Research

Author for correspondence:
Killian Martin
e-mail: killian.martin@ens-lyon.fr

THE ROYAL SOCIETY
PUBLISHING

Vocal complexity in a socially complex corvid: gradation, diversity and lack of common call repertoire in male rooks

Killian Martin[1], Francesca M. Cornero[2],
Nicola S. Clayton[2], Olivier Adam[3,4], Nicolas Obin[5] and
Valérie Dufour[1]

[1]PRC, UMR 7247, Ethologie Cognitive et Sociale, CNRS-IFCE-INRAE-Université de Tours, Strasbourg, France
[2]Department of Psychology, University of Cambridge, Cambridge, UK
[3]Institut Jean Le Rond d'Alembert, UMR 7190, CNRS-Sorbonne Université, 75005 Paris, France
[4]Institut des Neurosciences Paris-Saclay, UMR 9197, CNRS-Université Paris Sud, Orsay, France
[5]STMS Lab, IRCAM, CNRS-Sorbonne Université, Paris, France

KM, 0000-0002-8428-6012; FMC, 0000-0001-7825-4746;
NSC, 0000-0003-1835-423X; OA, 0000-0001-5192-982X;
NO, 0000-0002-5236-5306; VD, 0000-0003-1766-9598

Vocal communication is widespread in animals, with vocal repertoires of varying complexity. The social complexity hypothesis predicts that species may need high vocal complexity to deal with complex social organization (e.g. have a variety of different interindividual relations). We quantified the vocal complexity of two geographically distant captive colonies of rooks, a corvid species with complex social organization and cognitive performances, but understudied vocal abilities. We quantified the diversity and gradation of their repertoire, as well as the inter-individual similarity at the vocal unit level. We found that males produced call units with lower diversity and gradation than females, while song units did not differ between sexes. Surprisingly, while females produced highly similar call repertoires, even between colonies, each individual male produced almost completely different call repertoires from any other individual. These findings question the way male rooks communicate with their social partners. We suggest that each male may actively seek to remain vocally distinct, which could be an asset in their frequently changing social environment. We conclude that inter-individual similarity, an understudied aspect of vocal repertoires, should also be considered as a measure of vocal complexity.

1. Introduction

Living in a complex social system imposes communicative demands on individuals, ranging from the exchange of a wide variety of information to discriminating between conspecifics and heterospecifics [1,2]. The social complexity hypothesis for communicative complexity posits that there should be a relationship between social and communicative complexity [3,4]. The social intelligence hypothesis [5,6] also posits that social species, including corvids, have acquired advanced cognitive skills to meet the demands of their sociality. Following both hypotheses, we can predict that social species with advanced cognitive aptitudes should also exhibit vocal complexity. Vocal complexity can be addressed by looking at several features of the repertoire of a species, including its diversity (e.g. the number of vocal units in the repertoire) and its flexibility [7]. Vocal flexibility can be either structural, such as with gradation, where vocalizations of different types exist at either end of a continuum, and vocalizations gradually shift in structure from one type to the other [7–9], or functional, such as when an individual produces acoustically identical vocalizations in several different contexts [10,11]. The notion of vocal flexibility is an intriguing concept. On the one hand, vocal flexibility should provide communicative advantages, as it potentially allows the transmission of more complex signals, with more nuances and expressivity [7]. On the other hand, vocalizing has the main communicative function of conveying a message to conspecifics. This is in general true of the single calls in the repertoire (hereafter referred to as call units), which serve functions as diverse as warning about predators, signalling food location, making social interactions or coordinating social activities [1,2]. To remain intelligible to listeners, call units should remain somewhat standardized in both acoustic and functional terms.

Aside from call units, the potential for vocal flexibility also exists in the songs of adult birds. Songs are composed of vocal units (hereafter referred to as song units), which can be reused call units or entirely different vocalizations. In songs, different structural arrangements of song units appear functionally equivalent, for example serving reproductive or territorial functions [12,13]. This may allow greater freedom for different individuals to flexibly produce different songs, either through variable song unit choice, or through variable sequential organization. However, the vast majority of bird species sing the same stereotyped song throughout their life after acquiring their songs as juveniles [14]. In extreme cases, an individual may produce only a single stereotyped song, as in the zebra finch *Taeniopygia guttata* [15,16] or the white-crowned sparrow *Zonotrichia leucophrys* [17]. Only comparatively few species, called open-ended learners, appear able to modify their song past the juvenile acquisition phase [18]. Vocal flexibility may thus lead to varying degrees of inter-individual similarity.

Aside from the diversity and flexibility of the repertoire, other inter-individual vocal variations may be markers of social complexity. One example is the phenomenon called 'vocal signature' [19–21], which does not impact function, but can serve to identify individuals (e.g. in birds [22–27]; primates [28–32]; mongoose [33]; cetaceans [34–36]), or can signal social group membership [36,37]. Another example are dialects, variations between geographically separate populations [12,38,39]. Dialects may be a by-product of vocal learning, as many species acquire their vocalizations by learning from social partners [12] such as parents and affiliated individuals. Given their communicative function, the vocalizations of different individuals or groups are not expected to greatly deviate from what is typical in the species' repertoire in terms of their acoustic structure, as it may otherwise hinder transmission of information to receivers of the vocalization.

Birds, with their diverse aggregation patterns, are ideal species to study the link between social systems and communicative complexity. The corvid family, in particular, includes some of the most socially complex species [40]. For instance, rooks (*Corvus frugilegus*) are pair-bonded birds that live in social groups throughout the year; multiple social groups merge into large colonies for breeding in spring and communal roosting in autumn but stay separate at other times of the year [40–44]. The social organization of rooks is thus highly complex, as individuals are involved in multiple layers of social connections in a frequently changing social environment where they may need to signal not only their own identity, but also what pair, social group, or even colony they belong to. We may therefore expect their repertoire to exhibit very high levels of diversity or flexibility that allow this identity signalling. At the group level, rooks possess a medium to large repertoire of calls [45,46], with evidence of an individual vocal signature in the most common 'caw' call unit produced by males [23]. Like several other corvids, they can also mimic other sounds, including anecdotal evidence of human voice mimicry [41,43,47] (N.S.C. and V.D., personal observations). Rooks also show long-term vocal recognition of conspecifics, have excellent learning, memory and planning skills, and more generally show good socio-cognitive skills [40,48,49]. They also appear to have the same song-related

neural circuits as other oscines [50]. Aside from their calls, adult rooks, and some other corvids, also produce soft 'songs' [41,47,51]. These 'songs' are sequences of song units organized into phrases and can last from a few seconds to several minutes. However, unlike most other cases of birdsong, corvid 'songs' appear unrelated to either territorial defence or courtship [12,47], but instead resemble the undirected songs exhibited by several other songbird species [52–55]. This vocally atypical and socially complex species may therefore be a good model to evaluate vocal complexity in relation to cognitive and social complexity.

Our goal was to evaluate the complexity of vocal production in rooks. To do so, we recorded the vocalizations of two distant captive colonies, in Strasbourg (France) and Cambridge (UK). In contrast with previous studies, which rely only on standard acoustic parameters [23,56–59], we chose to construct each individual vocal repertoire through validated state-of-the-art machine-learning approaches [60,61]. We hypothesized that the high social complexity found in rooks should be associated with high vocal complexity, which should yield high diversity (large number of vocal units) and/or high gradation. Differences may be found between males and females, or between call units and song units. If vocal complexity is closely connected to the social relationship and structure existing in this species, we should find higher vocal similarity in call units between closely related individuals (pairs/colonies) compared to more distant individuals. In addition, if each individual rook produces complex and flexible songs as suggested in the literature [41], we should find low song unit similarity between individuals.

2. Material and methods

2.1. Study groups

The rooks in this study belonged to two captive colonies, one in Strasbourg (France) and one in Cambridge (UK). The individuals of each colony had been housed together since their capture as part of long-term research projects, except for one male in the Strasbourg colony. All birds except for this male had been caught as wild pre-flight juveniles and were hand-raised (see electronic supplementary material, table S1, for details, and table S2 for a tentative comparison with wild conditions). At the time of the study, 15 individuals (7 females, 8 males) comprised the Strasbourg colony and 7 individuals (4 females and 3 males) comprised the Cambridge colony. DNA testing had been used to ascertain the sex of each bird after their original capture. Finally, all birds were given coloured and/or numbered leg rings for identification purposes. Both colonies had food, water and enrichment available ad libitum, with food and water renewed daily.

In addition to the vocal complexity analysis in this study, we also analysed the social structure of the Strasbourg colony through proximity scans, to ascertain whether there were differences between certain individuals in social complexity (e.g. the total number and diversity of relations they had with other individuals). The results of this social analysis are shown in electronic supplementary material, figure S5.

2.2. Data collection

Both colonies were recorded and observed in several sessions with the same protocol. Autonomous recorders (Song Meter 4, Wildlife Acoustics, USA) recorded several hours each day. Each recorder bore two microphones on 3 m cables, and were spaced throughout the aviary to minimize distance from vocalizing birds (approximate maximum distance of 10 m). The dataset was thus composed of multichannel audio (2, 4 or 6 channels), digitized at 48 kHz with 16-bit resolution. Recordings from different recorders were manually synchronized before annotation.

Several recording sessions were attended by a human observer, throughout 2020 and 2021 for the Strasbourg colony and between February and April 2022 for the Cambridge colony. Vocalizations were noted in real time using a custom Python script, along with the identity of the emitter. Behavioural contexts for the vocalizations were determined by observation (based on [41]) and included when possible. Strasbourg sessions were taped using a video camera to help with the higher number of individuals. An expert observer then used these notes to annotate the recordings as spectrograms with the Audacity software (v3.1.3) [62]. We defined vocal units as vocalizations corresponding to continuous traces on the spectrogram, such as separating every element of a song sequence as song units, or as sequences of extremely short elements that only occurred together. We also made a distinction between two types of vocal units: songs (defined as series of at least 5 vocal units of at least

2 perceptually different types produced by one individual, with less than 10 s between successive vocal units), and calls (all other vocalizations).

This manually annotated dataset was used to train a deep neural network. This neural network is described in a previous paper [63]. In brief, this network replicates the manual annotation process (extracting start and end timestamps and individual identity for each vocal unit) using Mel-scale spectrograms corresponding to short chunks of audio. The network trained on the above dataset achieved approximately 72% retrieval of vocal units and 85% accuracy in identifying emitters in the retrieved vocal units. The same expert observer who annotated the previous dataset validated the network results by comparing them with the full audio recordings to ensure the same quality as the manual annotations, such that no false positive detections, false negative misses, or misidentifications remained when incorporating these results for the analysis.

The final dataset included 31 148 vocalizations from 47.9 h of audio recordings (29.1 h from the Strasbourg colony, 18.8 h from the Cambridge colony), of which 3 h from each colony came from unattended recording sessions.

2.3. Audio representation

The vocal units were extracted from the sound recordings and denoised with a dynamic spectral gating algorithm [61], then the microphone with the highest signal-to-noise ratio was selected. The audio was pre-emphasized and filtered, then converted to a spectrogram using a 10 ms Hamming window with 80% overlap. The spectrogram was then Mel-scaled to 80 frequency coefficients and constrained to above 100 Hz as a final noise-removing step. The spectrogram was trimmed to remove silent beginnings and ends, log-scaled, thresholded to a dynamic range of 20 dB below its maximum value, and finally normalized to between 0 and 1. This normalization removed most background noise and controlled for differences in amplitude and recording conditions between colonies and between different days [64]. Finally, only vocalizations from identified individuals that did not overlap other sounds were used in the analysis, keeping 25 220 vocalizations for further analysis.

2.4. Acoustic analysis

In this study, we sought to quantify and visualize the vocal complexity of rooks, which required the construction of a catalogue of the different vocal units produced by the species. We followed the procedure outlined here: define a measure of the acoustic distance between pairs of vocal units, compute the pairwise acoustic distances between the vocal units in the dataset, apply a dimensionality reduction algorithm followed by clustering to define groups of similar vocal unit, and finally use the clustering results to compute vocal complexity measures. Each step of this procedure is further detailed below and in the electronic supplementary material.

We considered vocal units to be similar if their spectrograms could be aligned in time and frequency. We selected a dynamic frequency–time warping (DFTW) distance to compute this alignment [65], although we used a constrained variant of the DFTW algorithm to reduce computational complexity (see electronic supplementary material). DFTW measures acoustic similarity between vocal units even if, for instance, their lengths or their frequency distribution differ. DFTW finds the optimal alignment between spectrograms by minimizing a cost function corresponding to the distance between the aligned spectrograms, and the value associated with this optimal alignment was used as the acoustic distance in our analysis.

The DFTW distance matrix, obtained from the acoustic distances between all pairs of vocal units in the data, contains the relative positions of the vocal units in the dataset with respect to one another. We used the UMAP dimensionality reduction algorithm [66] to project these positions to absolute coordinates in two-dimensional space for visualization. The projection was then used in clustering using the HDBSCAN algorithm [67,68] to group similar vocalizations together. HDBSCAN has the particularity of allowing soft clustering, where each vocal unit is not assigned to a single cluster, but instead to all clusters with varying probabilities. Both UMAP and HDBSCAN have become state-of-the-art algorithms due to their performance and robustness [69,70].

This analysis was carried out in Python (v3.8) using custom code for the DFTW, and the *umap-learn* (v0.5.3) [66] and *hdbscan* (v0.8.28) [67] packages for UMAP and HDBSCAN, respectively. Parameters were left at their default values in these packages, except for the following: UMAP *n_neighbors* = 30, *min_dist* = 0; HDBSCAN *min_samples* = 30, *min_cluster_size* = 30. UMAP parameters were validated by checking that the UMAP projection preserved the structure of the dataset with the trustworthiness and continuity

measures [71]. Trustworthiness was 0.96 and continuity was 0.92, where 1 represents perfect preservation, so we proceeded with the analysis. HDBSCAN parameters were chosen to avoid small clusters that may emerge purely by chance.

2.5. Quantifying repertoire complexity

We used the soft clustering results to quantify three aspects of complexity: repertoire diversity (the number of different vocal units in the repertoire of an individual), repertoire gradation (a measure of separation between categories of vocal units; gradation is low if categories are well-separated with no intermediate vocal units), and inter-individual repertoire similarity (the extent to which two individuals share their repertoires).

We quantified diversity and gradation using two information theory indices, Shannon's diversity index and Simpson's diversity index [72] (see also electronic supplementary material), which represent the effective number of same-size clusters in the results from HDBSCAN. Shannon's diversity weighs all clusters equally, while Simpson's diversity gives more importance to large clusters. These indices can take values between 1 and the total number of clusters. We normalized the values to between 0 and 1 to make them comparable with future results, so that a vocal unit attributed to only one cluster corresponded to a value of 0, and a vocal unit attributed to all clusters with equal probability corresponded to a value of 1. For repertoire gradation, we computed one index value per vocal unit. For repertoire diversity, we computed one index value per individual by averaging the soft clustering results of all vocal units produced by this individual. We note that this may cause a strong correlation between the two indices, since we compute them from essentially the same data. In practice, this stresses the importance of using not just one or the other, but both indices, to avoid ambiguities (e.g. mistaking a highly graded repertoire for a highly diverse repertoire) by looking at the data through multiple perspectives.

For inter-individual similarity, we used the Morisita–Horn overlap index [73] (see also electronic supplementary material), which represents the probability that two vocal units, one each randomly drawn from the repertoire of different individuals, fall in the same cluster. The index is bounded between 0 and 1, with 0 representing complete dissimilarity (no common clusters between the individuals) and 1 representing complete similarity (both individuals produce vocalizations from the same clusters with the same distribution). We computed the index for each pair of individuals by averaging the soft clustering results of all vocal units produced by each individual.

2.6. Statistical analysis

Statistical analysis of the distribution of gradation, diversity and similarity was carried out using generalized linear models (GLMs) or generalized linear mixed models (GLMMs). Models were fitted in R (v4.1.1) [74] with the *glmmTMB* package (v1.1.7) [75]. We chose ordered beta regression as the response variables were the indices, represented as continuous variables bounded between 0 and 1 [76]. We built full models with sex, vocal unit type and colony as predictors, as well as all second-order interactions. For diversity and gradation, individual identity was also included as a random factor. For inter-individual similarity, the sex and vocal unit type included comparisons within and between levels (e.g. for sex, we compared males with other males, females with other females and females with males); the colony variable only compared within the same colony or between different colony (we did not distinguish between the Strasbourg and Cambridge colonies due to the lower number of singing birds in the latter). We then performed model selection with the *MuMIn* package (v1.43.17) [77], fitting all models obtainable by removing variables from the original full model (including the null model by removing all variables), and ranking them by increasing AICc. In this procedure, the best model has the lowest AICc, and any model more than 2 AICc units above this model is significantly worse at predicting the data [78]. The quality of the selected best model was then assessed visually through QQ plots of the model residuals against quantiles of a normal distribution with mean 0 and variance 1. We finally performed pairwise tests on the best model with the *emmeans* package (v1.8.0) [79], and Tukey's method was used to correct *p*-values for multiple testing.

3. Results

3.1. Dataset distribution

Vocalization rates (vocal units produced per individual per hour recorded) were lower for females than males in both colonies, and were also lower for the Cambridge colony in general compared to the
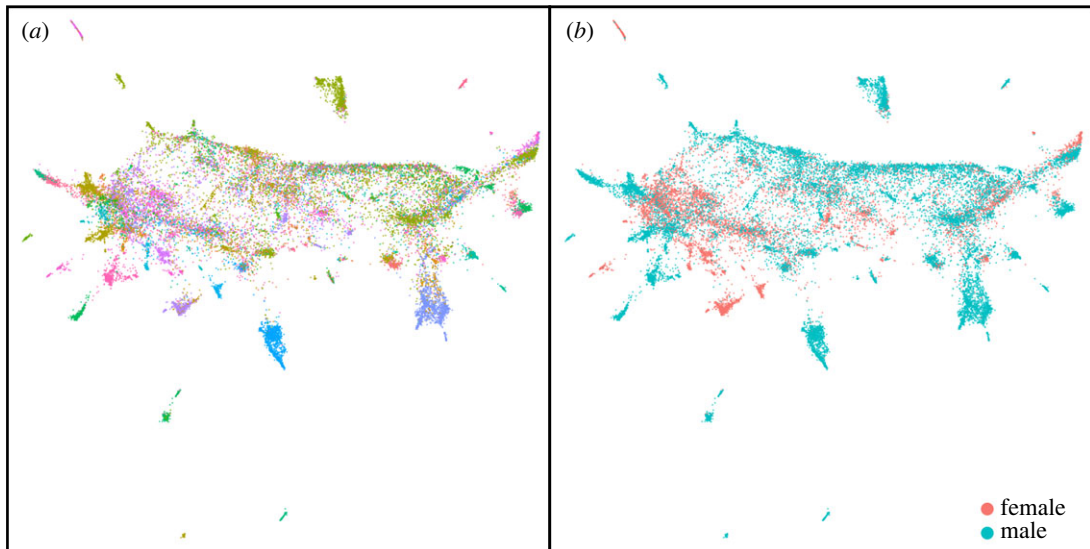
Figure 1. UMAP embeddings of the vocal units in the data, with different colour schemes: (*a*) individual identity (legend omitted to reduce clutter) and (*b*) sex. See electronic supplementary material, figure S2, for projections also coloured by colony and vocal unit type.

Strasbourg colony (electronic supplementary material, table S3). This meant that some females were sampled far less (in particular, one female from the Cambridge colony accounted for less than 200 vocal units). The dataset was split approximately equally between call (45.9%) and song (54.1%) units. Both males and females produced song.

Contexts for the vocal units were determined from observing the behaviour of the emitter at the time of production. Only 33% of vocal units in the dataset could be contextualized in this manner, with equal repartition between males (52% of contextualized vocal units) and females (48%). The specific contexts associated with these vocal units differed between sexes. Contextualized vocal units from females were most often associated with a bow and tail fanning posture (27% of these vocal units), nest building (26%), receiving food (20%) and nest calling (10%). Contextualized vocal units from males were most often associated with a bow and tail fanning posture (58% of these vocal units), play or object exploration (21%) and foraging (11%). Furthermore, 53% of contextualized vocal units correspond to song, with 85% produced by males. Among these song units, 9% were also associated with a bow and tail fanning posture, 7% with play or object exploration, 1% with a copulatory display and under 0.1% with mounts, foraging and feeding another bird.

3.2. Projection and clustering results

All vocal units were projected into two-dimensional space with UMAP, showing several well-defined clusters around a more diffuse region (figure 1). Many clusters appeared homogeneous, including only one individual and either call units or song units. HDBSCAN found 80 clusters, accounting for 68.3% of the vocal units. The remaining 31.7% of vocal units were too sparse to assign to a cluster. To avoid dropping them from the analysis, we used the soft clustering results, where each vocal unit is assigned to all clusters with varying probabilities.

3.3. Repertoire gradation

We quantified repertoire gradation with Shannon's diversity and Simpson's diversity. Both indices resulted in similar statistical results; we therefore solely discuss Shannon's diversity here (see electronic supplementary material, table S4 and figure S3, for the same analysis with Simpson's diversity). The best model included only sex, vocal unit type and the interaction between them. The null model was significantly worse at predicting repertoire gradation (425.4 AICc units above the best model).

Pairwise tests (figure 2*a*) showed that male call units had significantly lower gradation than either male song units (odds ratio, call units to song units: 0.70 ± 0.01, $z = -20.30$, $p < 0.0001$) or female call units (odds ratio, females to males: 1.87 ± 0.20, $z = 5.90$, $p < 0.0001$). By contrast, female call units
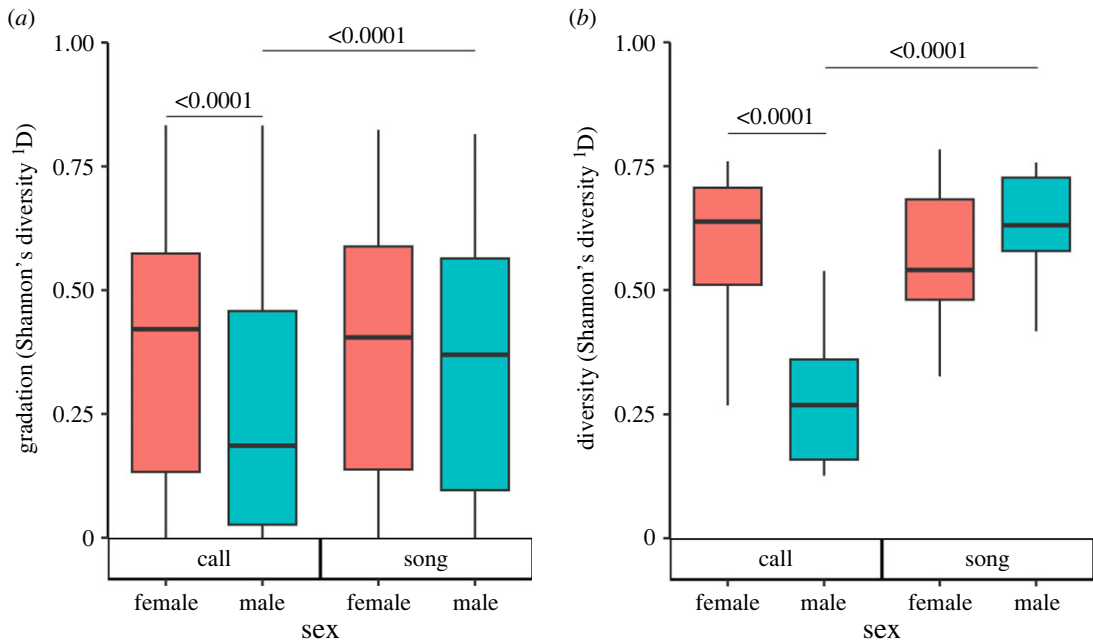
Figure 2. Analysis of (*a*) repertoire gradation and (*b*) repertoire diversity by sex and vocal unit type, measured by Shannon's diversity. The two colonies were not statistically different and so were pooled together. Horizontal lines represent pairwise tests with the associated *p*-value above. For clarity, only tests where one variable varies are shown (e.g. comparing call units between males and females, or comparing between call units and song units in males, but not between male call units and female song units).

were not less graded than female song units (odds ratio, call units to song units: 1.08 ± 0.03, $z = 2.40$, $p = 0.08$), and male song units were not less graded than female song units (odds ratio, female to male: 1.20 ± 0.13, $z = 1.72$, $p = 0.31$).

3.4. Repertoire diversity

Like with repertoire gradation, similar statistical results for repertoire diversity were obtained with both Shannon's diversity and Simpson's diversity indices; we therefore only discuss Shannon's diversity here (see electronic supplementary material, table S4 and figure S3, for the analysis using Simpson's diversity). The best model included only sex, vocal unit type and the interaction between them. The null model was significantly worse at predicting repertoire diversity (25.6 AICc units above the best model).

Pairwise tests (figure 2*b*) showed that male call units were significantly less diverse than either male song units (odds ratio, call units to song units: 0.23 ± 0.04, $z = -7.81$, $p < 0.0001$) or female call units (odds ratio, females to males: 4.01 ± 1.01, $z = 5.51$, $p < 0.0001$). By contrast, male song units were no more diverse than female song units (odds ratio, females to males: 0.82 ± 0.24, $z = -0.67$, $p = 0.91$), and female call units were no more diverse than female song units (odds ratio, call units to song units: 1.12 ± 0.24, $z = 0.57$, $p = 0.94$).

3.5. Inter-individual repertoire similarity

We assessed inter-individual similarity with the Morisita–Horn overlap index, which quantifies to what extent the vocalizations of two individuals fall in the same clusters (figure 3). The best model was the full model (electronic supplementary material, table S4). The null model was significantly worse at predicting inter-individual repertoire similarity (296.4 AICc units above the best model).

The most striking result concerned the extremely low inter-individual similarities in the call units of males (figure 4), even within the same colony. Male call unit repertoire similarity was significantly lower than either female call unit similarity (odds ratio, females to males within the same colony: 8.04 ± 1.39, $z = 12.03$, $p < 0.0001$) or male song unit similarity (odds ratio, call units to song units within the same colony: 0.11 ± 0.02, $z = -11.89$, $p < 0.0001$). By contrast, females were far more similar
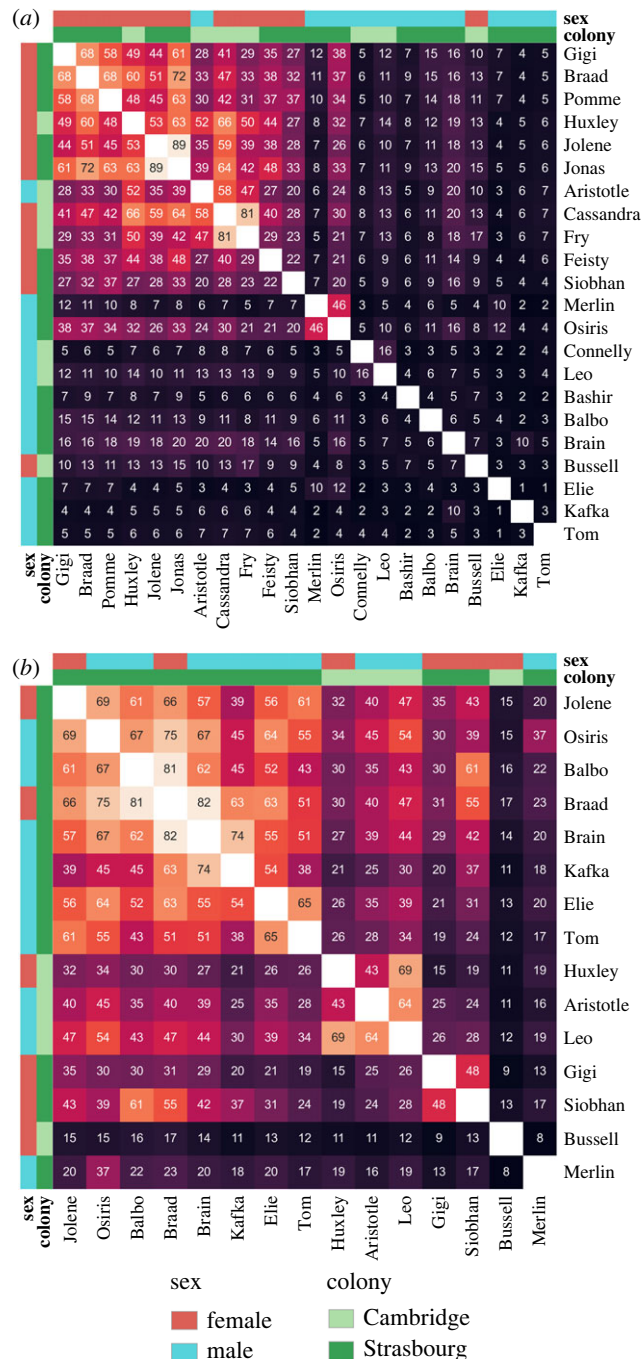
Figure 3. Pairwise inter-individual repertoire similarity, measured by the Morisita–Horn overlap index. Each heatmap corresponds to a similarity matrix, where each row and column represents one individual. (a) For call units. (b) For song units. Each matrix was reordered through hierarchical clustering so individuals were grouped by their pattern of similarity with respect to other individuals. Light cells indicate high similarity, with the associated index value inside the cell in percent.

in their call unit repertoires, even between different colonies (odds ratio, female to male between different colonies: 4.82 ± 0.85, $z = 8.87$, $p < 0.0001$). This was a very consistent pattern, with very few exceptions (figure 3a). For instance, two Strasbourg males were more similar to each other (Merlin and Osiris, 46% overlap index), and one Cambridge female was highly dissimilar to all other individuals in either colony (Bussell, 317% overlap index). Moreover, while many individuals in both colonies were in breeding pairs (electronic supplementary material, table S1), there was low call unit similarity between males and females in general (figure 3a, figure 4 leftmost panel), which precluded similarity between the members of a pair. To illustrate these results, examples of the most frequently uttered call units from every individual can be found in the electronic supplementary material (see also electronic

Figure 4. Analysis of inter-individual repertoire similarity by sex, vocal unit type, and colony, measured by the Morisita–Horn overlap index. Boxplots indicate distributions of similarity values. Horizontal lines represent pairwise tests with the associated *p*-value above. Analogous comparisons within and between colonies, when they are associated with the same result (e.g. top left, second comparison from the top: the comparison of male call units versus female call units are significant both within the same colony and between colonies, so they are represented on the same line, with the *p*-value associated with the 'within colony' test in red and the *p*-value associated with the 'between colony' test in blue). For clarity, only tests where one variable varies are shown and grouped by which variable is being compared (for instance, the bottom three comparisons are 'by colony', i.e. comparisons between individuals 'within' the same colony, or between individuals 'between' different colonies, but involving the same vocal units, e.g. female song units for the leftmost of these comparisons).

supplementary material, figure S6). Unlike call units, song unit repertoire similarity did not differ between sexes (odds ratio, females to males within the same colony: 0.80 ± 0.18, $z = -0.95$, $p = 1$; females to males between different colonies: 0.48 ± 0.11, $z = -3.07$, $p = 0.17$). Most individuals had intermediate to high song unit repertoire similarity, even between colonies (figure 3*b*).

Comparing between song units and call units (i.e. whether the same vocal units may be used both during song and as call units), the similarity matrix (electronic supplementary material, figure S4) suggested that males, despite higher song unit repertoire similarity, did not incorporate call units in their song unit repertoire. Females were significantly more similar to one another in this respect (odds ratio, females to males within the same colony: 3.21 ± 0.47, $z = 7.90$, $p < 0.001$; between different colonies: 1.92 ± 0.31, $z = 4.07$, $p = 0.006$). These results excluded self-similarities (i.e. between the call and song units of the same individual); nonetheless, self-similarities showed no apparent pattern: some individuals were highly self-similar (e.g. Aristotle 76%, Bussell 98%, Braad 85%, Merlin 93%), others much less so (e.g. Tom 6%, Kafka 17%), with no effect of sex or colony.

Despite being included in the best model, few pairwise tests on the effect of the colony were significant. Only three tests showed significantly higher similarity within the same colony (figure 4, black *p*-values): in female song units (odds ratio, within the same colony to between different colonies: 2.69 ± 0.50, $z = 5.37$, $p < 0.0001$), between female call units and female song units (odds ratio, within to between: 1.72 ± 0.24, $z = 3.88$, $p = 0.012$), and between male song units and female song units (odds ratio, within to between; 1.93 ± 0.29, $z = 4.37$, $p = 0.002$). Otherwise, there was no difference in similarity when comparing individuals from the same colony and individuals from different colonies.

4. Discussion

We evaluated the vocal complexity of the calls and songs of two colonies of rooks, a species with high social complexity and cognitive ability, through the analysis of the diversity, gradation and inter-individual similarity in their vocal repertoire. Male call units had the lowest gradation and diversity, while female song units, female call units, and male song units were not significantly different in either gradation or diversity. Thus, we detected vocal complexity but not necessarily where we had hypothesized. Indeed, we expected vocal complexity to mirror social complexity, through call or song similarity at the pair and colony levels. In the majority of vocal species, a vocal repertoire is composed in majority by clusters reflecting different call types, with some degree of gradation potentially present, but we found large degrees of gradation that precluded segregation into call types, especially for female call units. Finally, the vast majority of call units in this study could not be conclusively associated with specific behaviour contexts, which is consistent with other literature on corvid vocal communication [45–47]. However, one particular call type could be conclusively associated with a behavioural context: the nest call, produced only by brooding females. This association between acoustic structure and behavioural context was confirmed in our analysis, as almost all nest calls were projected close to other nest calls in the UMAP projection (see electronic supplementary material, figure S7). We note that this study is the first to quantify vocal complexity in corvids, although a fairly similar study has been conducted in macaques [80]. Nevertheless, these results show high vocal complexity via diversity, gradation and functional uncertainty in the repertoire [7], which still may reflect some elements of social complexity in the life of rooks.

Our most striking result, and directly contrary to our hypotheses, is the observation that call unit repertoires of individual males were all almost completely dissimilar even within the same colony, unlike the highly similar call unit repertoires shared by females even between different colonies. Interestingly, most of the female calls that could be contextualized were either food-related or associated with a breeding context (e.g. nest calls emitted by brooding females). These calls may be more biologically significant in females, who depend entirely on their mates feeding them during much of the breeding season [81,82]. As such, nest calls may be under stronger pressure for inter-individual homogeneity within and between individuals and groups. These essential communicative functions may explain why the call units of the females are more similar than those of the males. Unlike females, males in both colonies exhibited highly individualized call unit repertoires. Our study is the first to quantify this individual heterogeneity using objective measurements, expanding on the previous literature on the repertoire of rooks [23,45] and other corvids such as American crows [22], and Hawaiian crows [83], where this heterogeneity was suspected but could not be quantified. In rooks, the most frequently uttered 'caw' call unit has been reported to vary greatly between individuals, especially in its duration [23,45], while two other call units, the 'gull' and 'squalling' calls, varied in frequency modulation [45]. However, previous repertoires grouped calls into functional categories rather than by acoustic structure. In most bird species, individuals produce similar call repertoires [2], which arise through vocal learning during development [81,84,85] or through vocal convergence between social partners [32,86–88]. Inter-individual heterogeneity has been reported in very few cases, such as in the contact calls of parrots [25,89], the signature whistles of bottlenose dolphins [35], or the food begging calls of juvenile chipping sparrows [90]. However, it is not always clear if the reported differences are due to heterogeneity as we observe in this study or to individual vocal signatures [21], and none of these previous results were at the scale of the entire call repertoire. The large acoustic heterogeneity detected in male rooks suggests that they are weakly influenced by vocal productions of others, and thus reduces the probability of detecting group signatures or dialects in their calls. This heterogeneity, and the lower diversity found in male call units, could not be linked to lower social complexity in males compared to females (see electronic supplementary material, figure S5). While analogous data are currently lacking in wild rooks, personal observations and the existence of the same patterns in two independent captive colonies mean that our results are unlikely to be artefacts of either captivity or atypical vocal behaviour. In particular, rooks are capable of vocal mimicry of others and environmental noises [41,43,47], thus inter-male heterogeneity could have been explained by mimicry of and in different environments. However, we found no evidence that individuals from the same colony at the time of the study, or individuals captured from the same colony (see electronic supplementary material, table S1), were more similar, as could have arisen from social exposure to conspecifics, or common mimicry of environmental sounds at the respective locations of the colonies. In addition, female call units were remarkably similar between colonies despite the geographical distance between the colonies. The high inter-individual heterogeneity in

male rook calls raises the question of how they communicate within the group. Communication *a priori* requires common vocalizations for adequate information exchange, which was not the case in males here. One alternative is that rooks may memorize the calls used by each male in a given context. Corvids are among the only nonhuman species to have demonstrated episodic-like memory, the ability to recall not only what an event was, but also where and when it occurred [91–93], and so might have the ability to memorize the repertoires of other individuals.

Unlike call units, song unit repertoires were more similar between individuals, and might thus be better indicators of colonial origin. However, we again found no support for this hypothesis, as there was no clear dissimilarity between colonies. Despite this, rook song appears atypical compared to other birdsong. Notably, both sexes sang with similar degrees of unit diversity and gradation, indicating similar song complexity, and in circumstances unrelated to territorial defence or courtship. In fact, rooks often actively seek to perch alone to sing, and sing very softly compared to their calls [41]. For instance, in our study, a singer was often only heard on the microphones closest to its position in the aviary, but not on microphones only a few metres away. In general, corvid song has rarely been studied in the literature beyond passing mentions [41,47,51], and female birdsong has historically been understudied compared to male song [94]. What does exist questions the function of rook songs compared to what is most often described in the literature [1]. One possibility is that rooks may actively seek vocal individuality in order to remain distinct in frequently changing social environments, such as when social groups merge into larger colonies [41]. Vocal avoidance could then promote vocal innovation, itself encouraged by regular song practice, and in turn explain both high call heterogeneity and high song gradation and diversity. Nevertheless, this unusual manner and context for singing is an integral part of the vocal repertoire, and thus contributes to the vocal complexity of this species. Further studies are needed to evaluate individual innovation by looking at individual patterns or rhythm and their stability in the songs of this species.

High inter-individual heterogeneity may exist in other species, but may have remained undetected due to methodological limitations. The computational capability to compare large datasets of complete representations of vocal units, such as entire spectrograms, in a reasonable timeframe has emerged only recently [60,61,65,95], and previous studies usually used either human classification or specific acoustic parameters, such as spectral or temporal energy, fundamental frequency or frequency contour trajectory [61]. The methodology used here is based on this principle of complete representation. As such, inter-individual differences can be highlighted, without many of the constraints in acoustic parameter-based approaches. For instance, individual vocal signatures can be controlled for if they are based only on frequency or time differences (e.g. pitch or duration; see electronic supplementary material, figure S7). This approach allows fine-grained analysis of vocal units for studying vocal complexity in any animal species, representing a highly effective tool, in both objectivity (by largely removing human-based decisions in choice of parameters or classification of vocal units, whether functional or structural) and scalability (by limiting the time needed to examine a corpus of data to classify vocal units). We have demonstrated its use to measure not only repertoire diversity and gradation but also inter-individual similarity, which we suggest should also be included as a component of vocal complexity.

To our knowledge, this study is the first to focus on constructing the vocal repertoire of a corvid species based on individually identified birds. Moreover, our findings are all the more robust as they are based on two independent colonies that had never been in contact. Further studies are now needed to determine the main factors (ecological, cognitive, or social) that promote this vocal complexity. Approaches such as the one used here prove powerful and robust, and may be applied to other bird species with high vocal complexity, such as parrots [10,86], the nightingale *Luscinia megarhynchos* [96], the brown thrasher *Toxostoma rufum* [97] the lyrebird *Menura alberti* [98] or the mynah *Gracula religiosa* [99]. More generally, contrasting the vocal repertoire and flexibility of species with varying social organization complexity and cognitive performances should provide insights into the determinants of vocal complexity and its links with social complexity in various animal species.

References

1. Chen Z, Wiens JJ. 2020 The origins of acoustic communication in vertebrates. *Nat. Commun.* **11**, 369. (doi:10.1038/s41467-020-14356-3)

2. Marler P. 2004 Bird calls: a cornucopia for communication. In *Nature's music: the science of birdsong*, pp. 132–177. San Diego, CA: Academic Press. (doi:10.1016/B978-012473070-0/50008-6)

3. Freeberg TM, Dunbar RI, Ord TJ. 2012 Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B* **367**, 1785–1801. (doi:10.1098/rstb.2011.0213)

4. Peckre L, Kappeler PM, Fichtel C. 2019 Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behav. Ecol. Sociobiol.* **73**, 11. (doi:10.1007/s00265-018-2605-4)

5. Humphrey NK. 1976 The social function of intellect. In *Growing points in ethology* (eds P Bateson, R Hinde), pp. 303–317. Cambridge, UK: Cambridge University Press.

6. Dunbar RI. 1998 The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190. (doi:10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8)

7. Rebout N, Lone J-C, De Marco A, Cozzolino R, Lemasson A, Thierry B. 2021 Measuring complexity in organisms and organizations. *R. Soc. Open Sci.* **8**, 200895. (doi:10.1098/rsos.200895)

8. Wadewitz P, Hammerschmidt K, Battaglia D, Witt A, Wolf F, Fischer J. 2015 Characterizing vocal repertoires—hard vs. soft classification approaches. *PLoS ONE* **10**, e0125785. (doi:10.1371/journal.pone.0125785)

9. Fischer J, Wadewitz P, Hammerschmidt K. 2017 Structural variability and communicative complexity in acoustic communication. *Anim. Behav.* **134**, 229–237. (doi:10.1016/j.anbehav.2016.06.012)

10. Montes-Medina AC, Salinas-Melgoza A, Renton K. 2016 Contextual flexibility in the vocal repertoire of an Amazon parrot. *Front. Zool.* **13**, 40. (doi:10.1186/s12983-016-0169-6)

11. Clay Z, Archbold J, Zuberbuhler K. 2015 Functional flexibility in wild bonobo vocal behaviour. *PeerJ* **3**, e1124. (doi:10.7717/peerj.1124)

12. Catchpole CK, Slater PJ. 2008 *Bird song: biological themes and variations*, 2nd edn. Cambridge, UK: Cambridge University Press.

13. Nelson CMV, Sherwin WB, Ord TJ. 2022 Why does the complexity of functionally equivalent signals vary across closely related species? *Behav. Ecol.* **33**, 926–936. (doi:10.1093/beheco/arac050)

14. Beecher MD, Brenowitz EA. 2005 Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* **20**, 143–149. (doi:10.1016/j.tree.2005.01.004)

15. Templeton CN, Laland KN, Boogert NJ. 2014 Does song complexity correlate with problem-solving performance inflocks of zebra finches? *Anim. Behav.* **92**, 63–71. (doi:10.1016/j.anbehav.2014.03.019)

16. Boogert NJ *et al.* 2011 Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Anim. Behav.* **81**, 1209–1216. (doi:10.1016/j.anbehav.2011.03.004)

17. Marler P, Tamura M. 1962 Song 'dialects' in three populations of white-crowned sparrows. *Condor* **64**, 368–377. (doi:10.2307/1365545)

18. Brenowitz EA, Margoliash D, Nordeen KW. 1997 An introduction to birdsong and the avian song system. *J. Neurobiol.* **33**, 495–500. (doi:10.1002/(SICI)1097-4695(19971105)33:5<495::AID-NEU1>3.0.CO;2-#)

19. Bradbury JW, Vehrencamp SL. 1998 *Principles of animal communication*, 2nd edn. Sunderland, MA: Sinauer Associates, Inc.

20. Beecher MD. 1989 Signalling systems for individual recognition: an information theory approach. *Anim. Behav.* **38**, 248–261. (doi:10.1016/S0003-3472(89)80087-9)

21. Linhart P, Osiejuk TS, Budka M, Šálek M, Špinka M, Polícht R, Syrová M, Blumstein DT. 2019 Measuring individual identity information in animal signals: overview and performance of available identity metrics. *Methods Ecol. Evol.* **10**, 1558–1570. (doi:10.1111/2041-210X.13238)

22. Mates EA, Tarter RR, Ha JC, Clark AB, Mcgowan KJ. 2015 Acoustic profiling in a complexly social species, the American crow: caws encode information on caller sex, identity and behavioural context. *Bioacoustics* **24**, 63–80. (doi:10.1080/09524622.2014.933446)

23. Benti B, Curé C, Dufour V. 2019 Individual signature in the most common and context-independent call of the rook (Corvus frugilegus). *Wilson J. Ornithol.* **131**, 373. (doi:10.1676/18-41)

24. Elie JE, Theunissen FE. 2018 Zebra finches identify individuals using vocal signatures unique to each call type. *Nat. Commun.* **9**, 4026. (doi:10.1038/s41467-018-06394-9)

25. Berg KS *et al.* 2012 Vertical transmission of learned signatures in a wild parrot. *Proc. R. Soc. B* **279**, 585–591. (doi:10.1098/rspb.2011.0932)

26. Vignal C, Mathevon N, Mottin S. 2008 Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behav. Processes* **77**, 191–198. (doi:10.1016/j.beproc.2007.09.003)

27. Smeele SQ, Senar JC, Aplin LM, Mcelreath MB. 2023 Evidence for vocal signatures and voice-prints in a wild parrot. *bioRxiv*. (doi:10.1101/2023.01.20.524864)

28. Keenan S *et al.* 2020 The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Anim. Behav.* **169**, 9–21. (doi:10.1016/j.anbehav.2020.08.024)

29. Price JJ. 2015 Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Front. Ecol. Evol.* **3**, 40. (doi:10.3389/fevo.2015.00040)

30. Bouchet H, Blois-Heulin C, Lemasson A. 2013 Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Front. Psychol.* **4**, 390. (doi:10.3389/fpsyg.2013.00390)

31. Spillmann B, Van Schaik CP, Setia TM, Sadjadi SO. 2017 Who shall I say is calling? Validation of a caller recognition procedure in Bornean flanged male orangutan (*Pongo pygmaeus wurmbii*) long calls. *Bioacoustics* **26**, 109–120. (doi:10.1080/09524622.2016.1216802)

32. Candiotti A, Zuberbuhler K, Lemasson A. 2012 Convergence and divergence in Dianä monkey vocalizations. *Biol. Lett.* **8**, 382–385. (doi:10.1098/rsbl.2011.1182)

33. Jansen DA, Cant MA, Manser MB. 2012 Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biol.* **10**, 97. (doi:10.1186/1741-7007-10-97)

34. Kershenbaum A, Sayigh LS, Janik VM. 2013 The encoding of individual identity in dolphin signature whistles: how much information is needed? *PLoS ONE* **8**, e77671. (doi:10.1371/journal.pone.0077671)

35. Quick NJ, Janik VM. 2012 Bottlenose dolphins exchange signature whistles when meeting at sea. *Proc. R. Soc. B* **279**, 2539–2545. (doi:10.1098/rspb.2011.2537)

36. Kremers D, Lemasson A, Almunia J, Wanker R. 2012 Vocal sharing and individual acoustic distinctiveness within a group of captive orcas (*Orcinus orca*). *J. Comp. Psychol.* **126**, 433–445. (doi:10.1037/a0028858)

37. Monteiro RD, Ferreira CD, Perbiche-Neves G. 2021 Vocal repertoire and group-specific signature in the smooth-billed ani, *Crotophaga ani* Linnaeus, 1758 (Cuculiformes, Aves). *Papeis Avulsos de Zoologia* **61**, e20216159. (doi:10.11606/1807-0205/2021.61.59)

38. Mundinger PC. 1982 Microgeographic and macrogeographic variation in the acquired vocalizations of birds. In *Acoustic communication in birds* (eds DE Kroodsma, EH Miller), pp. 147–208. New York, NY: Academic Press.

39. Lachlan RF, Verhagen L, Peters S, Cate C. 2010 Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taenopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *J. Comp. Psychol.* **124**, 92–108. (doi:10.1037/a0016996)

40. Clayton NS, Emery NJ. 2007 The social life of corvids. *Curr. Biol.* **17**, 652–656. (doi:10.1016/j.cub.2007.05.070)

41. Coombs CJ. 1960 Observations on the rook *Corvus frugilegus* in southwest Cornwall. *Ibis* **102**, 394–419. (doi:10.1111/j.1474-919X.1960.tb08417.x)

42. Kelley JL, Morrell LJ, Inskip C, Krause J, Croft DP. 2011 Predation risk shapes social networks in fission-fusion populations. *PLoS ONE* **6**, e24280. (doi:10.1371/journal.pone.0024280)

43. Bradbury JW, Balsby TJ. 2016 The functions of vocal learning in parrots. *Behav. Ecol. Sociobiol.* **70**, 293–312. (doi:10.1007/s00265-016-2068-4)

44. Boucherie PH, Mariette MM, Bret C, Dufour V. 2016 Bonding beyond the pair in a monogamous bird: impact on social structure in adult rooks (*Corvus frugilegus*). *Behaviour* **153**, 897–925. (doi:10.1163/1568539X-00003372)

45. Roskaft E, Espmark Y. 1982 Vocal communication by the rook *Corvus frugilegus* during the breeding season. *Ornis Scandinavica* **13**, 38–46. (doi:10.2307/3675971)

46. Munteanu A, Papuc I, Federspiel I, Clayton N, Emery N. 2009 The social context of contact calls by rooks (*Corvus frugilegus*). *Cluj Veterinary Journal* **15**, 33–43. (doi:10.52331/cvj.v15i1.7)

47. Brown ED. 1985 The role of song and vocal imitation among common crows (*Corvus brachyrhynchos*). *Z. Tierpsychol.* **68**, 115–136. (doi:10.1111/j.1439-0310.1985.tb00119.x)

48. Bird CD, Emery NJ. 2009 Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proc. Natl Acad. Sci. USA* **106**, 10 370–10 375. (doi:10.1073/pnas.0901008106)

49. Dally JM, Clayton NS, Emery NJ. 2006 The behaviour and evolution of cache protection and pilferage. *Anim. Behav.* **72**, 13–23. (doi:10.1016/j.anbehav.2005.08.020)

50. Wang R et al. 2009 Song control nuclei in male and female large-billed crows (*Corvus macrorhynchos*). *Zool. Sci.* **26**, 771–777. (doi:10.2108/zsj.26.771)

51. Goodwin D. 1986 *Crows of the world*. Bury St Edmunds, UK: Emdunsbury Press.

52. Sossinka R, Böhner J. 1980 Song types in the zebra finch *Poephila guttata castanotis*. *Z. Tierpsychol.* **53**, 123–132. (doi:10.1111/j.1439-0310.1980.tb01044.x)

53. Dunn AM, Zann RA. 1996 Undirected song in wild zebra finch flocks: contexts and effects of mate removal. *Ethology* **102**, 529–539. (doi:10.1111/j.1439-0310.1996.tb01145.x)

54. Riters LV, Spool JA, Merullo DP, Hahn AH. 2019 Song practice as a rewarding form of play in songbirds. *Behav. Process.* **163**, 91–98. (doi:10.1016/j.beproc.2017.10.002)

55. Singh UA, Iyengar S. 2022 The role of the endogenous opioid system in the vocal behavior of songbirds and its possible role in vocal learning. *Front. Physiol.* **13**, 823152. (doi:10.3389/fphys.2022.823152)

56. Boeckle M, Bugnyar T. 2012 Long-term memory for affiliates in ravens. *Curr. Biol.* **22**, 801–806. (doi:10.1016/j.cub.2012.03.023)

57. Yorzinski JL, Vehrencamp SL, Mcgowan KJ, Clark AB. 2006 The inflected alarm caw of the American crow: differences in acoustic structure among individuals and sexes. *Condor* **108**, 518–529. (doi:10.1093/condor/108.3.518)

58. Stowell D et al. 2016 Bird detection in audio: a survey and a challenge. In *IEEE 26th Int. Workshop on Machine Learning for Signal Processing, Vietri sul Mare, Italy, 13–16 September 2016*. (doi:10.1109/MLSP.2016.7738875)

59. Laiolo P, Palestrini C, Rolando A. 2000 A study of choughs' vocal repertoire: variability related to individuals, sexes and ages. *Journal fur Ornithologie* **141**, 168–179. (doi:10.1046/j.1439-0361.2000.00074.x)

60. Elie JE, Theunissen FE. 2016 The vocal repertoire of the domesticated zebra finch: a data driven approach to decipher the information-bearing acoustic features of communication signals. *Anim. Cogn.* **19**, 285–315. (doi:10.1007/s10071015-0933-6)

61. Sainburg T, Thielk M, Gentner TQ. 2020 Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLoS Comput. Biol.* **16**, e1008228. (doi:10.1371/journal.pcbi.1008228)

62. Audacity Team. 2021 Audacity. Version 3.1.3. See https://audacityteam.org.

63. Martin K et al. 2022 Rookognise: acoustic detection and identification of individual rooks in field recordings using multi-task neural networks. *Ecol. Informatics* **72**, 101818. (doi:10.1016/j.ecoinf.2022.101818)

64. Thomas M et al. 2022 A practical guide for generating unsupervised, spectrogram-based latent space representations of animal vocalizations. *J. Anim. Ecol.* **91**, 1567–1581. (doi:10.1111/1365-2656.13754)

65. Somervuo P. 2019 Time–frequency warping of spectrograms applied to bird sound analyses. *Bioacoustics* **28**, 257–268. (doi:10.1080/09524622.2018.1431958)

66. Mcinnes L, Healy J, Melville J. 2018 UMAP: uniform manifold approximation and projection for dimension reduction. (https://arxiv.org/abs/1802.03426)

67. Mcinnes L, Healy J, Astels S. 2017 hdbscan: hierarchical density based clustering. *J. Open Source Softw.* **2**, 205. (doi:10.21105/joss.00205)

68. Campello RJ et al. 2015 Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Trans. Knowl. Discov. Data* **10**, 5. (doi:10.1145/2733381)

69. Sainburg T, Gentner TQ. 2021 Toward a computational neuroethology of vocal communication: from bioacoustics to neurophysiology, emerging tools and future directions. *Front. Behav. Neurosci.* **15**, 811737. (doi:10.3389/fnbeh.2021.811737)

70. Best P, Paris S, Glotin H, Marxer R. 2023 Deep audio embeddings for vocalisation clustering. *PLoS ONE* **18**, e0283396. (doi:10.1371/journal.pone.0283396)

71. Kaski S, Nikkilä J, Oja M, Venna J, Törönen P, Castrén E. 2003 Trustworthiness and metrics in visualizing similarity of gene expression. *BMC Bioinf.* **4**, 48. (doi:10.1186/1471-2105-4-48)

72. Chao A, Chiu CH, Jost L. 2014 Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through Hill numbers. *Ann. Rev. Ecol. Evol. Syst.* **45**, 297–324. (doi:10.1146/annurev-ecolsys-120213-091540)

73. Chao A et al. 2006 Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* **62**, 361–371. (doi:10.1111/j.1541-0420.2005.00489.x)

74. R Core Team. 2021 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

75. Brooks ME et al. 2017 glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal* **9**, 378–400. (doi:10.32614/RJ-2017-066)

76. Kubinec R. 2022 Ordered beta regression: a parsimonious, well-fitting model for continuous data with lower and upper bounds. *Political Anal.* **31**, 519–536. (doi:10.1017/pan.2022.20)

77. Barton K. 2020 MuMIn: multi-model inference. See https://cran.r-project.org/package=MuMIn.

78. Anderson DR. 2004 Multimodel inference understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304. (doi:10.1177/0049124104268644)

79. Lenth RV. 2022 emmeans: estimated marginal means, aka least-squares means. See https://cran.r-project.org/package=emmeans.

80. Rebout N et al. 2020 Tolerant and intolerant macaques show different levels of structural complexity in their vocal communication. *Proc. R. Soc. B* **287**, 20200439. (doi:10.1098/rspb.2020.0439)

81. Wright TF, Dahlin CR. 2017 Vocal dialects in parrots: patterns and processes of cultural evolution. *Emu* **118**, 50–66. (doi:10.1080/01584197.2017.1379356)

82. Roskaft E. 1983 Sex-role partitioning and parental care by the rook *Corvus frugilegus*. *Ornis Scandinavica* **14**, 180–187. (doi:10.2307/3676151)

83. Tanimoto AM, Hart PJ, Pack AA, Switzer R. 2017 Vocal repertoire and signal characteristics of 'Alalā, the Hawaiian crow (*Corvus hawaiiensis*). *Wilson J. Ornithol.* **129**, 25–35. (doi:10.1676/1559-4491-129.1.25)

84. Adret-Hausberger M. 1982 Social influences on the whistled songs of starlings. *Behav. Ecol. Sociobiol.* **11**, 241–246. (doi:10.1007/BF00299300)

85. Kroodsma DE. 2004 The diversity and plasticity of birdsong. In *Nature's music: the science of birdsong*, pp. 108–131. San Diego, CA: Academic Press. (doi:10.1016/B978-012473070-0/50007-4)

86. Walløe S *et al.* 2015 Differences in short-term vocal learning in parrots, a comparative study. *Behaviour* **152**, 1433–1461. (doi:10.1163/1568539X-00003286)

87. Sewall KB. 2009 Limited adult vocal learning maintains call dialects but permits pair-distinctive calls in red crossbills. *Anim. Behav.* **77**, 1303–1311. (doi:10.1016/j.anbehav.2009.01.033)

88. Luef EM, Ter Maat A, Pika S. 2017 Vocal similarity in long-distance and short-distance vocalizations in raven pairs (*Corvus corax*) in captivity. *Behav. Processes* **142**, 1–7. (doi:10.1016/j.beproc.2017.05.013)

89. Cortopassi KA, Bradbury JW. 2006 Contact call diversity in wild orange-fronted parakeet pairs, *Aratinga canicularis*. *Anim. Behav.* **71**, 1141–1154. (doi:10.1016/j.anbehav.2005.09.011)

90. Liu WC, Wada K, Nottebohm F. 2009 Variable food begging calls are harbingers of vocal learning. *PLoS ONE* **4**, e5929. (doi:10.1371/journal.pone.0005929)

91. Clayton NS, Dickinson A. 1999 Scrub jays (*Aphelocoma coerulescens*) remember the relative time of caching as well as the location and content of their caches. *J. Comp. Psychol.* **113**, 403–416. (doi:10.1037/0735-7036.113.4.403)

92. Griffiths D, Dickinson A, Clayton N. 1999 Episodic memory: what can animals remember about their past? *Trends Cogn. Sci.* **3**, 74–80. (doi:10.1016/S1364-6613(98) 01272-8)

93. Clayton NS, Yu KS, Dickinson A. 2003 Interacting cache memories: evidence for flexible memory use by western scrub-jays (*Aphelocoma californica*). *J. Exp. Psychol.* **29**, 14–22. (doi:10.1037/0097-7403.29.1.14)

94. Austin VI *et al.* 2021 Avian vocalisations: the female perspective. *Biol. Rev.* **96**, 1484–1503. (doi:10.1111/brv.12713)

95. Goffinet J, Brudner S, Mooney R, Pearson J. 2021 Low-dimensional learned feature spaces quantify individual and group differences in vocal repertoires. *eLife* **10**, e67855. (doi:10.7554/eLife.67855)

96. Hultsch H, Kipper S, Mundry R, Todt D. 2004 Long-term persistence of song performance rules in nightingales (*Luscinia megarhynchos*): a longitudinal field study on repertoire size and composition. *Behaviour* **141**, 371–390. (doi:10.1163/156853904322981914)

97. Boughey MJ, Thompson NS. 1981 Song variety in the brown thrasher (*Toxostoma rufum*). *Z. Tierpsychol.* **56**, 47–58. (doi:10.1111/j.1439-0310.1981.tb01283.x)

98. Backhouse F *et al.* 2022 Higher-order sequences of vocal mimicry performed by male Albert's lyrebirds are socially transmitted and enhance acoustic contrast. *Proc. R. Soc. B* **289**, 20212498. (doi:10.1098/rspb.2021.2498)

99. Bertram B. 1970 The vocal behaviour of the Indian hill mynah, *Gracula religiosa*. *Anim. Behav. Monogr.* **3**, 79-IN1. (doi:10.1016/s0066-1856(70)80003-6)

100. Martin K, Cornero FM, Clayton NS, Adam O, Obin N, Dufour V. 2024 Vocal complexity in a socially complex corvid: gradation, diversity and lack of common call repertoire in male rooks. Figshare. (doi:10.6084/m9.figshare.c.6989956)