Hz based on manual inspection of typical alligator bellows. In addition, we extracted mel-spectrograms and mel-frequency cepstral coefficients (MFCCs) with the tuneR package (Ligges et al., 2018), using a window of 40 ms, a step of 10 ms, 32 filters and no pre-emphasis. Although MFCCs are traditionally used in speech recognition, they provide a very efficiently coded, sparse representation of the spectrum of an entire bellow or some portion thereof.

Unsupervised Classification

We began by obtaining a distance matrix between all recordings based on multivariate dynamic time warp (DTW) of the feature matrix. The features consisted of eight biologically and acoustically meaningful spectral descriptives (Table 1) and 31 MFCCs. DTW matches time series allowing for some time shift or distortion: for instance, an imperfect imitation of a musical tune will still produce a close match despite some irregularities in timing and intonation. DTW was performed with R package dtw (Giorgino, 2009), and the output was a measure of normalized acoustic distance between

Table 1
Acoustic measures

Name	Definition	Meaning
AM depth	Relative strength or depth of amplitude modulation relative to the amplitude of a smoothed amplitude envelope, 0 to 1	Strongly modulated alligator bellows sound trill- or purr-like
AM frequency	Frequency of amplitude modulation defined as the highest peak in the spectrum of the smoothed amplitude envelope, (Hz)	AM frequency measures how slowly or rapidly the alligator is 'purring'
Average f_0	Average fundamental frequency or pitch, Hz	Like musical tones, alligator bellows can be low or high in pitch
Duration	Duration, s	Bellows can be shorter or longer in duration
Harmonics-to-noise ratio	Relative amplitude of the autocorrelation function, a measure of how tonal or noisy a sound is, (dB)	The tone can be relatively clear (high harmonicity), or it can be rough and noisy (low harmonicity)
Intensity SD	Standard deviation of the root-mean-square amplitude, 0 to 1	This variable shows how much the loudness of a bellow varies over time
Peak frequency	Frequency bin with maximum spectral energy, Hz	Peak frequency corresponds to the most prominent frequency component in a bellow
Spectral centroid	Centre of gravity of the spectrum, Hz	Spectral centroid shows the overall balance between high and low frequencies
Voiced frames	Proportion of voiced frames, 0 to 1	Some bellows are mostly voiced, while others have longer unvoiced components such as some 'purring' before the voiced proportion begins
MFCCs	Mel-frequency cepstral coefficients	Sets of MFCCs calculated at different times provide a sparse representation of the spectrogram. When averaged over an entire bellow, MFCCs capture the overall shape of its spectrum

two bellows. The dimensionality of the resulting distance matrix was reduced with the state-of-the-art algorithm of uniform manifold approximation and projection (UMAP) implemented in R with the uwot package version 0.1.11 (Melville et al., 2020). UMAP is conceptually similar to traditional methods such as principal component or multidimensional scaling, but it is based on topological analysis, uncovers geometrically complex clustering, preserves global structure and within-cluster distances, and scales well to large data sets, making this method a popular choice in bioacoustics (Sainburg et al., 2020).

The two-dimensional UMAP projection was plotted (Fig. 2) and used as input for hierarchical clustering. We performed hierarchical clustering with the dbscan package (Hahsler et al., 2019) with the minimum number of points set to six (the same as the smallest number of calls per individual in the final analysis). The number of clusters was consistently smaller than the true number of

individuals to ensure that performance would not be inflated due to splitting all observations into unique clusters. Cluster purity was defined as the proportion of recordings correctly assigned to their own cluster, that is, to the cluster in which this alligator was the most common one (0 = pure noise, 1 = perfect clustering). Interestingly, clustering the DTW-based distance matrix (described in the first paragraph of this section) directly, without passing it through UMAP, resulted in a considerably poorer clustering performance. Likewise, clustering purity decreased if we used Euclidian distances between MFCC matrices without DTW. The presented pipeline for unsupervised classification (DTW − UMAP − hierarchical clustering) is thus the one that in this case resulted in the best separation between individuals.

We explored several possibilities in terms of the number of MFCCs and timeframes, as well as using complete melspectrograms instead of MFCCs. Optimization was performed for
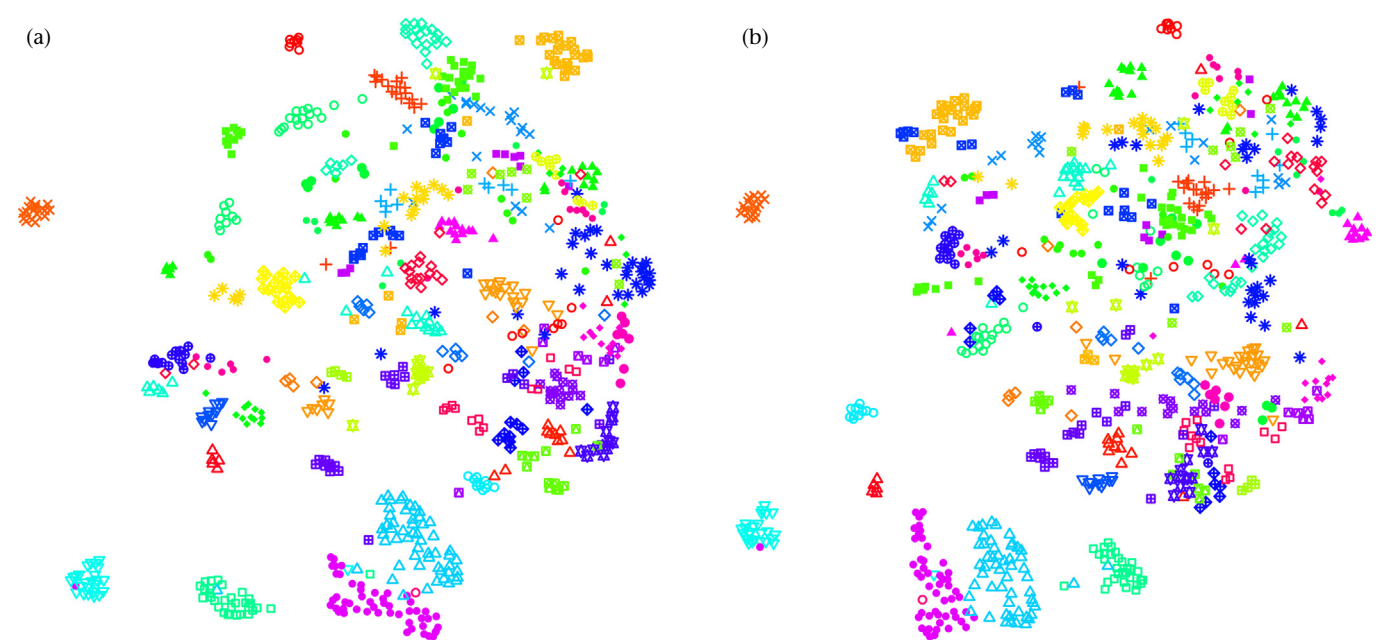

Figure 2. Unsupervised classification of 814 bellows from 47 animals produces comparable clustering both (a) with and (b) without acoustic features that were verified manually, namely voicing and pitch. Each point represents one bellow; the colour and shape encode the true identity of the caller.

unsupervised classification based on cluster purity from two-dimensional UMAP embedding of MFCC matrices only (without including other spectral features or performing DTW). Thirty-one MFCCs and 10 timeframes provided good resolution without overextending the number of predictors to unmanageable levels for a data set of this size (814 recordings).

Supervised Classification

All supervised classification, whether by individual or location (on land versus in water), was performed with a random forest classifier (Breiman, 2001). We considered various acoustic predictors of individuality. Manually verified F0 contours were time-standardized to 10 points per call, thus providing 10 features. Spectral descriptors provided another eight features (Table 1), followed by MFCCs, which could be time-averaged across the entire vocalization ($N = 31$) or time-standardized to a fixed number of frames per vocalization (we used $N = 31 \times 10 = 310$). Because the number of recorded calls per animal varied greatly (from 6 to 54), we calculated balanced accuracy by averaging the true positive rates across all animals. Balanced accuracy is a popular performance metric in bioacoustics (Arnaud et al., 2023) because it ensures that recognition accuracy is not inflated due to better recognition of the most common individuals. For example, in a data set with one very common and two very uncommon classes, the overall hit rate may be close to 100% if the algorithm always picks the most common class. However, averaging hit rates across the three classes would bring the performance down to the chance level of 1/3. Note that, throughout the text and merely as an intuitive benchmark, we report chance levels calculated simply as 1/number of classes (e.g. number of individuals); more appropriate odds ratios for original versus scrambled data sets are reported in our analysis coding scripts (see Data Availability statement below).

Each random forest model was rerun 1000 times with randomly selected training and testing samples. At each iteration, two-thirds of the available number of calls per individual were used for training, and one-third for testing the algorithm. The results were summarized as the mean balanced accuracy and 95% coverage interval (95% CI). To account for any temporal autocorrelation, bellows from the same animal—date combination (calls recorded on the same day from the same individual) were put in either the training set or the testing set at each iteration, but never in both. Recording conditions may vary day by day, introducing a confound in the sense that accidental day-specific acoustic features (e.g. characteristic background noise) may inflate classification accuracy (Lehmann et al., 2022). Accordingly, we trained and tested the classifiers on recordings from different days. The bellows from individuals used in the analysis had between 2 and 11 days of recordings (mean = 6.3 days, median = 6). Including two localities in our analysis (U.S.A. and Sweden) could, in theory, artificially inflate the classification accuracy of our algorithm, because the different recording environments, not individual characteristics, would have led to acoustic differences. We therefore ran our analyses twice, with or without the smaller Swedish population, to assess whether the classification accuracy would be affected.

Number of Recordings Required per Subject and Effect of Location

To test how many recordings per animal were needed to achieve robust recognition, we fitted a binomial regression model predicting the proportion of correctly identified bellows per individual as a function of the number of recorded bellows. The accuracy data were derived from the main random forest model (using absolute F0 contours, eight time-averaged spectral features and 31 time-averaged MFCCs as predictors), taking accuracy per animal in the

test data set from 1000 simulations. This Bayesian multilevel model was fitted with a random intercept for each individual and its predictions are shown in Fig. 3d.

To test whether location, on land or in water, could interfere with individual recognition, we selected 13 alligators that produced at least six bellows in both locations and trained the same random forest classifier with training and testing sets segregated by location. We then compared these testing and training sets with sets of mixed locations.

The Main Acoustic Predictors of Individuality

Some acoustic characteristics of bellows, for example formants, may contain individuality, but require manual scoring of each vocalization and adjustments to the frequency window in which to expect a certain number of formants, based on the estimated length of an individual's vocal tract (Reber et al., 2017). In the current study we instead focused on readily measurable and meaningful properties of the voiced part of each bellow, particularly the contour of F0 and eight other acoustic descriptors (Table 1). To determine what acoustic features best enabled individual recognition, we fitted a series of random forest models with different combinations of acoustic predictors (Table 2). To distinguish between the effects of average F0 and its dynamics, we also normalized intonation contours per bellow to semitones (12 semitones = 1 octave) above or below the mean.

Acoustic Markers of Sex and Size

Finally, we wanted to test for differences in specific acoustic predictors depending on the location (land or water), sex and body length of alligators. To do so, we fitted three multivariate mixed models, each with a random intercept for each individual predicting individual acoustic features as a function of location, body length or sex (Fig. 4). The acoustic features were normalized to have a mean of 0 and SD of 1 to put all changes on the same relative scale.

RESULTS

Unsupervised Classification

The most successful unsupervised algorithm measured differences between pairs of recordings with the DTW, which is a technique that allows for partial temporal misalignments and distortions when comparing time sequences such as intonation contours. Here we applied DTW not only to F0 contours, but also to eight other spectral features (Table 1) and MFCC matrices, which can be seen as a sparse representation of the entire spectrogram. When visualized with UMAP, a powerful method for dimensionality reduction, the resulting distance matrix produced visually obvious clustering by individual (Fig. 2). We confirmed this formally with hierarchical clustering, which resulted in a cluster purity of 0.61 (Fig. 2a), confirming that fully automated analyses should be adequate for unsupervised classifications of bellows by the individual.

Supervised Classification

A series of random forest models for classifying 47 individual alligators achieved out-of-sample accuracy of up to 66% when using different combinations of intonation contours (Fig. 3b), spectral descriptors and MFCCs as predictors (Fig. 3c, Table 2). Importantly, we accounted for temporal autocorrelation by training and testing the classifiers on bellows recorded on different days. Further, the inclusion of bellowing data from two different localities (U.S.A. and
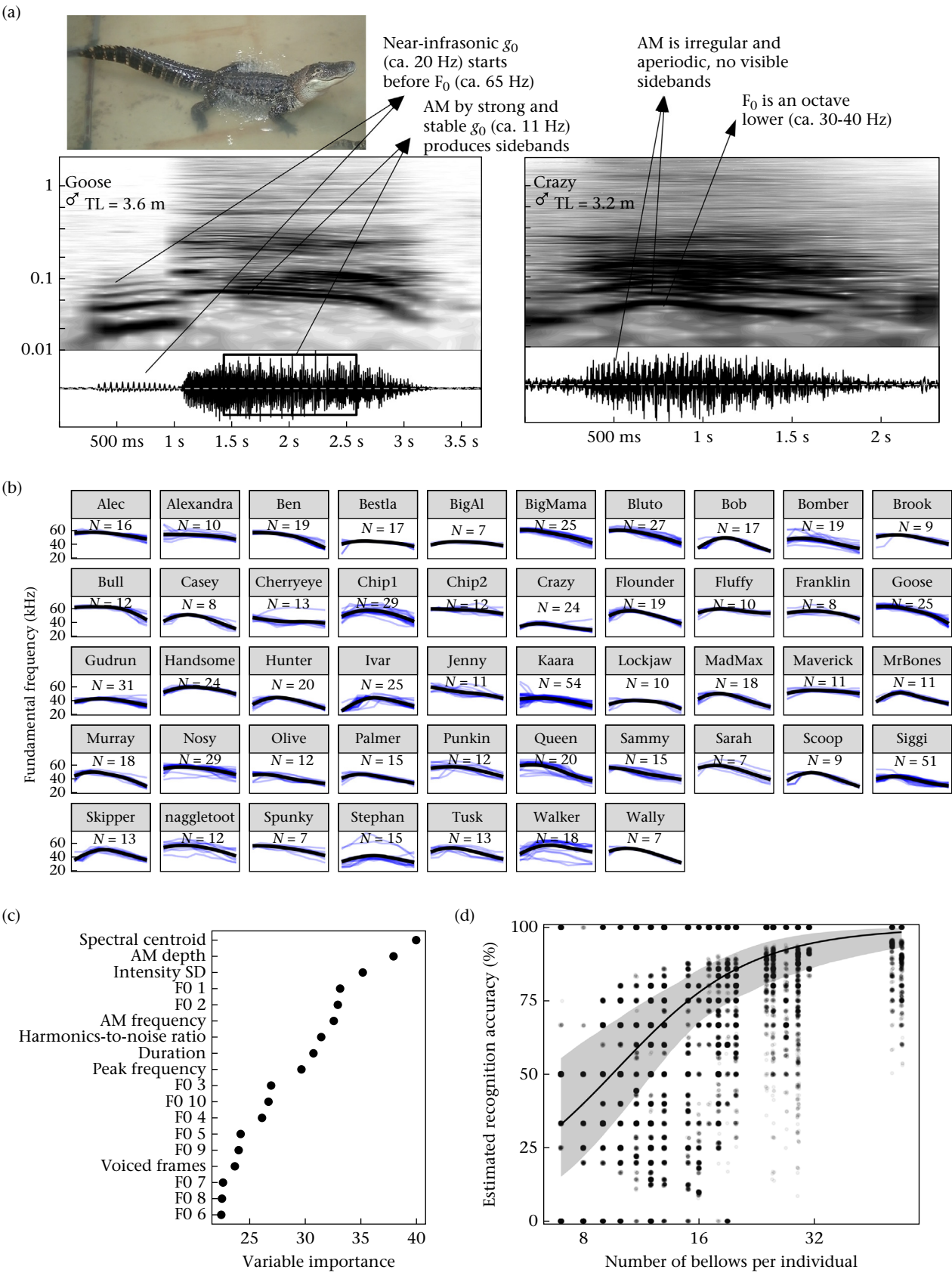
Figure 3. The encoding of individual identity in alligator bellows. (a) Spectrograms of bellows from two different alligators highlighting some acoustic signatures of the caller's identity (400 ms Gaussian window, frequency on a logarithmic scale, TL = total length; see Supplementary material for audio examples). (b) Typical intonation contour per animal (solid black line) and individual bellows (blue lines). Both mean F0 and its change over time (intonation contour) vary across individuals. The numbers show how many bellows were recorded from each individual. (c) The contribution of individual acoustic features towards individual recognition based on internal random forest estimation. (d) Supervised classification with random forest shows that 47 alligators can be distinguished nearly perfectly if at least 15–20 calls are recorded per individual: median of the posterior distribution of predicted accuracy per animal (solid line) with its 95% credible interval (shaded area) and individual simulations (points).

Table 2
Individual recognition accuracy: out-of-sample balanced accuracy averaged across 47 individuals (chance level = 1/47 = ca. 2.1%) by random forest models with different sets of acoustic predictors

Acoustic predictor set	Accuracy in test set (%) [95% CI][a]	Interpretation
Relative F0 contours	25 [21, 29]	Different individuals bellow with different intonation
Absolute F0 contours	36 [31, 42]	Intonations combined with the absolute F0 level are more individually distinct
Eight[b] time-averaged spectral features	32 [26, 37]	Bellows are also individually distinguishable by the average values of spectral features (most importantly duration, amplitude modulation, and spectral centroid)
Absolute F0 contours + 8 time-averaged spectral features	52 [45, 59]	Recognition accuracy improves when intonation contours and time-averaged spectral features are combined
Time-standardized MFCC matrices (31 coefficients × 10 time steps)	36 [30, 41]	Bellows are also recognizable when only compressed representation of the entire spectrogram are used (an alternative approach, which can be very effective provided that the data set is large enough)
31 time-averaged MFCCs (31 coefficients × 1 time step)	58 [51, 64]	A more compact representation (essentially, the average spectrum) may improve recognition accuracy despite sacrificing dynamic information (at least in the current sample with 6—54 calls per individual)
Absolute F0 contours + 8 time-averaged spectral features + 31 time-averaged MFCCs	66 [60, 73]	The combination of intonation contours, spectral descriptors, and MFCCs produced the best results. This approach depends on some manual intervention in the acoustic analysis
Seven[c] time-averaged robust spectral features + 31 time-averaged MFCCs	61 [54, 67]	Alligator bellows are individually distinct to a degree that robust spectral features and MFCCs, which can be extracted fully automatically without expert knowledge, yield results nearly as accurate as the best classifier

[a] The 95% coverage intervals were calculated by rerunning each random forest model 1000 times, each time with different samples in the training and testing data set. To avoid pseudoreplication, the models were trained and tested on bellows recorded on different days.
[b] These eight features are: duration, proportion of voiced frames, AM depth, AM frequency, intensity SD, harmonics-to-noise ratio, peak frequency and spectral centroid (see Table 1).
[c] These seven features are similar to those above but calculated for all frames rather than only for the voiced frames (and excluding the proportion of voiced frames).

Sweden) did not artificially inflate the classification accuracy of our algorithm, as performance improved somewhat when alligators from Sweden were excluded from the analysis (U.S.A. and Sweden ($N = 47$): accuracy 66% [60, 73]/U.S.A. alone ($N = 42$): 70% [63, 78]). We therefore believe that our algorithm can be applied to American alligator recordings in general.

Number of Recordings Required per Subject and Effect of Location

The number of recordings was strongly related to recognition accuracy (Fig. 3d). Hence, for practical applications, it would be advisable to obtain more recordings from each animal to ensure reliable recognition, perhaps in the range of 15—20. As expected, we also found that some acoustic features of bellows behaved differently on land than in water. A random forest classifier using F0, eight spectral features (Table 1) and time-averaged MFCCs classified bellows by location with an accuracy of ca. 95% (chance = 50%). Calls produced in the water were considerably less tonal, shorter in duration and had a higher peak frequency within each individual (Fig. 4). This raises the question of whether location might interfere with individual recognition, that is, whether we would be able to recognize an individual on land after learning their call only in the water, or vice versa. When we ran our algorithms on calls in water only and tested on land calls, or vice versa, we noticed poor transfer in both cases. Models trained and tested on a mixture of land and water bellows performed similarly in the training set and testing set (out-of-bag balanced accuracy of individual recognition 92.6% [86.1, 96.2] and 87.8% [78.2, 95.3], respectively, chance level = 1/13 ~= 7.7%). In contrast, the balanced accuracy in the testing set dropped markedly when trained on land bellows and tested on water bellows (to 70.6% [59.1, 80.6]) or vice versa (70.8% [57.7, 80.5]). The reduction in the size of the training data set does not explain this lack of transfer as out-of-bag accuracy in the training sample remained high (ca. 83%).

In sum, bellows produced on land appear to differ from bellows produced by the same animal in water, and this difference is significant enough to impact individual recognition. As such, ideally a sufficient number of vocalizations per animal should be recorded in both locations. Yet, individuality remains by far the largest source of acoustic variation in the analysed bellows.

The Main Acoustic Predictors of Individuality

As shown in Table 2, the absolute F0 contours of bellows (with frequency in Hz) were sufficient for recognizing individual alligators with a balanced accuracy of 36%, 95% CI [31, 42], which is worse than the performance of the full model above (66%), but still much better than chance (2.1%). The F0 contours normalized relative to the mean frequency of each bellow (i.e. converted from absolute frequencies to intonation) still enabled above-chance recognition (25% [21, 29]), indicating that individuality is encoded both in the average F0 and in the typical intonation of the bellowing. In particular, the first three of 10 F0 values (as all intonation contours were standardized to 10 points) had the highest variable importance (Fig. 3c). The intonation in the first third of a bellow is thus particularly distinct across individuals, whereas the terminal part of a bellow does not provide much useful information because F0 typically declined towards the end of the bellow in all recorded animals (Fig. 3b). Adding these eight acoustic descriptors listed in Table 1 to F0 contours in the random forest model in the Supervised classification improved its performance from 36% to 52% [45, 59]. The most important variables were, once again, the intonation in the first third of the bellow, but also spectral centroid and amplitude modulation, which is very pronounced in alligator bellows and apparently highly individual specific (Fig. 3a).

Acoustic Markers of Sex and Size

Alligators in the sample varied from 1.9 to 3.9 m in total body length (mean ± SD = 3.0 ± 0.6 m), and males were typically larger than females (mean 3.3 m versus 2.3 m). Controlling for body size, male alligators were found to produce longer bellows (+0.88 SD
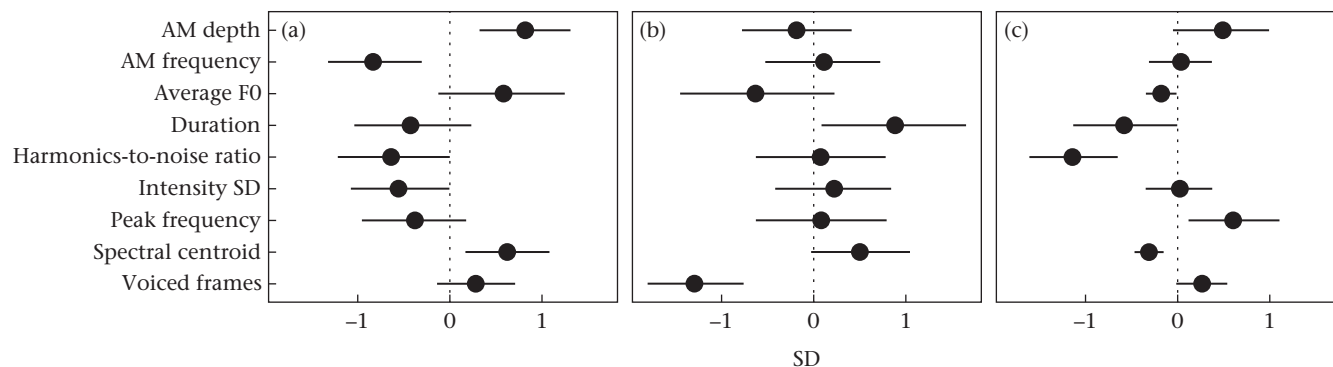
Figure 4. Acoustic differences related to body size, sex and location during bellowing. Predicted change in acoustic features, in standard deviations, for (a) a 1 m increase in body size controlling for sex, (b) a male versus a female alligator of the same size and (c) an alligator in the water versus on land: medians of posterior distribution and 95% coverage interval. $N = 47$ alligators for body length and sex and $N = 13$ for water versus land.

[0.09, 1.65]) with a smaller proportion of voiced frames (-1.29 SD [-1.80, -0.76]), likely due to the unvoiced water dance in the beginning of male bellows. The greatest differences between the bellows of small and large alligators of the same sex were captured by two characteristics: spectral centroid and amplitude modulation. Larger alligators produced bellows with considerably more energy in high frequencies (spectral centroid $+0.62$ [0.17, 1.08]). This may be because the calls of large animals are louder; we cannot measure this directly because the original loudness was not measured, but strong upper harmonics are often indicative of a more powerful sound source (Traunmuller & Eriksson, 2000). As for amplitude modulation, we found that larger animals produced stronger ($+0.82$ [0.32, 1,31]) amplitude modulation at a lower frequency (-0.83 [-1.32, -0.30]), compared to smaller animals (Fig. 4). Interestingly, F0 was not clearly indicative of either body size ($+0.58$ [-0.12, 1.25], that is, possibly higher in larger animals) or sex (-0.63[-1.45, 0.23] in males).

DISCUSSION

We provide the first evidence of acoustic cues to individual identity in adult crocodylians. The bellows of American alligators are highly distinct, making individual discrimination very accurate with or without pretraining the model on manually labelled vocalizations. It is thus possible to use acoustic identity cues for passive acoustic monitoring for research and conservation efforts. The next steps will be to test the alligators' ability to use cues in bellows to discriminate between familiar and unfamiliar conspecifics, and eventually to test whether they can truly recognize other individuals. Such skills appear to be beneficial for navigating their social environment, for example to avoid previous rivals (Carlson et al., 2020; Temeles, 1994; Tibbetts, 2004).

Bellows may be particularly suitable for advertising individual identity because they are long in duration and contain several steps that are strenuous to perform. The contours of the calls may therefore harbour many identity cues that may be caused by mere changes in the anatomy of individuals as they grow (Boughmann & Moss, 2003). Especially the initial third of F0 was highly individually specific, which is also when most force is put into the call. After this, the bellow slowly drops in intensity as the lungs get depleted of air. This is similar to other animals, such as giant pandas, for which individuality appears to be coded in F0 and the amplitude modulation (Charlton et al., 2009). Likewise, frequency parameters are important in the acoustic individuality of African bush elephants, *Loxodonta africana* (Wierucka et al., 2021), and amplitude modulation has high significance in lamb, *Ovis aries*, bleats (Sèbe

et al., 2018). Like in some mammals studied, F0 in alligator bellows is not a good predictor of body size (Grawunder et al., 2018; Reby & McComb, 2003). Generally, the characteristics of alligator vocal anatomy and their bellows share more similarities with large mammal vocalizations than with bird calls, despite their closer phylogenetic relationship with the latter (Reber, 2020; Reber et al., 2017; Riede et al., 2015). However, their F0 is not indicative of sex, which is often the case in mammals (Taylor & Reby, 2010). Crocodylian vocal folds are very small and essentially as long as the diameter of the glottis (Reese, 1945), which explains the decoupling of size and F0. However, tissues above the glottis in the pharyngeal cavity can provide additional oscillators responsible for the observed amplitude modulation, and they presumably keep growing as the animal increases in size.

We found no indication that alligators have any elaborate control over their vocal production (compared to, for example, birds), which makes crocodylians a good extant model for understanding the evolution of sound production in Archosauria. A recent study on the ossified larynx of an ankylosaur, *Pinacosaurus grangeri*, showed that the larynx morphology of this species might have allowed it to produce a similar diversity of controlled vocalizations as some birds (Yoshida et al., 2023b). The crocodylian hyolaryngeal apparatus is mainly cartilaginous, but all archosaurs have the same, although modified, elements in their vocal production system (Riede et al., 2015; Yoshida et al., 2023a). Hence, nonavian dinosaurs might have been able to produce sounds intermediate between modern birds and crocodylians. Today, crocodylians appear to be the only extant archosaurs producing advertisement calls that become individually distinct by putting extensive strain on the laryngeal muscles (Riede et al., 2019). In alligators, in addition to by-product distinctiveness (Boughmann & Moss, 2003), different growth rates at key developmental stages, bodily damage, illnesses or similar factors could also contribute to the individual identity cues. For instance, one of our subjects ('Maverick') had recovered from severe pneumonia that damaged one of his lungs; he subsequently produced bellows of approximately half the duration of similarly sized males (e.g. 'Goose'; mean bellow duration ± SE: Maverick: 1.16 ± 0.05 s, $N = 11$; Goose: 2.43 ± 0.07 s, $N = 25$).

As part of their individual identity, bellows also contain information about a caller's size, strength and endurance, and probably its sex (Fig. 4). Larger animals of either sex had a more stable and deeper amplitude modulation, and their calls had more energy in high-frequency bands (Fig. 4). Previous research has shown that the vocal tract resonances of American alligators reliably show a negative correlation with body size (Reber et al., 2017), and that the dominant frequency in the bellows of the closest relative, the

Chinese alligator, *Alligator sinensis*, is primarily caused by tissue vibrations (Reber et al., 2015). The dominant frequency is not a cue to body size in American alligators (Reber et al., 2017) and in the current study, we found that neither is F0. This shows that the mass and consistency of tissues above the glottis are barely contributing to the information about a caller's attributes. Body size plays a major role in the mate choice of crocodylians as females only mate with males larger than themselves, and males prefer large females because they can lay more eggs. When bellowing, only males produce a water dance containing subaudible vibrations, which precedes the audible bellow, and are most likely perceived with their integumentary sensory organs (Reber, 2020). The whole bellowing display might therefore act as a two-part signal: first, the water dance indicates a male producer, and second, the bellow itself contains a range of cues to size. That bellows signal size, strength and endurance is consistent with previous findings that bellowing alligators actively pull their larynx towards the sternum to elongate the vocal tract and thereby seal the palatal valve against the palatal plate (Reber et al., 2017). Stronger animals can be expected to be able to hold this position more consistently while air is pushed through the glottis and into the pharyngeal space, making the tissues inside vibrate more stably. A stronger subglottal pressure would also lead to a bigger proportion of this pharyngeal tissue, and probably tissue on the basiohyoid, to vibrate, which would result in deeper frequencies. This suggestion, that holding this bellowing position requires strain on muscles pulling the larynx down, is supported by our finding that bellows produced by the same individual differ when produced on land versus in water: The bellows on land were longer, more tonal and had a lower peak frequency, likely due to muscle capacity being freed up when not used for stabilization and balancing in water. Alligators consistently produce lower first formants on land than they do in the water, and on land their vocal tract configuration more closely resembles a stretched tube, at least mathematically (Reber et al., 2017). Water pressure against the pharynx and the strain of holding the tail up out of the water (a behaviour not shown on land) may prevent the animals from pulling the larynx maximally down. The same strain might shorten the call duration and impact the ability to sustain regular vibrations.

In addition to testing whether alligators can recognize other individuals based on acoustics, it would be interesting to see how strong their recognition of conspecifics is. For instance, we now found that bellows sounded markedly different depending on whether they were produced on land or in water. Would subjects generalize across production contexts in a playback experiment? If so, their recognition may be more fine-tuned than our classification algorithms. There is no direct evidence in our material that any element of the bellowing display evolved to specifically advertise an animal's identity: rather, identity cues likely constitute a by-product of individually unique vocal morphology and anatomical changes that accumulate progressively throughout life. Given the universality of the crocodylian call repertoire within Crocodylia (Vergne et al., 2009), we can expect that the roars and bellows of other species in this order also have several parameters rich in identity cues. This particularly applies to the 26 species of alligatorids and crocodiles, which have a similar mating system to American alligators. For the same reason, we would also not expect any selective pressures for evolving acoustic identity cues in these species. A potentially quite different example is the Indian gharial, *Gavialis gangeticus*, which has a mating system akin to resource defence polygyny: the female chooses a mate (or mates) each season at whose territory she may nest. The hatching of eggs in this species is highly synchronized between all nests, and the young are protected by both males and females. Parental protection, including protection by males, is essential for the survival of gharial

hatchlings, a breeding system reminiscent of some palaeognath birds, for example the ostrich, *Struthio camelus* (Davis, 2002; Jailabdeen & Lang, 2018; Lang, 2015; Melgar et al., 2022). Here, the size and strength of the male may not be the only qualities to which a female attends. Over several seasons, a male might build prestige as a good caretaker; and in such a system, there could be selective pressure to evolve individuality cues in advertisement calls.

Our findings have practical applications, particularly in view of the great robustness of individual recognition by the tested models. For example, it is useful to have an automated analysis pipeline for automatically tracking individual alligators in a specific area, such as a national park. One way to achieve this functionality is to train a supervised classifier to recognize known individuals and then continue to detect their presence in future recordings, as with our supervised random forest models. As we demonstrated, however, the bellows of many individual alligators formed compact and highly distinct clusters in our unsupervised analysis. Although the recordings in this case were segmented manually and we verified our F0 measurements, classification based on standard, automatically extracted acoustic measurements (requiring no expert knowledge) was shown to be nearly as effective. It is therefore technically possible to estimate the number of individual crocodylians in a population without human intervention in the analysis, and indeed without knowing the identities of individuals (e.g. based on passive bioacoustics monitoring). As the conservation methods of American alligators involve alligator farming (Frechette, 2001) and sustainable harvest of wild populations (Grigg & Kirshner, 2015), being able to accurately estimate population densities is vital to preventing overharvesting (Joanen et al., 2021). Currently, the method for monitoring alligator populations is the eye-shine survey, involving the counting of alligators at night via the reflections of light from their eyes, which does not sufficiently discriminate between age classes (Subalusky et al., 2009). In contrast, bellowing is only performed by healthy, adult individuals, or animals close to sexual maturity, which means that a population survey based on bellowing behaviour provides a better assessment of the 'effective population size' (i.e. breeding individuals). A potential caveat is the distinction between land and water bellows, potentially leading to land and water bellows by one caller being classified as two individuals if the sample size is low. As such, it is important to collect a good number of calls before estimating population densities.

In conclusion, American alligator bellows display several clear cues making them individually distinct. These relate to frequency measures, but the most prominent identity cues lie in the temporal development of these measures throughout a bellow. We developed an algorithm that can identify these cues in an unsupervised manner and reliably estimate the numbers of callers in a sample. This has applications for research and conservation biology and opens doors to future studies on acoustic individual recognition by crocodylians, which would be the first evidence of its kind in any nonavian reptile.

Author Contributions

S.A.R. and A.A. conceived the study. S.A.R. and T.R.J. collected the data. A.A. and T.R.J. performed the data analysis and wrote the first draft of the manuscript. M.O. helped conceptualize the findings. All authors contributed substantially to the revisions of the paper.

Data Availability

Data and code used in this study are available at Open Science Framework via https://doi.org/10.17605/OSF.IO/KFTD510. Original

audio files are not directly publicly available for copyright reasons but can be made available upon request.

Declaration of Interest

Acknowledgments

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.anbehav.2023.11.00910.

References

Anikin, A. (2019). Soundgen: An open-source tool for synthesizing nonverbal vocalizations. *Behavior Research Methods, 51*(2), 778–792. https://doi.org/10.3758/s13428-018-1095-7

Arnaud, V., Pellegrino, F., Keenan, S., St-Gelais, X., Mathevon, N., Levrero, F., & Coupe, C. (2023). Improving the workflow to crack Small, Unbalanced, Noisy, but Genuine (SUNG) datasets in bioacoustics: The case of bonobo calls. *PLoS Computational Biology, 19*(4). https://doi.org/10.1371/journal.pcbi.1010325

ASAB Ethical Committee/ABS Animal Care Committee. (2023). Guidelines for the ethical treatment of nonhuman animals in behavioural research and teaching. *Animal Behaviour, 195*. I–Xi https://doi.org/10.1016/j.anbehav.2022.09.006.

Basken, J. N., Connor, N. P., & Ciucci, M. R. (2012). Effect of aging on ultrasonic vocalizations and laryngeal sensorimotor neurons in rats. *Experimental Brain Research, 219*(3), 351–361. https://doi.org/10.1007/s00221-012-3096-6

Bee, M. A., Kozich, C. E., Blackwell, K. J., & Gerhardt, H. C. (2001). Individual variation in advertisement calls of territorial male green frogs, *Rana clamitans*: Implications for individual discrimination. *Ethology, 107*(1), 65–84. https://doi.org/10.1046/j.1439-0310.2001.00640.x

Benti, B., Cure, C., & Dufour, V. (2019). Individual signature in the most common and context-independent call of the rook (*Corvus frugilegus*). *Wilson Journal of Ornithology, 131*(2), 373–381. https://doi.org/10.1676/18-41

Boughmann, J., & Moss, C. F. (2003). Social sounds: Vocal learning and development of mammal and bird calls. In A. M. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (Vol. 16, pp. 138–224). Springer.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, J. C., Smaragdis, P., & Nousek-McGregor, A. (2010). Automatic identification of individual killer whales. *Journal of the Acoustical Society of America, 128*(3), El93–El98. https://doi.org/10.1121/1.3462232

Carlson, N. V., Kelly, E. M., & Couzin, I. (2020). Individual vocal recognition across taxa: A review of the literature and a look into the future. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1802). https://doi.org/10.1098/rstb.2019.0479

Charlton, B. D., Ellis, W. A. H., Brumm, J., Nilsson, K., & Fitch, W. T. (2012). Female koalas prefer bellows in which lower formants indicate larger males. *Animal Behaviour, 84*(6), 1565–1571. https://doi.org/10.1016/j.anbehav.2012.09.034

Charlton, B. D., Zhang, Z. H., & Snyder, R. J. (2009). Vocal cues to identity and relatedness in giant pandas (*Ailuropoda melanoleuca*). *Journal of the Acoustical Society of America, 126*(5), 2721–2732. https://doi.org/10.1121/1.3224720

Christie, P. J., Mennill, D. J., & Ratcliffe, L. M. (2004). Chickadee song structure is individually distinctive over long broadcast distances. *Behaviour, 141*, 101–124. https://doi.org/10.1163/156853904772746628

Clemins, P. J., Johnson, M. T., Leong, K. M., & Savage, A. (2005). Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. *Journal of the Acoustical Society of America, 117*(2), 956–963. https://doi.org/10.1121/1.1847850

Davis, S. J. J. F. (2002). *Ratites and tinamous*. Oxford University Press.

Fant, G. (1960). *Acoustic theory of speech production*. Mouton.

Ferreira, A. C., Silva, L. R., Renna, F., Brandl, H. B., Renoult, J. P., Farine, D. R., … Doutrelant, C. (2020). Deep learning-based methods for individual recognition in small birds. *Methods in Ecology and Evolution, 11*(9), 1072–1085. https://doi.org/10.1111/2041-210x.13436

Frechette, D. L. (2001). Gatornomics: Profitable and sustainable use of alligators in the southeastern United States. *Marine Resource Economics, 16*(2), 127–142. https://doi.org/10.1086/mre.16.2.42628834

Garrick, L. D., & Lang, J. W. (1977). Social signals and behaviors of adult alligators and crocodiles. *American Zoologist, 17*(1), 225–239. https://doi.org/10.1093/icb/17.1.225

Garrick, L. D., Lang, J. W., & Herzog, H. A., Jr. (1978). Social signals of adult American alligators. *Bulletin of the American Museum of Natural History, 160*, 3.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software, 31*(7), 1–24. https://doi.org/10.18637/jss.v031.i07

Gradisek, A., Slapnicar, G., Sorn, J., Lustrek, M., Gams, M., & Grad, J. (2017). Predicting species identity of bumblebees through analysis of flight buzzing sounds. *Bioacoustics: The International Journal of Animal Sound and Its Recording, 26*(1), 63–76. https://doi.org/10.1080/09524622.2016.1190946

Grawunder, S., Crockford, C., Clay, Z., Kalan, A. K., Stevens, J. M. G., Stoessel, A., & Hohmann, G. (2018). Higher fundamental frequency in bonobos is explained by larynx morphology. *Current Biology, 28*(20), R1188–R1189. https://doi.org/10.1016/j.cub.2018.09.030

Grendeus, S., & Reber, S. A. (2020). Crocodylia cognition. In J. Vonk, & T. K. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior*. Springer. https://doi.org/10.1007/978-3-319-47829-6_926-1.

Grigg, G., & Kirshner, D. (2015). *Biology and evolution of crocodylians*. Cornell University Press.

Hahsler, M., Piekenbrock, M., & Doran, D. (2019). dbscan: fast density-based clustering with R. *Journal of Statistical Software, 91*(1), 1–30. https://doi.org/10.18637/jss.v091.i01

Hambalkova, L., Policht, R., Horak, J., & Hart, V. (2021). Acoustic individuality in the hissing calls of the male black grouse (*Lyrurus tetrix*). *PeerJ, 9*, Article e11837. https://doi.org/10.7717/peerj.11837

Jailabdeen, A., & Lang, J. W. (2018). Acoustic communication in Gharials (*Gavialis gangeticus*) [Conference presentation abstract October 8-13]. In *XXVI International Bioacoustics Council Congress*.

Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences of the United States of America, 103*(21), 8293–8297. https://doi.org/10.1073/pnas.0509918103

Joanen, T., Merchant, M., Griffith, R., Linscombe, J., & Guidry, A. (2021). Evaluation of effects of harvest on alligator populations in Louisiana. *Journal of Wildlife Management, 85*(4), 696–705. https://doi.org/10.1002/jwmg.22028

Lang, J. W. (2015). Behavioral ecology of the gharial, *Gavialis gangeticus*. *Integrative and Comparative Biology, 55*(Issue Suppl_1). https://doi.org/10.1093/icb/icv011. E105–E105.

Lehmann, K. D. S., Jensen, F. H., Gersick, A. S., Strandburg-Peshkin, A., & Holekamp, K. E. (2022). Long-distance vocalizations of spotted hyenas contain individual, but not group, signatures. *Proceedings of the Royal Society B: Biological Sciences, 289*(1979). https://doi.org/10.1098/rspb.2022.0548

Lengagne, T., Lauga, J., & Jouventin, P. (1997). A method of independent time and frequency decomposition of bioacoustic signals: Inter-individual recognition in four species of penguins. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences, 320*(11), 885–891. https://doi.org/10.1016/S0764-4469(97)80873-6

Ligges, U., Krey, S., & Schnackenberg, S. (2018). *tuneR: analysis of music and speech*. https://CRAN.R-project.org/package=tuneR.

Linhart, P., Mahamoud-Issa, M., Stowell, D., & Blumstein, D. T. (2022). The potential for acoustic individual identification in mammals. *Mammalian Biology, 102*, 667–683. https://doi.org/10.1007/s42991-021-00222-2

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., & Tyack, P. L. (2013). Estimating animal population density using passive acoustics. *Biological Reviews, 88*(2), 287–309. https://doi.org/10.1111/brv.12001

Martin, K., Adam, O., Obin, N., & Dufour, V. (2022). Rookognise: Acoustic detection and identification of individual rooks in field recordings using multi-task neural networks. *Ecological Informatics, 72*, Article 101818. https://doi.org/10.1016/j.ecoinf.2022.101818

Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford University Press.

McComb, K., & Reby, D. (2009). Communication in terrestrial animals. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (pp. 1167–1171). Academic Press.

Melgar, J., Schou, M. F., Bonato, M., Brand, Z., Engelbrecht, A., Cloete, S. W., & Cornwallis, C. K. (2022). Experimental evidence that group size generates divergent benefits of cooperative breeding for male and female ostriches. *eLife, 11*. https://doi.org/10.7554/eLife.77170. ARTN e77170.

Melville, J., Lun, A., & Djekidel, M. (2020). *uwot: The uniform manifold approximation and projection (UMAP) method for dimensionality reduction*. https://CRAN.R-project.org/package=uwot.

Mitani, J. C., GrosLouis, J., & Macedonia, J. M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology, 17*(4), 569–583. https://doi.org/10.1007/Bf02735192

Parsons, S., & Jones, G. (2000). Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *Journal of Experimental Biology, 203*(17), 2641–2656. https://doi.org/10.1242/jeb.203.17.2641

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Reber, S. A. (2018). Crocodilia communication. In J. Vonk, & T. K. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior*. Springer. https://doi.org/10.1007/978-3-319-47829-6_950-1.

Reber, S. A. (2020). Crocodilians are promising intermediate model organisms for comparative perception research. *Comparative Cognition & Behavior Reviews, 15*, 111—129. https://doi.org/10.3819/CCBR.2020.150004

Reber, S. A., Janisch, J., Torregrosa, K., Darlington, J., Vliet, K. A., & Fitch, W. T. (2017). Formants provide honest acoustic cues to body size in American alligators. *Scientific Reports, 7*(1816), 1—11. https://doi.org/10.1038/s41598-017-01948-1

Reber, S. A., Nishimura, T., Janisch, J., Robertson, M., & Fitch, W. T. (2015). A Chinese alligator in heliox: Formant frequencies in a crocodilian. *Journal of Experimental Biology, 218*(15), 2442—2447. https://doi.org/10.1242/jeb.119552

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour, 65*, 519—530. https://doi.org/10.1006/anbe.2003.2078

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences, 272*(1566), 941—947. https://doi.org/10.1098/rspb.2004.2954

Reese, A. M. (1945). The laryngeal region of *Alligator mississippiensis*. *Anatomical Record, 92*(3), 273—277. https://doi.org/10.1002/ar.1090920308

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F-0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *Journal of the Acoustical Society of America, 117*(2), 944—955. https://doi.org/10.1121/1.1848011

Riede, T., Li, Z. H., Tokuda, I. T., & Farmer, C. G. (2015). Functional morphology of the *Alligator mississippiensis* larynx with implications for vocal production. *Journal of Experimental Biology, 218*(7), 991—998. https://doi.org/10.1242/jeb.117101

Riede, T., Thomson, S. L., Titze, I. R., & Goller, F. (2019). The evolution of the syrinx: An acoustic theory. *PLoS Biology, 17*(2), Article 2006507. https://doi.org/10.1371/journal.pbio.2006507

Rohtla, E. A., Russell, A. P., & Bauer, A. M. (2019). Sounding off: Relationships between call properties, body size, phylogeny, and laryngotracheal form of geckos. *Herpetologica, 75*(3), 175—197. https://doi.org/10.1655/D-19-00021

Sainburg, T., Thielk, M., & Gentner, T. Q. (2020). Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLoS Computational Biology, 16*(10), Article 1008228. https://doi.org/10.1371/journal.pcbi.1008228

Sataloff, R. T. (2017). The effects of age on the voice. In R. T. Sataloff (Ed.), *Vocal health and pedagogy: Science, assessment, and treatment* (pp. 319—338). Plural Publishing.

Sèbe, F., Poindron, P., Ligout, S., Sèbe, O., & Aubin, T. (2018). Amplitude modulation is a major marker of individual signature in lamb bleats. *Bioacoustics: The International Journal of Animal Sound and Its Recording, 27*(4), 359—375. https://doi.org/10.1080/09524622.2017.1357146

Subalusky, A. L., Smith, L. L., & Fitzgerald, L. A. (2009). Detection of American alligators in isolated, seasonal wetlands. *Applied Herpetology, 6*(3), 199—210. https://doi.org/10.1163/157075408x386132

Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology, 280*(3), 221—236. https://doi.org/10.1111/j.1469-7998.2009.00661.x

Temeles, E. J. (1994). The role of neighbors in territorial systems: When are they 'dear enemies'? *Animal Behaviour, 47*(2), 339—350. https://doi.org/10.1006/anbe.1994.1047

Terry, A. M. R., Peake, T. M., & McGregor, P. K. (2005). The role of vocal individuality in conservation. *Frontiers in Zoology, 2*, 10. https://doi.org/10.1186/1742-9994-2-10

Thomsen, H. M., Balsby, T. J. S., & Dabelsteen, T. (2013). Individual variation in the contact calls of the monomorphic peach-fronted conure, *Aratinga aurea*, and its potential role in communication. *Bioacoustics: The International Journal of Animal Sound and Its Recording, 22*(3), 215—227. https://doi.org/10.1080/09524622.2013.779560

Thomsen, H. M., Balsby, T. J. S., & Dabelsteen, T. (2019). The imitation dilemma: Can parrots maintain their vocal individuality when imitating conspecifics? *Behaviour, 156*(5—8), 787—814. https://doi.org/10.1163/1568539x-00003548

Tibbetts, E. A. (2004). Complex social behaviour can select for variability in visual features: A case study in *Polistes* wasps. *Proceedings of the Royal Society B: Biological Sciences, 271*(1551), 1955—1960. https://doi.org/10.1098/rspb.2004.2784

Tibbetts, E. A., & Dale, J. (2007). Individual recognition: It is good to be different. *Trends in Ecology & Evolution, 22*(10), 529—537. https://doi.org/10.1016/j.tree.2007.09.001

Titze, I. R. (1994a). *Principles of voice production*. Prentice Hall.

Titze, I. R. (1994b). Toward standards in acoustic analysis of voice. *Journal of Voice, 8*(1), 1—7. https://doi.org/10.1016/S0892-1997(05)80313-3

Trapanotto, M., Nanni, L., Brahnam, S., & Guo, X. (2022). Convolutional neural networks for the identification of African lions from individual vocalizations. *Journal of Imaging, 8*(4), 96. https://doi.org/10.3390/jimaging8040096

Traunmuller, H., & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *Journal of the Acoustical Society of America, 107*(6), 3438—3451. https://doi.org/10.1121/1.429414

Vergne, A. L., Avril, A., Martin, S., & Mathevon, N. (2007). Parent-offspring communication in the Nile crocodile *Crocodylus niloticus*: Do newborns' calls show an individual signature? *Naturwissenschaften, 94*(1), 49—54. https://doi.org/10.1007/s00114-006-0156-4

Vergne, A. L., Pritz, M. B., & Mathevon, N. (2009). Acoustic communication in crocodilians: From behaviour to brain. *Biological Reviews, 84*(3), 391—411. https://doi.org/10.1111/j.1469-185X.2009.00079.x

Vliet, K. A. (1989). Social displays of the American alligator (*Alligator mississippiensis*). *American Zoologist, 29*(3), 1019—1031. https://doi.org/10.1093/icb/29.3.1019

Wierucka, K., Henley, M. D., & Mumby, H. S. (2021). Acoustic cues to individuality in wild male adult African savannah elephants (*Loxodonta africana*). *PeerJ, 9*. https://doi.org/10.7717/peerj.10736

Wyman, M. T., Walkenhorst, B., & Manser, M. B. (2022). Selection levels on vocal individuality: Strategic use or byproduct. *Current Opinion in Behavioral Sciences, 46*, Article 101140. https://doi.org/10.1016/j.cobeha.2022.101140

Yorzinski, J. L. (2017). The cognitive basis of individual recognition. *Current Opinion in Behavioral Sciences, 16*, 53—57. https://doi.org/10.1016/j.cobeha.2017.03.009

Yoshida, J., Kobayashi, Y., & Fiorillo, A. R. (2023). Evolutionary insights from an anatomical network analysis of the hyolaryngeal apparatus in extant archosaurs (birds and crocodilians). *Anatomical Record-Advances in Integrative Anatomy and Evolutionary Biology, 306*(7), 1631—1645. https://doi.org/10.1002/ar.25153

Yoshida, J., Kobayashi, Y., & Norell, M. A. (2023). An ankylosaur larynx provides insights for bird-like vocalization in non-avian dinosaurs. *Communications Biology, 6*(1), 152. https://doi.org/10.1038/s42003-023-04513-x