Research

Author for correspondence:
Michael A. Pardo
e-mail: map385@cornell.edu

THE ROYAL SOCIETY PUBLISHING

Female African elephant rumbles differ between populations and sympatric social groups

Michael A. Pardo[1], David S. Lolchuragi[2], Joyce Poole[3], Petter Granli[3], Cynthia Moss[4], Iain Douglas-Hamilton[2] and George Wittemyer[1,2]

[1]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO, USA
[2]Save The Elephants, Nairobi, Kenya
[3]ElephantVoices, Sandefjord, Norway
[4]Amboseli Elephant Research Project, Nairobi, Kenya

MAP, 0000-0001-6978-848X; GW, 0000-0003-1640-5355

Vocalizations often vary in structure within a species, from the individual to population level. Vocal differences among social groups and populations can provide insight into biological processes such as vocal learning and evolutionary divergence, with important conservation implications. As vocal learners of conservation concern, intraspecific vocal variation is of particular interest in elephants. We recorded calls from individuals in multiple, wild elephant social groups in two distinct Kenyan populations. We used machine learning to investigate vocal differentiation among individual callers, core groups, bond groups (collections of core groups) and populations. We found clear evidence for vocal distinctiveness at the individual and population level, and evidence for much subtler vocal differences among social groups. Social group membership was a better predictor of call similarity than genetic relatedness, suggesting that subtle vocal differences among social groups may be learned. Vocal divergence among populations and social groups has conservation implications for the effects of social disruption and translocation of elephants.

1. Introduction

Many vocal animals exhibit intraspecific variation in the properties of their vocalizations. Geographic variation in

vocalizations occurs on a wide range of spatial scales and may be graded or discrete. The songs of red-faced cisticolas (*Cisticola erythrops*) vary gradually across sub-Saharan Africa over more than 6500 km [1], while mourning warblers (*Geothlypis philadelphia*) have discrete, regionally specific dialects spanning up to 2000 km [2]. At the other end of the geographical scale, little hermit hummingbirds (*Phaethornis longuemareus*) exhibit multiple microgeographic dialects within a single lek that span just a few tens of metres [3]. Geographic vocal variation can be adaptive; for example, by facilitating mating with locally adapted individuals [4] or optimizing vocalizations for the sound propagation properties and noise profile of the local habitat [5]. However, it can also be selectively neutral [6,7].

Animals may also exhibit vocal differences among social groups that overlap in home range but are socially distinct from one another. Some cetaceans have group-specific repertoires of discrete acoustic signals, in which some but not all call types are shared with other social groups [8–10]. More commonly, vocal group signatures result from a single species-wide call type being slightly more similar in structure among members of the same group [11–17].

A common function of vocal group signatures is to facilitate social recognition. While individual recognition allows more fine-scale discrimination among conspecifics [18], group recognition may be less cognitively demanding, as it requires learning fewer distinct signals. A shared signature of group identity can act as a 'password', allowing animals to easily assess the group membership of others without having to recognize them all individually [14,19,20].

Geographic vocal variation can occur regardless of whether calls are innate or learned. Differences in body size or other morphological characteristics between populations can lead to geographic vocal differences, as vocal parameters such as fundamental frequency, formants and maximum duration are tied to vocal cord mass, vocal tract length and lung capacity, respectively [21]. Social learning of vocalizations often results in distinct regional dialects even in the presence of substantial gene flow [22]. Animals with some degree of behavioural plasticity in vocal production may adjust their vocalizations to propagate more efficiently in the local habitat even if they do not socially learn their calls, although learning ability seems to facilitate such acoustic adaptation [5,23,24]. Most documented examples of vocal differences among sympatric social groups involve social learning [12,13,17,25], although it is theoretically possible for vocal group signatures to be genetic.

Vocal learning can be broadly divided into usage learning and production learning, both of which can lead to vocal divergence among groups [13,26,27]. Vocal usage learning involves either learning to modify the context in which existing calls are produced or learning to modify the temporal patterning of calls, while vocal production learning involves modifying the acoustic structure of vocalizations based on auditory experience [27]. Vocal production learning is rarer than usage learning and exists on a spectrum of complexity [27]. Some animals, such as meerkats (*Suricata suricatta*) [28], domestic goats (*Capra hircus*) [16] and non-human primates [29], can learn slight modifications to species-typical vocalizations that are otherwise innate. Other vocal learners, such as swamp sparrows (*Melospiza georgiana*), require auditory input to develop normal songs but only learn songs that are typical of their species [30], while still others, such as superb lyrebirds (*Menura novaehollandiae*), can mimic heterospecific sounds [31].

Intraspecific vocal variation is of interest in elephants for both practical and theoretical reasons. As managers frequently translocate elephants between populations, understanding how elephant populations and social groups differ in their vocal behaviour could be valuable for their conservation [32]. Moreover, studies on captive individuals have shown that elephants are among the few mammals capable of mimicking heterospecific sounds [33,34], but it is unknown how vocal production learning manifests in wild elephants. One possibility is that elephants evolved vocal learning to facilitate social recognition through the development of group signatures, which could be particularly beneficial for elephants given their large and multi-tiered social networks (figure 1).

In African savannah elephants (*Loxodonta africana*), mothers and their dependent offspring form the most fundamental unit of social organization, multiple (usually related) mother–offspring units form a 'core group' led by the oldest adult female, multiple (usually related) core groups form a 'bond group', and multiple bond groups form a 'clan' [35–38]. Elephants vocally discriminate among different tiers of social affiliates, and if vocal group signatures exist this might facilitate recognition, especially of distant affiliates [39].

Animals with multi-tiered social structures may exhibit group signatures at one or more levels of social organization. In killer whales (*Orcinus orca*) and sperm whales (*Physeter macrocephalus*), different tiers of social organization can be distinguished by the number of call types they have in common, with individuals from the same core unit sharing the most call types [8,40]. In greater spear-nosed bats (*Phyllostomus hastatus*), calls cluster by social group within a cave, and all groups from the
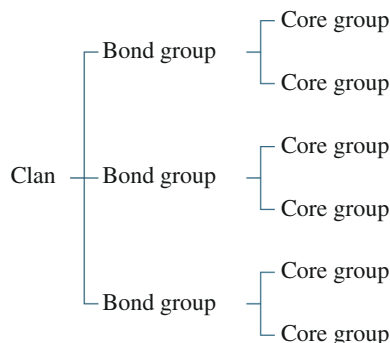
Figure 1. Illustration of hierarchically nested social organization of female African savannah elephants. Wittemyer *et al*. [35] reported a mean of 7.46 individuals per core group, 2.0 core groups per bond group and 3.25 bond groups per clan in Samburu.

same cave cluster more closely with one another than with calls from other caves [14]. By contrast, geladas (*Theropithecus gelada*) exhibit call convergence at the level of the band (the second tier of social organization) but not at the level of smaller social units within a band [17]. Most elephant vocalizations are low-frequency, harmonic calls known as rumbles [41]. Rumbles are individually specific [42–46], but no prior study has investigated whether rumbles also exhibit group-specific acoustic signatures at any level of social affiliation or intraspecific geographic variation.

We tested the hypotheses that the acoustic structure of rumbles produced by female African savannah elephants differs between populations, bond groups, core groups and individuals. We also hypothesized that vocal differences among social groups are learned rather than genetically determined. The predictions associated with these hypotheses are summarized in table 1.

2. Methods

2.1. Data collection

We recorded rumbles from wild adult female elephants in Amboseli National Park, Kenya ('Amboseli') in 1986–1990 and 1997–2006 and in Samburu and Buffalo Springs National Reserves, Kenya ('Samburu') in November 2019–March 2020 and June 2021–April 2022. These two populations are 390 km apart with no current gene flow between them due to intervening urban development [47]. Both populations have been continuously monitored for decades and all individuals can be individually identified by external ear morphology [35,36]. We focused on adult females (10+ years of age) to ensure that any vocal differences between populations or social groups were not an artefact of age or sex. Our final dataset included calls from 21 adult females in Amboseli (mean ± s.d. age = 26.2 ± 12.9 years) and 81 adult females in Samburu (mean ± s.d. age = 25.6 ± 11.2 years). The field recording methods for this dataset [48] have been previously published [49].

We recorded the identity of the caller and the behavioural context of each call. The caller was identified using behavioural and contextual cues, such as an open mouth, flapping ears or being the only individual who was not a young calf in the immediate vicinity (calf calls are easily distinguished from adult/subadult calls due to their higher frequency and shorter duration) [41]. We only included in the analysis calls for which we were able to identify the caller with certainty. Behavioural context was originally scored using slightly different ethograms in Amboseli [41] and Samburu [49]. To facilitate comparison between these two datasets, we concatenated behavioural context into nine categories shared across both populations (electronic supplementary material, table S1).

2.2. Definition of social groups

We determined the group membership of the elephants in Samburu following a previously published protocol [35]. In brief, we calculated simple ratio association indices [50] between all adult females in the population using observational data collected between January 2019 and April 2022, excluding individuals that were seen less than 20 times during this period. We performed a Ward's hierarchical cluster analysis on association indices, plotted the cumulative number of bifurcations as a function of bifurcation distance, and identified the most significant knot by visual inspection of the plot (*sensu*

Table 1. Hypotheses and predictions tested in this study. The majority classifier is a model that always guesses the most numerous category in the training data, and is a more conservative baseline than the weighted expectation (chance).

hypotheses	predictions
1. Rumbles differ between allopatric populations	1a. Calls can be assigned to population with better accuracy than majority classifier
	1b. Calls from the same population are more similar on average than calls from different populations
2. Rumbles exhibit acoustic signatures at the bond group level	2a. Calls can be assigned to bond group with better accuracy than majority classifier
	2b. Calls from different core groups in the same bond group are more similar on average than calls from different bond groups
3. Rumbles exhibit acoustic signatures at the core group level, which are stronger than the acoustic signatures at the bond group level	3a. Calls can be assigned to core group with better accuracy than majority classifier
	3b. Calls from same core group are more similar on average than calls from different core groups in the same bond group or from different bond groups
4. Rumbles differ among individual callers	4a. Calls can be assigned to individual callers with better accuracy than majority classifier
5. Vocal differences among social groups are learned, not genetic	5a. Social group (core or bond group) membership is a significant predictor of call similarity but genetic relatedness is not

[35]). We cut the dendrogram at the height corresponding to this knot and designated all clusters below this height as separate core groups. To identify bond groups, we repeated the procedure using only the matriarch (oldest female) of each core group. We did not include Amboseli data in the analysis of core groups and bond groups because we did not have comparable association data for Amboseli, and because 92% of our Amboseli recordings came from individuals belonging to the same (subjectively defined) core group. Data analysis was conducted using R v. 4.2.2 [51].

2.3. Call measurement

We measured 94 acoustic features on each call describing the distribution of energy across time and frequency in the mel spectrogram (electronic supplementary material, table S2, Supplemental Methods). A mel spectrogram is similar to a traditional spectrogram (raster plot with time on the x-axis, frequency on the y-axis, and amplitude indicated by pixel darkness) but with frequency transformed to the logarithmic mel scale [52]. While the mel scale was designed to approximate human hearing sensitivity, most other mammals, including elephants, perceive frequency on a similar logarithmic scale [53].

While the measurements we took from the mel spectrogram capture more of the variation in the calls than traditional measurements such as fundamental frequency and formants, they also capture background noise and thus may be more susceptible to influence from the recording equipment. To ensure that differences between Amboseli and Samburu could not be attributed to the different recording gear used in each population, we also traced the second harmonic (f_1) contour of each call, which is highly robust to recording equipment differences [54]. We calculated nine summary statistics of the f_1 contour (electronic supplementary material, table S2).

2.4. Individual, social unit and population assignment accuracy

To test the hypothesis that Amboseli and Samburu elephants exhibit population-level acoustic differences, we ran a random forest (500 trees, six variables/node, 60% of observations/tree, minimum node size = 1, no maximum tree depth) to predict population as a function of the mel spectrogram acoustic

features. A random forest is a machine learning model comprising many decision trees, each using a randomly selected fraction of the data [55]. This analysis included 938 calls from 81 individuals in Samburu, and 414 calls from 21 individuals in Amboseli. As random forests are biased towards more numerous classes [56], we balanced the dataset by randomly subsampling the Samburu observations so there were an equal number of observations in each class. To ensure that the model could only predict population using cues that generalized across callers, we randomly selected 20% of the callers with at least five calls each from each population and allocated all calls from these callers to the test set, with the remaining calls allocated to the training set [57]. We calculated the proportion of observations in the test set that were classified correctly (classification accuracy) and ran a one-tailed exact binomial test comparing the classification accuracy with the proportion that would have been classified accurately if the model always guessed the most common group in the training set ('majority' or 'zero rate' classifier). The majority classifier is a standard baseline model used in machine learning [58]. When there are the same number of observations in each class, the majority classifier is mathematically equivalent to the weighted expectation (guessing each class with a probability equal to the proportion of the data comprised by that class). However, when the data are imbalanced, the majority classifier always outperforms the weighted expectation. We repeated the above process 10 000 times and calculated the median p-value across all runs. The number of calls allocated to the test set varied across runs because different callers produced different numbers of calls. The mean ± s.d. proportion of the calls allocated to the test set across 10 000 runs was 0.16 ± 0.05. Note that although the full dataset was balanced by subsampling an equal number of calls from each population, the data were not perfectly balanced within the training and test sets, because the number of calls from each population allocated to the test set depended on the number of calls per caller. Moreover, the majority population in the training set was not always the majority population in the test set. Thus, the majority classifier accuracy could be greater or less than 50%.

To determine if vocal differences between the two populations could be an artefact of the different recording equipment used in each population, we ran the same model using the f_1 contour measurements instead of the mel spectrogram measurements. We also ran a logistic regression model with population as the response variable and the f_1 contour measurements and caller age as the regressors, to determine which acoustic features had a significant relationship to population.

To test the hypotheses that elephants exhibit vocal signatures of group identity at the bond group or core group level, we ran two additional random forest models (same hyperparameters) to predict bond group and core group, respectively, as a function of the mel spectrogram acoustic features, using only data from Samburu. To ensure that the bond group model could only use acoustic features that generalized across the entire bond group, rather than features specific to core groups, to predict bond group, we restricted the dataset to bond groups that contained at least two core groups with at least five calls each in our dataset (six bond groups, 931 calls) [57]. For each iteration of the model, we randomly selected one core group from each bond group and allocated all calls from those core groups to the test set. To ensure that the core group model could only use acoustic features that generalized across the entire core group, rather than features specific to individual callers, to predict core group, we restricted the dataset to core groups that contained at least two individuals with at least five calls each in our dataset (seven core groups, 794 calls) [57]. For each iteration of the model, we randomly selected 20% of the callers with at least five calls each from each core group and allocated all calls from those individuals to the test set. The data were severely imbalanced across bond groups and core groups, but we were unable to balance it by subsampling the more numerous classes because doing so would have reduced the sample size excessively. We ran 10 000 iterations for each model, calculating the classification accuracy and p-value for each run as before. The mean ± s.d. proportion of the calls allocated to the test set was 0.32 ± 0.17 for the bond group model and 0.23 ± 0.08 for the core group model.

To test the hypothesis that elephant rumbles are individually specific, we ran a fifth random forest (same hyperparameters) to predict individual caller identity as a function of the mel spectrogram acoustic features. As calls produced by the same caller on the same date might exhibit similar features due to temporary circumstances such as the caller's internal state, behavioural context and ambient conditions, we randomly selected one date for each caller and held out all calls from these caller dates as the test set [49,57]. We used callers from both populations for this analysis, but only included callers that produced at least three calls on at least two different dates each (427 calls from 15 callers in Samburu, 218 calls from nine callers in Amboseli). We calculated the classification accuracy and p-value for each of 10 000 iterations as before. The mean ± s.d. proportion of calls allocated to the test set was 0.20 ± 0.02.

2.5. Assessment of call similarity

Elephant rumbles vary with the behavioural context and the individual identity and age of the caller [41,59,60]. To determine if there were acoustic differences among populations, bond groups or core groups that could not be explained by behavioural context, caller identity or caller age, we calculated random forest proximity scores between each possible pair of calls. The random forest proximity score for a given pair of calls was the proportion of trees for which both calls were classified in the same terminal node, adjusted for the size of the node, and represented a metric of call similarity in terms of the acoustic features most relevant to predicting the response variable [61]. We calculated proximity scores from four different random forests (population with mel spectrogram features, population with f_1 contour features, bond group and core group), using the same hyperparameters and subsets of the data as before except that we increased the number of trees to 8000 and did not balance the data by subsampling or hold out any observations as a test set (population models: all observations, n = 1352; bond group model: all bond groups in Samburu containing at least two core groups with at least five calls each, n = 931; core group model: all core groups in Samburu containing at least two individuals with at least five calls each, n = 794).

For each set of proximity scores, we ran a generalized linear mixed model with a gamma error distribution and a log link function. In each model, pairwise call proximity score was the response variable and pair ID (unique identifier for a given pair of callers) was a random effect. Fixed effects were 'pair class' (whether the two calls in a given pair came from the same population, bond group or core group, depending on the model) and the scaled and centred absolute value of the age difference between the two callers. For the population models, the 'pair class' variable was binary: a pair of calls could either be from the same population or not. For the bond group and core group models, the 'pair class' variable had three possible states: same core group, different core groups within the same bond group or different bond groups within the Samburu population. To ensure that differences between populations or social groups were not an artefact of individual identity or behavioural context, we only included pairs of calls with the same behavioural context and different callers. We also only included calls for which we were certain of the behavioural context (see electronic supplementary material, Supplemental Methods). This resulted in a sample size of 149 640 call pairs for the two population models, 97 586 call pairs for the bond group model and 72 437 call pairs for the core group model. As proximity scores could be 0, we added 0.00001 to all proximity scores so all the values would be positive. For the bond group and core group models, we used the R package 'emmeans' [62] to examine the pairwise contrasts for each level of the 'pair class' variable.

To assess whether vocal similarity between individuals was better explained by social affiliation or genetic relatedness, we ran two additional gamma regressions on call pairs, modelling call proximity score as a function of 'binary pair class' (see below), caller age difference and caller genetic relatedness. One gamma model used proximity scores extracted from the random forest trained to predict core groups and the other used proximity scores extracted from the random forest trained to predict bond groups. These two sets of proximity scores represented the pairwise similarity between calls in terms of the features most relevant to predicting core group membership or bond group membership, respectively. For the model using proximity scores extracted from the bond group random forest, binary pair class indicated whether the two calls in a pair were from the same bond group. For the model using proximity scores extracted from the core group random forest, binary pair class indicated whether the two calls in a pair were from the same core group. These gamma models could only be run on the subset of callers for which we had genetic relatedness data (13 individuals for the bond group model, eight for the core group model). Genetic relatedness was calculated in a previously published study using 20 microsatellite loci extracted from tissue or dung samples collected before 2006 [37]. Due to social disruption caused by poaching, many core groups and bond groups in the Samburu population now include unrelated individuals, which made it possible to independently assess the effects of social affiliation and relatedness on call similarity [37].

All statistical analyses were performed in R v. 4.2.2 [51] and 0.05 was used as the significance threshold for all tests.
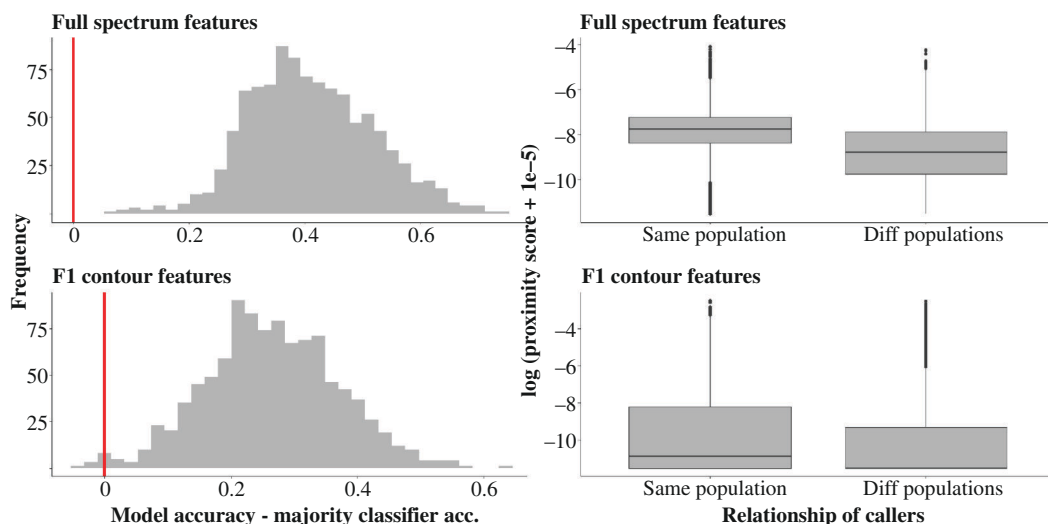
Figure 2. Evidence for acoustic differences between Amboseli and Samburu populations. Top row: models using acoustic features derived from mel spectrogram. Bottom row: models using acoustic features derived from second harmonic contour. Histograms represent distribution of classification accuracies of 10 000 random forest models trained to predict population from the acoustic features. Positive values (to the right of the red line) indicate that the random forest outperformed the majority classifier (binomial exact tests; median $p < 0.0001$ for both models). Boxplots represent acoustic similarity of calls within versus between populations (gamma regression; $p < 0.0001$ for both models). Acoustic similarity is represented by call proximity scores extracted from the random forest. The log of the proximity scores is plotted on the y-axis to facilitate visualization of the data (less negative values = greater call similarity). Centre lines = medians, box edges = interquartile ranges, whiskers = $1.5 \times$ interquartile range.

Table 2. Results of logistic regression modelling population as a function of acoustic features derived from the second harmonic contour (Hz) and caller age (days). χ^2 statistics and p-values are derived from analysis of deviance on the model. Significant p-values in bold.

regressor	coefficient	χ^2 statistic	p-value
mean of second harmonic	−0.664	4.96	**0.026**
s.d. of second harmonic	1.16	8.31	**0.004**
skew of second harmonic	−0.0335	0.0628	0.802
kurtosis of second harmonic	0.0440	0.688	0.407
10th percentile of second harmonic	−0.000548	0.0000	0.997
90th percentile of second harmonic	0.0101	0.0027	0.959
frequency at 25% of call duration	0.0617	0.513	0.474
frequency at 50% of call duration	0.0758	0.675	0.411
frequency at 75% of call duration	0.358	16.9	**<0.0001**
caller age	−0.0000692	19.6	**<0.0001**

3. Results

3.1. Vocal differences between allopatric populations

The random forest model trained to predict population from mel spectrogram acoustic features was significantly more accurate than the majority classifier that always guessed the population with the most calls in the training set (median observed classification accuracy = 0.79, median majority classifier accuracy = 0.39, median $p < 0.0001$) (figure 2). Similarly, the random forest model trained to predict population from f_1 contour features, which describe less of the variation in the call but are robust to influences of the recording equipment, was significantly more accurate than the majority classifier (median observed classification accuracy = 0.65, median majority classifier accuracy = 0.39, median $p < 0.0001$) (figure 2). Three parameters of the f_1 contour had a significant relationship with population in
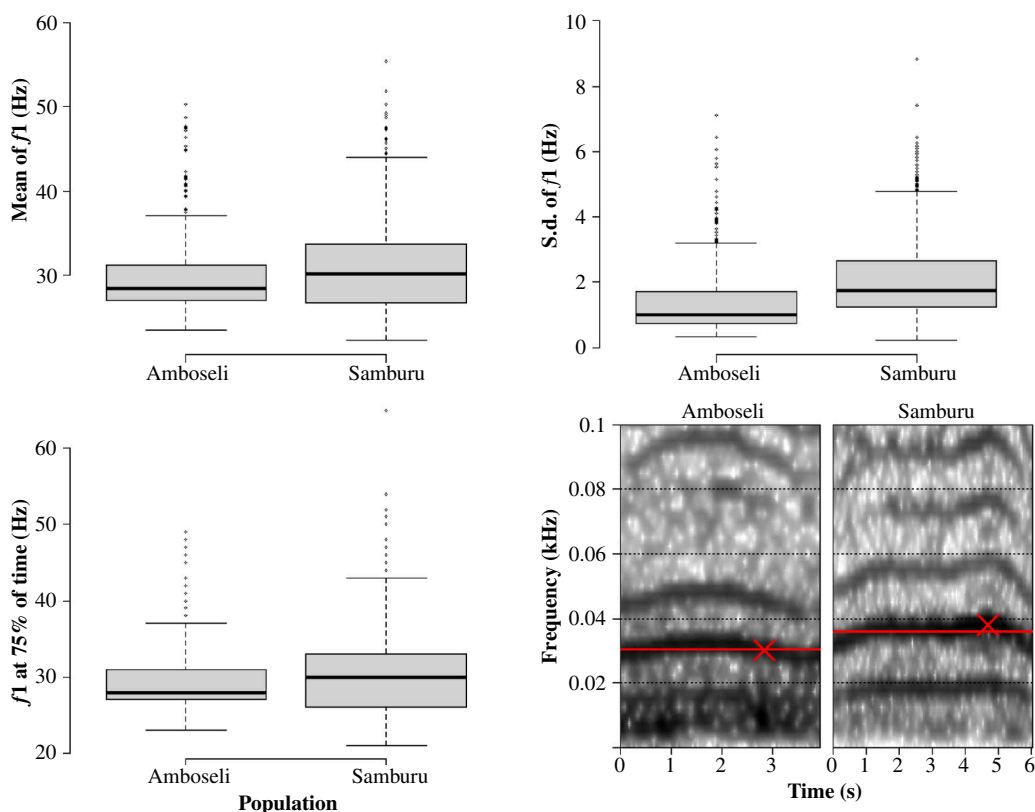
Figure 3. Population differences in the second harmonic (f_1) contour. Boxplots represent population differences in the three features of the second harmonic that differed between Amboseli and Samburu (logistic regression; mean of second harmonic: $p = 0.026$; s.d. of second harmonic: $p = 0.004$; frequency of second harmonic at 75% of call duration: $p < 0.0001$). Bottom right: spectrograms of rumbles from the same behavioural context (contact calling) made by a 17-year-old female in each population (2 kHz sampling rate, Hanning window, 800 samples/window, 90% overlap). Red lines: mean frequency of second harmonic; red Xs: frequency of second harmonic at 75% of call duration.

a logistic regression model after controlling for caller age: mean of the second harmonic ($x^2 = 5.0$, $p = 0.026$), s.d. of the second harmonic ($x^2 = 8.3$, $p = 0.004$), and value of the second harmonic at 75% of call duration ($x^2 = 16.9$, $p < 0.0001$). All three values were higher in Samburu than in Amboseli on average (figure 3, table 2).

Calls from different individuals within the same population were significantly more similar than calls from different populations after controlling for behavioural context, individual caller identity and the age difference between the callers (gamma regression; proximity scores calculated from mel spectrogram features: $x^2 = 493.0$, $p < 0.0001$; proximity scores calculated from f_1 contour features: $x^2 = 41.3$, $p < 0.0001$) (figure 2, table 3). This indicates that the vocal divergence between Samburu and Amboseli was not merely an artefact of differences in the age structure or prevalence of certain behavioural contexts in the two populations, but rather reflects a population-specific vocal difference. Call similarity also decreased as the age difference between the callers increased (gamma regression; proximity scores calculated from mel spectrogram features: $x^2 = 35.2$, $p < 0.0001$; proximity scores calculated from f_1 contour features: $x^2 = 80.8$, $p < 0.0001$) (table 3).

3.2. Vocal differences among sympatric social groups

The random forest trained to predict bond group in Samburu from mel spectral features correctly predicted bond group for 10% of calls on average, which was significantly more accurate than the majority classifier (median observed classification accuracy = 0.10, median majority classifier accuracy = 0.05, median $p = 0.001$) (figure 4). The random forest trained to predict core group in Samburu from mel spectral features correctly predicted core group for 19% of calls on average, which was no better than the majority classifier (median observed classification accuracy = 0.19, median majority classifier accuracy = 0.19, median $p = 0.70$) (figure 4). Analysis of call proximity scores indicated that

Table 3. Results of gamma regressions with call proximity score as the response variable. Rows are separate models and columns are regressors. Values in each cell include the coefficient, χ^2 statistic, and p-value for fixed effects (χ^2 and p based on analysis of deviance) and the s.d. for random effects. Caller age was scaled to make it more comparable in range to other regressors. Pair ID represented a unique pair of callers.

model #	RF used to generate proximity scores	levels of pair class (reference level bold)	pair class	caller age (scaled)	genetic relatedness	pair ID (random effect)
1	population ~ mel spectral features	same versus **different** population	Coef = 0.518; χ^2 = 493.0; $p <$ 0.0001	Coef = − 0.059; χ^2 = 35.2; p < 0.0001	NA	0.535
2	population ~ F1 contour features	same versus **different** population	Coef = 0.381; χ^2 = 41.3; $p <$ 0.0001	Coef = −0.233; χ^2 = 80.8; p < 0.0001	NA	5.074
3	bond group ~ mel spectral features	same core group versus same bond group versus **different bond groups**	Coef (same core group) = 0.181; Coef (same bond group) = 0.049; χ^2 = 13.2; p = 0.001	Coef = −0.111; χ^2 = 68.0; p < 0.0001	NA	0.658
4	core group ~ mel spectral features	same core group versus same bond group versus **different bond groups**	Coef (same core group) = 0.228; Coef (same bond group) = −0.061; χ^2 = 26.5; $p <$ 0.0001	Coef = −0.154; χ^2 = 82.8; p < 0.0001	NA	0.543
5	bond group ~ mel spectral features	same versus **different** bond group	Coef = 0.309; χ^2 = 3.54; p = 0.060	Coef = −0.048; χ^2 = 2.13; p = 0.144	Coef = 0.146; χ^2 = 0.133; p = 0.715	0.508
6	core group ~ mel spectral features	same versus **different** core group	Coef = 0.374; χ^2 = 5.85; p = 0.016	Coef = 0.013; χ^2 = 0.181; p = 0.670	Coef = 0.382; χ^2 = 1.49; p = 0.222	0.375

calls recorded from the same core group were more similar than calls recorded from different core groups in the same bond group or calls recorded from different bond groups, suggesting some vocal convergence at the core group level (tables 3 and 4). By contrast, calls from different core groups in the same bond group were no more similar than calls from different bond groups, suggesting a lack of vocal convergence at the bond group level (tables 3 and 4). In both the bond group model and the core group model, call similarity decreased as caller age difference increased (table 3).

3.3. Vocal differences among individual callers

The random forest trained to predict individual caller ID from mel spectral features performed significantly better than the majority classifier (median classification accuracy = 0.29, median majority classifier accuracy = 0.03, median p < 0.0001) (table 3, electronic supplementary material, figure S1).

3.4. Social versus genetic influences on call similarity

Social group membership was a better predictor of call similarity than genetic relatedness. For the proximity scores extracted from the random forest trained to predict core group, belonging to the same
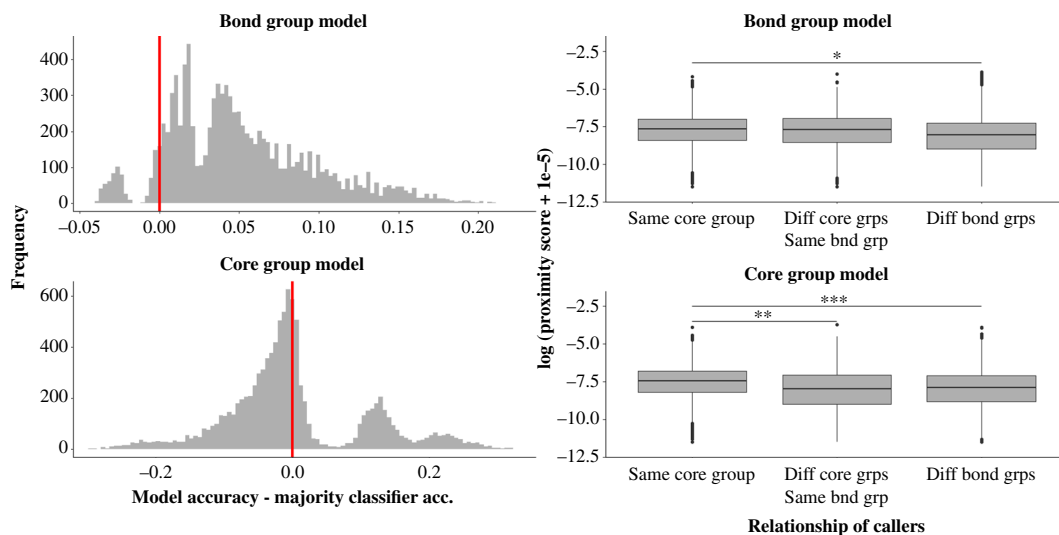
Figure 4. Limited evidence for vocal differences among sympatric social groups. Histograms represent distribution of classification accuracies of 10 000 random forest models trained to predict bond group or core group from the acoustic features. Positive values (to the right of the red line) indicate that the random forest outperformed the majority classifier (binomial exact tests; bond group model: median $p = 0.001$; core group model: median $p = 0.70$). Boxplots represent acoustic similarity of calls in same core group versus same bond group versus different bond groups (gamma regression, post hoc contrasts: $*p < 0.01$, $**p < 0.001$, $***p < 0.0001$). Acoustic similarity is represented by call proximity scores extracted from the random forest. The log of the proximity scores is plotted on the y-axis to facilitate visualization of the data (less negative values = greater call similarity). Centre lines = medians, box edges = interquartile ranges, whiskers = 1.5 ∗ interquartile range.

Table 4. Post hoc contrasts indicating the statistical significance of acoustic differences between core groups and bond groups. Both models in this table were of the form call proximity score (similarity score between a pair of calls) ~ pair class + scaled caller age difference + (1|pair ID), where pair class was a three-way categorical variable indicating whether the two calls in question were recorded from the same core group, different core groups in the same bond group or different bond groups. The models differed in which random forest (RF) model was used to generate the call proximity scores. In model 3, the proximity scores represented the pairwise similarity of calls in terms of the acoustic features most relevant to predicting bond group membership, while in model 4 the proximity scores represented the pairwise similarity of calls in terms of the acoustic features most relevant to predicting core group membership.

model #	RF used to generate proximity scores	same core group versus different core groups in same bond group	same core group versus different bond groups	different core groups in same bond group versus different bond groups
3	bond group ~ mel spectral features	$p = 0.144$	$p = 0.001$	$p = 0.627$
4	core group ~ mel spectral features	$p = 0.0001$	$p < 0.0001$	$p = 0.536$

core group was a significant predictor of call proximity (gamma regression, $\chi^2 = 5.9$, $p = 0.016$), but genetic relatedness was not (gamma regression, $\chi^2 = 1.5$, $p = 0.222$) (electronic supplementary material, figure S2, table 3). For the proximity scores extracted from the random forest trained to predict bond group, belonging to the same bond group was a marginally non-significant predictor of call similarity (gamma regression, $\chi^2 = 3.5$, $p = 0.060$) and genetic relatedness was not significant (gamma regression, $\chi^2 = 0.133$, $p = 0.715$) (electronic supplementary material, figure S2, table 3).

4. Discussion

The vocal flexibility of elephants is notable, but few studies have assessed geographic and social group variation in wild elephant calls. Our results provide evidence that elephant rumbles differ in structure between allopatric populations and suggest at least some vocal divergence among sympatric

core groups. Our results also indicate that to the extent that elephants exhibit vocal differences among social groups, these differences are probably due to social factors rather than genetically determined, as group membership, but not genetic relatedness, was a significant predictor of call similarity.

As 92% of the calls we recorded in Amboseli came from a single social unit, the effects of population and social group were somewhat confounded. However, random forest models were able to achieve much higher discriminability between the Amboseli and Samburu populations than among social groups within Samburu. This suggests that the two populations are more vocally divergent than social groups within a single population. Previous work has shown that elephant populations differ in the most common order of call combinations [63], and our present findings add to this by showing that populations also differ in the fine acoustic structure of rumbles.

There are several possible explanations for the difference in call structure between Amboseli and Samburu. Cultural evolution is a common driver of vocal differences between animal populations. For example, white-throated sparrows (*Zonotrichia albicollis*) and yellow-naped amazon parrots (*Amazona auropalliata*) exhibit regional dialects that probably stem partially from copying errors during vocal production learning [7,64]. Chimpanzees exhibit population differences in the order of call combinations that probably result from usage learning [26]. In elephants, vocal production learning seems more likely than usage learning to explain the population differences we observed, given that they were structural differences in a single call type. However, we cannot definitively rule out the possibility that Amboseli and Samburu elephants share the same repertoire of rumble subtypes but differ in the contexts in which each subtype is used, thus leading to an average difference in the fundamental frequency of rumbles recorded from the two populations.

Amboseli and Samburu exhibit significant genetic divergence in mitochondrial DNA (F_{ST} = 0.423, p < 0.0001) and to a lesser extent nuclear microsatellites (F_{ST} = 0.02, p < 0.05), so genetics may play a role as well [65]. Moreover, if Samburu elephants were more stressed than Amboseli elephants on average during the periods in which we recorded them, this might explain why Samburu rumbles were higher in pitch, as stress is correlated with elevated fundamental frequency in elephant rumbles [66]. Acoustic adaptation also cannot be definitively ruled out. If Samburu has more low-frequency noise for some reason, the Samburu elephants might have shifted the frequency of their rumbles upwards to compensate, as has been documented in many other species [67]. However, there is no obvious reason to expect differences in the acoustic environments of Samburu and Amboseli given that both are protected areas with similar habitat types surrounded by relatively low-density pastoralist communities [47].

One explanation that probably can be excluded is body size. There is little difference in growth asymptotes among female elephants in these two populations [68], and the vocal differences persisted when controlling for age. To the extent that there is any difference in body size between the populations, Samburu elephants are slightly taller on average [68], which if anything would be expected to result in lower fundamental frequencies [21]. However, we observed the opposite.

Elephants are often translocated between populations to facilitate gene flow, reduce ecological damage caused by local overcrowding or mitigate human–elephant conflict [32,69]. Our finding that elephant populations less than 400 km apart exhibit significant differences in the structure of their rumble vocalizations could thus have important implications for elephant conservation, especially if the differences are found to result from cultural or genetic divergence, rather than acute responses to local environmental conditions. A prior study found that elephants reacted to seismically transmitted playback of alarm rumbles recorded in their own population but ignored alarm rumbles from a different population, although it is unclear if this was due to failure to recognize alarm calls from another population, increased responsiveness to familiar callers or differences between the stimuli unrelated to population of origin [70]. Further study is warranted to determine whether vocal differences between elephant populations impede communication or social integration. If so, translocation efforts should attempt to quantify and account for behavioural compatibility between populations when deciding where to move individuals, to the extent that it is feasible to do so.

We found evidence for limited vocal differences among sympatric social groups. Analysis of call proximity scores indicated statistically significant vocal convergence among members of the same core group but found no evidence for vocal convergence among group members at the bond group level. Our random forest models correctly predicted the core group for 19% of calls and correctly predicted the bond group for 10% of calls, but only the bond group model was statistically significantly better than the majority classifier. This is probably because the training data were more unbalanced for the core group model on average, so the majority classifier was more accurate for this model, thus raising the threshold for a statistically significant improvement over the majority classifier.

Overall, our results suggest that elephants exhibit greater vocal convergence at the core group level than at the bond group level. This is similar to some cetaceans and bats, in which the greatest vocal convergence takes place at the closest tier of social affiliation [8,14,40]. However, it contrasts with geladas, in which vocal convergence takes place at the band level, a higher tier of social organization [17]. In geladas, bands typically forage and sleep together, and the function of call convergence at the band level is hypothesized to be maintaining social cohesion among less closely knit social affiliates who must coordinate their behaviour on a daily basis [17]. By contrast, in elephants, bond group members from different core groups are more often separate than together, so there is less opportunity and potentially less need to develop vocal convergence at the bond group level [35].

The low accuracy of the random forest models in predicting social group membership from the acoustic structure of rumbles suggests that while detectable vocal differences among social groups do exist (at least at the core group level), they probably have limited potential for facilitating social recognition. Vocal convergence among core group members in elephants may not have an adaptive function; for example, it could be a selectively neutral by-product of vocal learning that evolved for some other purpose. Meerkats have group signatures in their close calls but fail to discriminate them, indicating that vocal differences among social groups can develop even if they do not facilitate recognition [28]. Previous work has shown that elephants can discriminate between the calls of close social affiliates (core/bond group members), distant social affiliates (clan members) and non-affiliates [39]. If elephants do not rely on group signatures to make this discrimination, that means they can individually recognize the calls of at least 100 other elephants on average, including individuals with whom they have limited interaction, suggesting that they possess exceptional social memory [39].

The vocal differences we observed among elephant social groups and populations reflected minor variations in the fine structure of a single shared call type. This is more similar to the subtle group signatures found in some bats [13,71], ungulates [15,16], meerkats [28] and non-human primates [72] than to the dialects of some cetaceans and birds, where individuals from different social groups or geographical areas use categorically distinct repertoires of call/song types [2,7,8,40]. While we did not include call types other than rumbles in this study, rumbles comprise the vast majority of elephant vocalizations [41], and we have not noticed any differences among social groups or populations in the total repertoire of basic call types.

It is sometimes difficult to determine if subtle modifications to species-specific call types involve vocal production learning. For example, one study found that translocated captive chimpanzees (*Pan troglodytes*) slightly modified their food calls to be more like those of their new group members [73], but this may simply reflect changes in arousal [74]. We think it is unlikely that the vocal differences we observed among sympatric elephant core groups can be explained by group differences in arousal or other physiological states, as these groups are all subject to similar environmental stressors and we controlled for behavioural context in our analysis of call proximity scores, but we cannot conclusively rule it out.

Learning slight modifications to existing call types seems to require less neurological specialization than learning entirely new call types, which explains why many species traditionally considered non-vocal production learners, such as non-human primates, are capable of acquiring vocal group signatures [75]. Yet unlike these species, some captive elephants have displayed a remarkable ability to mimic novel heterospecific sounds, far beyond what is necessary to develop the subtle vocal convergence we observed in this study [33,34]. Similar to humans, who exhibit both subtle vocal convergence among social affiliates and more sophisticated forms of vocal production learning [72], it may be that wild elephants use vocal learning in multiple ways. The recent discovery that elephants address one another with name-like calls suggests another possible function for vocal production learning in elephants [49].

Our study replicates previous findings that rumbles are individually distinct [42–46] and that rumble structure is correlated with caller age [59]. However, our random forest model predicting individual caller identity from call structure only achieved 29% classification accuracy on average, in contrast to previous studies on captive elephants that achieved 55–60% accuracy using discriminant function analysis [42,45] and 82.5% accuracy using hidden Markov models [43]. This difference in performance may be an artefact of the number of individual callers included in each dataset (24 for our study versus 6 or 13 for previous studies). It is also possible that rumbles produced under natural conditions are more variable and therefore more difficult to assign to individual callers than rumbles produced in captivity, or that our field recordings were noisier and more difficult to classify than recordings made in captivity.

This study is among the first to investigate possible consequences of vocal learning in wild elephants and adds to a growing body of literature on vocal variation and plasticity in mammals [75,76]. Interspecific variation in elephant calls is well-documented [63,77,78], and one study described intraspecific variation in the syntax of elephant call combinations [63]. Geographic variation in vocal structure has also been documented for elephants' closest living relatives: sirenians and hyraxes [79,80]. However, prior to the present study, intraspecific variation in the structure of a single call type was largely unexplored in elephants. Longitudinal studies on translocated individuals and individuals who change groups within a population due to poaching or other social disruption could help clarify the role of vocal learning in wild elephant communication. Playback experiments could determine if elephant calls differ in function as well as form across populations.

References

1. Benedict L, Bowie RCK. 2009 Macrogeographical variation in the song of a widely distributed African warbler. *Biol. Lett.* **5**, 484–487. (doi:10.1098/rsbl.2009.0244)

2. Pitocchelli J. 2011 Macrogeographic variation in the song of the mourning warbler (*Oporornis philadelphia*). *Can. J. Zool.* **89**, 1027–1040. (doi:10.1139/z11-077)

3. Kapoor JA. 2016 The functional significance of microgeographic dialects in a hermit hummingbird. PhD thesis, Cornell University, Ithaca, NY.

4. MacDougall-Shackleton EA, Derryberry EP, Hahn TP. 2002 Nonlocal male mountain white-crowned sparrows have lower paternity and higher parasite loads than males singing local dialect. *Behav. Ecol.* **13**, 682–689. (doi:10.1093/beheco/13.5.682)

5. Ey E, Fischer J. 2009 The 'acoustic adaptation hypothesis'—a review of the evidence from birds, anurans and mammals. *Bioacoustics* **19**, 21–48. (doi:10.1080/09524622.2009.9753613)

6. Campbell P, Pasch B, Pino JL, Crino OL, Phillips M, Phelps SM. 2010 Geographic variation in the songs of neotropical singing mice: testing the relative importance of drift and local adaptation. *Evolution.* **64**, 1955–1972. (doi:10.1111/j.1558-5646.2010.00962.x)

7. Ramsay SM, Otter KA. 2015 Geographic variation in white-throated sparrow song may arise through cultural drift. *J. Ornithol.* **156**, 763–773. (doi:10.1007/s10336-015-1183-8)

8. Yurk H, Barrett-Lennard LG, Ford JKB, Matkin CO. 2002 Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Anim. Behav.* **63**, 1103–1119. (doi:10.1006/anbe.2002.3012)

9. Rendell LE, Whitehead H. 2003 Vocal clans in sperm whales (*Physeter macrocephalus*). *Proc. R. Soc. B* **270**, 225–231. (doi:10.1098/rspb.2002.2239)

10. Van Cise AM, Mahaffy SD, Baird RW, Mooney TA, Barlow J. 2018 Song of my people: dialect differences among sympatric social groups of short-finned pilot whales in Hawai'i. *Behav. Ecol. Sociobiol.* **72**. (doi:10.1007/s00265-018-2596-1)

11. Hile A, Plummer T, Striedter G. 2000 Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Anim. Behav.* **59**, 1209–1218. (doi:10.1006/anbe.1999.1438)

12. Keen SC, Meliza CD, Rubenstein DR. 2013 Flight calls signal group and individual identity but not kinship in a cooperatively breeding bird. *Behav. Ecol.* **24**, 1279–1285. (doi:10.1093/beheco/art062)

13. Knörnschild M, Nagy M, Metz M, Mayer F, von Helversen O. 2012 Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Anim. Behav.* **84**, 761–769. (doi:10.1016/j.anbehav.2012.06.029)

14. Boughman JW, Wilkinson GS. 1998 Greater spear-nosed bats discriminate group mates by vocalizations. *Anim. Behav.* **55**, 1717–1732. (doi:10.1006/anbe.1997.0721)

15. Volodin IA, Volodina EV, Lapshina EN, Efremova KO, Soldatova NV. 2014 Vocal group signatures in the goitred gazelle *Gazella subgutturosa*. *Anim. Cogn.* **17**, 349–357. (doi:10.1007/s10071-013-0666-3)

16. Briefer EF, McElligott AG. 2012 Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Anim. Behav.* **83**, 991–1000. (doi:10.1016/j.anbehav.2012.01.020)

17. Painter MC, Gustison ML, Snyder-Mackler N, Tinsley Johnson E, le Roux A, Bergman TJ. 2024 Acoustic variation and group level convergence of gelada, *Theropithecus gelada*, contact calls. *Anim. Behav.* **207**, 235–246. (doi:10.1016/j.anbehav.2023.10.002)

18. Carlson NV, Kelly EM, Couzin I. 2020 Individual vocal recognition across taxa: a review of the literature and a look into the future. Phil. Trans. R. Soc. B **375**, 20190479. (doi:10.1098/rstb.2019.0479)

19. Colombelli-Négrel D, Hauber ME, Robertson J, Sulloway FJ, Hoi H, Griggio M, Kleindorfer S. 2012 Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. *Curr. Biol.* **22**, 2155–2160. (doi:10.1016/j.cub.2012.09.025)

20. Hersh TA *et al*. 2022 Evidence from sperm whale clans of symbolic marking in non-human cultures. *Proc. Natl Acad. Sci. USA* **119**, e2201692119. (doi:10.1073/pnas.2201692119)

21. Ey E, Pfefferle D, Fischer J. 2007 Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. *Primates.* **48**, 253–267. (doi:10.1007/s10329-006-0033-y)

22. Wright TF, Wilkinson GS. 2001 Population genetic structure and vocal dialects in an Amazon parrot. Proc. R. Soc. B **268**, 609–616. (doi:10.1098/rspb.2000.1403)

23. Ríos-Chelén AA, Salaberria C, Barbosa I, Macías Garcia C, Gil D. 2012 The learning advantage: bird species that learn their song show a tighter adjustment of song to noisy environments than those that do not learn. *J. Evol. Biol.* **25**, 2171–2180. (doi:10.1111/j.1420-9101.2012.02597.x)

24. Colombelli-Négrel D, Smale R. 2018 Habitat explained microgeographic variation in little penguin agonistic calls. *Auk* **135**, 44–59. (doi:10.1642/AUK-17-75.1)

25. Rendell L, Mesnick SL, Dalebout ML, Burtenshaw J, Whitehead H. 2012 Can genetic differences explain vocal dialect variation in sperm whales, *Physeter macrocephalus*? *Behav. Genet.* **42**, 332–343. (doi:10.1007/s10519-011-9513-y)

26. Girard-Buttoz C, Bortolato T, Laporte M, Grampp M, Zuberbühler K, Wittig RM, Crockford C. 2022 Population-specific call order in chimpanzee greeting vocal sequences. *iSci.* **25**, 104851. (doi:10.1016/j.isci.2022.104851)

27. Vernes SC, Kriengwatana BP, Beeck VC, Fischer J, Tyack PL, ten Cate C, Janik VM. 2021 The multi-dimensional nature of vocal learning. *Phil. Trans. R. Soc. B.* **376**, 20200236. (doi:10.1098/rstb.2020.0236)

28. Townsend SW, Hollén LI, Manser MB. 2010 Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. *Anim. Behav.* **80**, 133–138. (doi:10.1016/j.anbehav.2010.04.010)

29. Fischer J, Wegdell F, Trede F, Dal Pesco F, Hammerschmidt K. 2020 Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. Proc. R. Soc. B **287**, 20202531. (doi:10.1098/rspb.2020.2531)

30. Marler P, Peters S. 1977 Selective vocal learning in a sparrow. *Science* **198**, 519–521. (doi:10.1126/science.198.4316.519)

31. Zann R, Dunstan E. 2008 Mimetic song in superb lyrebirds: species mimicked and mimetic accuracy in different populations and age classes. *Anim. Behav.* **76**, 1043–1054. (doi:10.1016/j.anbehav.2008.05.021)

32. Pinter-Wollman N, Isbell LA, Hart LA. 2009 Assessing translocation outcome: comparing behavioral and physiological aspects of translocated and resident African elephants (*Loxodonta africana*). *Biol. Conserv.* **142**, 1116–1124. (doi:10.1016/j.biocon.2009.01.027)

33. Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S. 2005 Elephants are capable of vocal learning. *Nature* **434**, 455–456. (doi:10.1029/2001GL014051)

34. Stoeger AS, Mietchen D, Oh S, de Silva S, Herbst CT, Kwon S, Fitch WT. 2012 An Asian elephant imitates human speech. *Curr. Biol.* **22**, 2144–2148. (doi:10.1016/j.cub.2012.09.022)

35. Wittemyer G, Douglas-Hamilton I, Getz WM. 2005 The socioecology of elephants: analysis of the processes creating multitiered social structures. *Anim. Behav.* **69**, 1357–1371. (doi:10.1016/j.anbehav.2004.08.018)

36. Moss CJ, Poole JH. 1983 Relationships and social structure in African elephants. In *Primate social relationships: an integrated approach* (ed. RA Hinde), pp. 315–325. Oxford, UK: Blackwell Science, Inc.

37. Wittemyer G, Okello JBA, Rasmussen HB, Arctander P, Nyakaana S, Douglas-Hamilton I, Siegismund HR. 2009 Where sociality and relatedness diverge: the genetic basis for hierarchical social organization in African elephants. Proc. R. Soc. B **276**, 3513–3521. (doi:10.1098/rspb.2009.0941)

38. Moss C. 1981 Social circles. *Wildlife News.* **16**.

39. McComb K, Moss C, Sayialel S, Baker L. 2000 Unusually extensive networks of vocal recognition in African elephants. *Anim. Behav.* **59**, 1103–1109. (doi:10.1006/anbe.2000.1406)

40. Gero S, Whitehead H, Rendell L. 2016 Individual, unit and vocal clan level identity cues in sperm whale codas. *R. Soc. Open Sci.* **3**, 150372. (doi:10.1098/rsos.150372)

41. Poole JH. 2011 Behavioral contexts of elephant acoustic communication. In *The Amboseli elephants: a long-term perspective on a long-lived mammal* (eds CJ CJ Moss, H Croze, PC Lee), pp. 125–159. Chicago, IL: University of Chicago Press. (doi:10.7208/chicago/9780226542263.003.0009)

42. Soltis J, Leong K, Savage A. 2005 African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. *Anim. Behav.* **70**, 589–599. (doi:10.1016/j.anbehav.2004.11.016)

43. 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**, 956–963. (doi:10.1121/1.1847850)

44. de Silva S. 2010 Acoustic communication in the Asian elephant, *Elephas maximus maximus*. *Behaviour* **147**, 825–852. (doi:10.1163/000579510X495762)

45. Stoeger AS, Baotic A. 2016 Information content and acoustic structure of male African elephant social rumbles. *Sci. Rep.* **6**, 27585. (doi:10.1038/srep27585)

46. Wierucka K, Henley MD, Mumby HS. 2021 Acoustic cues to individuality in wild male adult African savannah elephants (*Loxodonta africana*). *PeerJ* **9**, e10736. (doi:10.7717/peerj.10736)

47. Chiyo PI, Grieneisen LE, Wittemyer G, Moss CJ, Lee PC, Douglas-Hamilton I, Archie EA. 2014 The influence of social structure, habitat, and host traits on the transmission of *Escherichia coli* in wild elephants. *PLoS One* **9**, e93408. (doi:10.1371/journal.pone.0093408)

48. Pardo M, Lolchuragi D, Poole J, Granli P, Moss C, Douglas-Hamilton I, Wittemyer G. 2024 Female african elephant rumbles differ between populations and sympatric social groups. Dryad Digital Repository. (doi:10.5061/dryad.rv15dv4dt)

49. Pardo MA, Fristrup K, Lolchuragi DS, Poole JH, Granli P, Moss C, Douglas-Hamilton I, Wittemyer G. 2024 African elephants address one another with individually specific name-like calls. *Nat. Ecol. Evol.* **8**, 1353–1364. (doi:10.1038/s41559-024-02420-w)

50. Ginsberg JR, Young TP. 1992 Measuring association between individuals or groups in behavioural studies. *Anim. Behav.* **44**, 377–379. (doi:10.1016/0003-3472(92)90042-8)

51. R Core Team. 2022 R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.

52. Stevens SS, Volkmann J, Newman EB. 1937 A scale for the measurement of the psychological magnitude pitch. *J. Acoust. Soc. Am.* **8**, 185–190. (doi:10.1121/1.1915893)

53. Heffner RS, Heffner HE. 1982 Hearing in the elephant (*Elephas maximus*): absolute sensitivity, frequency discrimination, and sound localization. *J. Comp. Physiol. Psychol.* **96**, 926–944.

54. Maryn Y, Ysenbaert F, Zarowski A, Vanspauwen R. 2017 Mobile communication devices, ambient noise, and acoustic voice measures. *J. Voice* **31**, 248.(doi:10.1016/j.jvoice.2016.07.023)

55. Breiman L. 2001 Random forests. *Mach. Learn.* **45**, 5–32. (doi:10.1109/ICCECE51280.2021.9342376)

56. Bader-El-Den M, Teitei E, Perry T. 2019 Biased random forest for dealing with the class imbalance problem. *IEEE Trans. Neural Netw. Learn. Syst.* **30**, 2163–2172. (doi:10.1109/TNNLS.2018.2878400)

57. Lehmann KDS, Jensen FH, Gersick AS, Strandburg-Peshkin A, Holekamp KE. 2022 Long-distance vocalizations of spotted hyenas contain individual, but not group, signatures. *Proc. R. Soc. B* **289**, 20220548. (doi:10.1098/rspb.2022.0548)

58. Sim J, Mcgoverin C, Oey I, Frew R, Kebede B. 2023 Stable isotope and trace element analyses with non-linear machine-learning data analysis improved coffee origin classification and marker selection. *J. Sci. Food Agric.* **103**, 4704–4718. (doi:10.1002/jsfa.12546)

59. Stoeger AS, Zeppelzauer M, Baotic A. 2014 Age group estimation in free-ranging African elephants based on acoustic cues of low-frequency rumbles. *Bioacoustics* **23**, 231–246. (doi:10.1080/09524622.2014.888375)

60. McComb K, Reby D, Baker L, Moss C, Sayialel S. 2003 Long-distance communication of acoustic cues to social identity in African elephants. *Anim. Behav.* **65**, 317–329. (doi:10.1006/anbe.2003.2047)

61. Rhodes JS, Cutler A, Moon KR. 2023 Geometry- and accuracy-preserving random forest proximities. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 10947–10959. (doi:10.1109/TPAMI.2023.3263774)

62. Lenth R. 2019 emmeans: estimated marginal means, aka least-squares means. R package version 1.8.4-1. https://CRAN.R-project.org/package=emmeans.

63. Pardo MA, Poole JH, Stoeger AS, Wrege PH, O'Connell-Rodwell CE, Padmalal UK, de Silva S. 2019 Differences in combinatorial calls among the 3 elephant species cannot be explained by phylogeny. *Behav. Ecol.* **30**, 809–820. (doi:10.1093/beheco/arz018)

64. Wright TF, Dahlin CR, Salinas-Melgoza A. 2008 Stability and change in vocal dialects of the yellow-naped amazon. *Anim. Behav.* **76**, 1017–1027. (doi:10.1016/j.anbehav.2008.03.025)

65. Okello JBA *et al*. 2008 Population genetic structure of savannah elephants in Kenya: conservation and management implications. *J. Hered.* **99**, 443–452. (doi:10.1093/jhered/esn028)

66. Viljoen JJ, Ganswindt A, Reynecke C, Stoeger AS, Langbauer WR. 2015 Vocal stress associated with a translocation of a family herd of African elephants (*Loxodonta africana*) in the Kruger National Park, South Africa. *Bioacoustics* **24**, 1–12. (doi:10.1080/09524622.2014.906320)

67. Roca IT *et al*. 2016 Shifting song frequencies in response to anthropogenic noise: a meta-analysis on birds and anurans. *Behav. Ecol.* **27**, 1269–1274. (doi:10.1093/beheco/arw060)

68. Parker JM, Wittemyer G. 2022 Orphaning stunts growth in wild African elephants. *Conserv. Physiol.* **10**, coac053. (doi:10.1093/conphys/coac053)

69. Goldenberg SZ *et al*. 2022 Social integration of translocated wildlife: a case study of rehabilitated and released elephant calves in Northern Kenya. *Mamm. Biol.* **102**, 1299–1314. (doi:10.1007/s42991-022-00285-9)

70. O'Connell-Rodwell CE, Wood JD, Kinzley C, Rodwell TC, Poole JH, Puria S. 2007 Wild African elephants (*Loxodonta africana*) discriminate between familiar and unfamiliar conspecific seismic alarm calls. *J. Acoust. Soc. Am.* **122**, 823–830. (doi:10.1121/1.2747161)

71. Boughman JW. 1998 Vocal learning by greater spear-nosed bats. Proc. R. Soc. B **265**, 227–233. (doi:10.1098/rspb.1998.0286)

72. Ruch H, Zürcher Y, Burkart JM. 2018 The function and mechanism of vocal accommodation in humans and other primates. *Biol. Rev. Camb. Philos. Soc.* **93**, 996–1013. (doi:10.1111/brv.12382)

73. Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE. 2015 Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* **25**, 495–499. (doi:10.1016/j.cub.2014.12.032)

74. Fischer J, Wheeler BC, Higham JP. 2015 Is there any evidence for vocal learning in chimpanzee food calls? *Curr. Biol.* **25**, R1028–R1029. (doi:10.1016/j.cub.2015.09.010)

75. Janik VM, Knörnschild M. 2021 Vocal production learning in mammals revisited. *Phil. Trans. R. Soc. B* **376**. (doi:10.1098/rstb.2020.0244)

76. Lameira AR, Delgado RA, Wich SA. 2010 Review of geographic variation in terrestrial mammalian acoustic signals: human speech variation in a comparative perspective. *J. Evol. Psychol.* **8**, 309–332. (doi:10.1556/JEP.8.2010.4.2)

77. Hedwig D, Poole J, Granli P. 2021 Does social complexity drive vocal complexity? Insights from the two African elephant species. *Animals* **11**, 1–21. (doi:10.3390/ani11113071)

78. Stoeger AS, Silva S. 2014 African and Asian elephant vocal communication: a cross-species comparison. In *Biocommunication of animals* (ed. G Witzany), pp. 233–247. Dordrecht, The Netherlands: Springer Netherlands. (doi:10.1007/978-94-007-7414-8)

79. Kershenbaum A, Ilany A, Blaustein L, Geffen E. 2012 Syntactic structure and geographical dialects in the songs of male rock hyraxes. Proc. R. Soc. B **279**, 2974–2981. (doi:10.1098/rspb.2012.0322)

80. Reyes-Arias JD *et al*. 2023 Vocalizations of wild West Indian manatee vary across subspecies and geographic location. *Sci. Rep.* **13**, 11028. (doi:10.1038/s41598-023-37882-8)

81. Pardo M, Lolchuragi DS, Poole J, Granli P, Moss C, Douglas-Hamilton I *et al*. 2024. Data from: Female African elephant rumbles differ between populations and sympatric social groups. Figshare. (doi:10.6084/m9.figshare.c.7456496)