

Comparative Analysis of the Vocal Repertoire of *Eulemur*: A Dynamic Time Warping Approach



Marco Gamba¹ · Olivier Friard¹ ·
Isidoro Riondato¹ · Roberta Righini¹ ·
Camilla Colombo¹ · Longondraza Miaretsoa² ·
Valeria Torti¹ · Bakri Nadhurou¹ ·
Cristina Giacoma¹

Received: 12 February 2015 / Accepted: 15 July 2015 / Published online: 18 September 2015
© Springer Science+Business Media New York 2015

Abstract The diversity of qualitative approaches and analytical methods has often undermined comparative research on primate vocal repertoires. The purpose of the present work is to introduce a quantitative method based on dynamic time warping to the study of repertoire size in *Eulemur* spp. We obtained a large sample of calls of *E. coronatus*, *E. flavifrons*, *E. fulvus*, *E. macaco*, *E. mongoz*, *E. rubriventer*, and *E. rufus*, recorded between 1999 and 2013 from captive and wild lemurs. We inspected recordings visually using spectrograms, then cut and saved high-quality vocal emissions to single files for further analysis. We extracted the acoustic features of all vocalizations of a species using the Hidden Markov Model Toolkit, an application of dynamic time warping, and then compared cepstral coefficients (a feature widely used in automatic speaker recognition) pairwise. We analyzed the results using affinity propagation clustering. We found that *Eulemur* species share most of their vocal repertoire but species-specific calls determine repertoire size differences. Repertoire size varied from 9 to 14 vocalization types among species, with a mean of 11.

Keywords Acoustic analysis · Cluster analysis · Species-specific differences · True lemurs · Unsupervised classification · Vocal communication

Electronic supplementary material The online version of this article (doi:10.1007/s10764-015-9861-1) contains supplementary material, which is available to authorized users.

✉ Marco Gamba
marco.gamba@unito.it

¹ Department of Life Sciences and Systems Biology, University of Torino, 10123 Torino, Italy

² Faculté de Sciences, University of Mahajanga, 401 Mahajanga, Madagascar

Introduction

Vocal repertoires provide essential information to the study of how communication systems evolve (Maynard Smith and Harper 2003). For example, studies of nonhuman primate vocal communication have provided valuable contributions to the debate about the basis for the evolution of language in humans (Dunbar 2009). Nonhuman primate vocal repertoire size correlates with time spent grooming and with group size (McComb and Semple 2005), providing support for the theory that the complexity of human language has gradually evolved with the increase of social complexity (Dunbar 2009). However, comparative studies of repertoire size are often undermined by two factors. First, vocal repertoire data are derived from studies using different methods (McComb and Semple 2005). Second, identification of the signal categories has traditionally relied on human observers' assessment of differences among vocalizations, and is thus subject to individual criteria. Although multivariate techniques have demonstrated that such categories may be appropriate (Fuller 2014; Gamba and Giacoma 2007; Maretti *et al.* 2010; Range and Fischer 2004), human assessment of vocalization types may reflect differences perceived by humans but not necessarily by the species (Fuller 2014; Green 1975; Hauser 1996).

New methodologies in the study of acoustic communication allow standardization across large datasets with limited assumptions (Clemins *et al.* 2006). These methods provide researchers with computer tools for exploring large databases without the disadvantages of subjective *a priori* classification, and are often referred to as “unsupervised” (Kogan and Margoliash 1997; Stathopoulos *et al.* 2014; Stowell and Plumbley 2014). Among the many methods (Garcia and Reyes Garcia 2003; Koolagudi *et al.* 2012), some used for automatic speech recognition, such as dynamic time warping, are increasingly used to investigate animal communication. Dynamic time warping has been useful for the classification of animal sounds in amphibians (Chen *et al.* 2012), birds (Anderson *et al.* 1996; Clemins and Johnson 2006; Ranjard and Ross 2008; Tao *et al.* 2008; Trawicki *et al.* 2005), marine mammals (Brown and Miller 2007), and primates (Riondato *et al.* 2013). These methods can be used to investigate the vocal repertoire across populations and species (Mercado and Handel 2012; Ranjard *et al.* 2010) and improve our ability to make inferences about the evolution of human language (Fedurek and Slocombe 2011). Although unsupervised classification cannot guarantee to classify calls in a way that is meaningful to animals, it does ensure quantitative objective classification (Pozzi *et al.* 2010).

Owing to their unique evolutionary history, lemurs are important subjects for comparative studies of vocal communication and may provide insights into the selective pressures that may have linked social and vocal complexity (Oda 2009). True lemurs (*Eulemur* spp.) are conspicuously vocal and their vocal repertoire comprises low-pitched and high-pitched sounds (Gamba and Giacoma 2005; Macedonia and Stanger 1994; Petter and Charles-Dominique 1979). The presence of various call variants and combinations has also been demonstrated qualitatively (Macedonia and Stanger 1994). Previous studies showed that vocal repertoire may differ between species in *Eulemur fulvus* (Paillette and Petter 1978), *E. mongoz* (Curtis 1997), *E. macaco* (Gosset *et al.* 2001), and *E. coronatus* (Gamba and Giacoma 2007).

The aim of this study was to investigate objectively the vocal repertoire across *Eulemur* species to understand whether different species show different repertoire size and vocalization types. We used an algorithm based on dynamic time warping to assess sound similarity (Ranjard *et al.* 2010). We then applied cluster analysis to identify groups of similar calls. To understand whether vocal repertoire size differs across *Eulemur* species we applied the same analytical process to datasets for different species, including the brown lemur (*E. fulvus*), the mongoose lemur (*E. mongoz*), the black lemur (*E. macaco*), and the crowned lemur (*E. coronatus*), whose repertoires were investigated in previous studies. We also analyzed three species that were not included in previous quantitative vocal repertoire studies: the red-bellied lemur (*E. rubriventer*), the rufous brown lemur (*E. rufus*), and the blue-eyed black lemur (*E. flavifrons*). Qualitative studies of *Eulemur* species have shown a degree of similarity in the acoustic structure of the calls but shed little light on the quantitative evaluation of similarities and differences, and suffered from subjective identification of the call types (Gamba and Giacoma 2005; Macedonia and Stanger 1994). No previous study has combined, to our knowledge, the study of lemurs' vocal repertoire across different species using a quantitative unsupervised methodology.

We tested whether or not our unsupervised analyses identified the same vocalization types as previously described. Human sound recognition mechanisms are robust against noise changes and integrate many factors, resulting in accurate low-level acoustic classification. Humans can differentiate calls as discrete types when an unsupervised program, and possibly other species, would recognize a single type (Hauser 1996; Lippmann 1997). We, therefore, predicted that unsupervised clustering would find fewer vocalization types than previous studies. We also predicted that more variable vocalization types mask variation at a lower level, as in a clustering analysis of Guinea baboon calls (*Papio papio*: Maciej *et al.* 2013). Alternatively, cluster analysis may highlight variants of vocal types showing a particular contextual occurrence and other types that overlap with the *a priori* classification.

Methods

Subjects, Study Sites, Equipment, Data Collection, and Analysis

The recordings analyzed for the purpose of this study were part of a large collection of lemur sounds at the Department of Life Sciences and Systems Biology, University of Torino. The recordings originate from various recording campaigns focused on lemur vocal behavior that took place between 1999 and 2013. They were recorded in the wild and in captivity. The number of recording campaigns (hereafter corpora) and the number of calls within a corpus vary with species. We considered only calls emitted by adults. Detailed information about the corpora, sampling, data collection, and associated references is given in the Electronic Supplementary Material (ESM) Appendix S1.

Clustering Analyses

To identify independent groupings and to visualize emerging vocal types (Nowicki and Nelson 1990), we clustered vocalizations of each species on the basis of their degree of

Table 1 Distribution of the vocalizations indicated *a priori* and as they emerged from the cluster analysis

Vocalization	<i>E. fulvus</i>	<i>E. rufus</i>	<i>E. rubriventer</i>	<i>E. mongoz</i>	<i>E. coronatus</i>	<i>E. flavifrons</i>	<i>E. macaco</i>
Alarm call	0	1		0	4	5	1
Alarm long grunt		0		1			
Chatter	1		0	0		0	0
Click	0	0	0	1			0
Group cohesion call	1	1		1			
Grunted hoot	0	1	0	1	0	2	0
Grunt	2	1	1	1	3	0	1
Grunt-tonal call	1	0	0	1	1	1	
Gurgle			1				
Hoot	0	0	1	0	0	0	0
Hoot-tonal call			2				
Long grunt			1	1		0	1
Long grunt-tonal call	1					2	1
Scream	2	3	2	1	4		5
Snort-grunt			0				
Snort-grunt-tonal call			3				
Snort	0	0	0	0			0
Territorial advertisement call	3	3		1			
Tonal call-grunt	0						
Tonal call	0	0	3	0	1	0	1
Number of clusters	11	10	14	9	13	10	10
Adjusted Rand Index	0.27	0.32	0.18	0.24	0.32	0.32	0.25

The numbers indicate the number of exemplars chosen during the clustering analysis for that particular vocalization. Gray-shaded cells show where a particular vocalization has not been assessed during the *a priori* classification. The number of clusters indicates the total number of clusters emerged during the Affinity Propagation process and the Adjusted Rand Index quantify the agreement between the *a priori* classification and the clustering analysis

dissimilarity, as measured by the pairwise comparison using dynamic time warping (Ranjard *et al.* 2010). Detailed information about the calculation of dissimilarity indices is given in ESM Appendix S1. We used the affinity propagation tool (Frey and Dueck 2007) of the *apcluster* package in R (Bodenhofer *et al.* 2011; Hornik 2013). We labeled clusters with the “representative” vocalization (the “exemplar”), which was automatically chosen during the affinity propagation clustering process (see ESM Appendix S2). The cluster analysis used a squared negative Euclidean distance to measure dissimilarity and identify clusters. This clustering algorithm is based on similarities between pairs of data points. Affinity propagation clustering simultaneously considers all the data points as potential cluster centers (exemplars) and then chooses

the final centers through an iterative process, after which the corresponding clusters also emerge. Although we did not define the number of clusters or the number of exemplars (Bodenhofer *et al.* 2011), the preference (p) with which a data point is chosen as a cluster center influences the number of clusters in the final solution. Because affinity propagation clustering does not automatically converge to an optimal clustering solution, we used two external validation procedures. The first validation was based on the q -scanning process (where q corresponds to the sample quantile of p , modified from Wang *et al.* 2007; see also Bodenhofer *et al.* 2011). We evaluated the clusters obtained using different preferences using the Adjusted Rand Index (Hubert and Arabie 1985) to assess the stability of successive cluster solutions (Hennig 2007). The second cluster validation procedure was based on the Silhouette Index, which reflects the compactness and separation of clusters in the final solution (Maciej *et al.* 2013). When ranked and averaged between species both procedures indicated the median of all the similarities between data points to be the optimal value for the preference. We kept all the analysis settings the same across all datasets. We used the calls used as exemplars in the final clustering solution to label the respective clusters.

***A Posteriori* Evaluation**

We evaluated the agreement between the clustering analyses and the *a priori* classification using the Adjusted Rand Index (Hubert and Arabie 1985; Table I).

The terminology we use in the description of the polar dendrograms refers to Drout and Smith (2013). Each branch of the polar dendrogram is termed a “branch” or a “clade” while the terminal portion of each clade is called a “leaf.” Two-leaved clades are called “bifolious,” but the number of leaves in a clade is not limited. Although the horizontal orientation of dendrograms is irrelevant, its vertical arrangement is meaningful. The vertical position of the branch points indicates how similar or different they are from each other. Branches departing from the same branch point are most similar and belong to the same “level.” In the polar dendrograms, levels are numbered from the center (root) to the outer ring.

We also ran a stepwise discriminant function analysis (sDFA, IBM SPSS Statistics 21; Lehner 1996) using the acoustic parameters measured (ESM Appendix S3; see Gamba and Giacoma 2007 for details) using Praat (University of Amsterdam; Boersma and Weenink 2014). We used the sDFA to identify the weight of the different parameters contributing to the clustering process, although the acoustic analysis does not necessarily simulate feature extraction during the dynamic time warping. We ran the sDFA with the cluster information as the grouping variable to estimate how the acoustic parameters contributed to the classification of calls using leave-one-out cross-validation.

Results

Vocal Repertoire

The cluster analysis showed variation in both the number of clusters and the distribution of calls across clusters with species (Table I; see ESM Appendix S5). Vocalizations

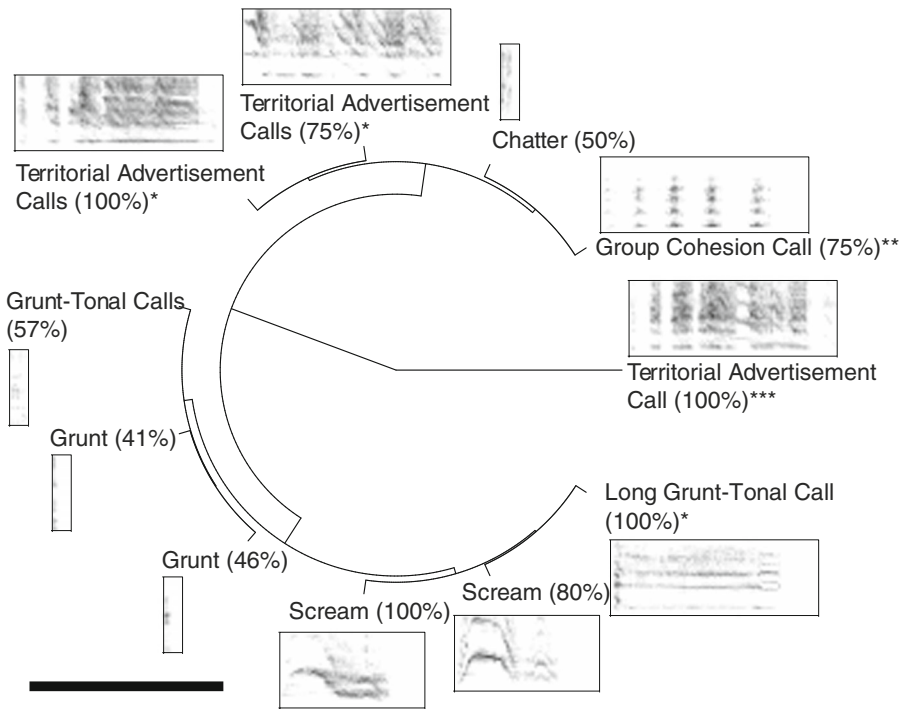


Fig. 1 Polar dendrogram (center) showing how vocalizations of *Eulemur fulvus* cluster together (see ESM Appendix S4 for a detailed description of cluster topology). For each cluster, we show a spectrogram (the horizontal axis represents time; the vertical axis represents frequency) of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–2.50 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Exceptions are indicated as follows: * for 1.25 s, ** for 1.50 s, *** for 2.50 s. Values in parentheses indicate the percentage of the exemplar's vocalization type in a cluster. Additional information is given in ESM Appendixes S4 to S6.

of *Eulemur fulvus* were grouped into 11 clusters (Fig. 1; Table I). sDFA showed an overall correct classification of 84.2 % (cross-validated) when we used the clusters as the grouping variable. Signal duration (on the first discriminant function) and the first formant (F1, on the second discriminant function) had the highest loads in the model (Table II).

Vocalizations of *Eulemur rufus* grouped into 10 clusters (Fig. 2; Table I). sDFA showed an overall correct classification of 94.7 % (cross-validated) when we used the clusters as the grouping variable. Signal duration (on the first discriminant function) and minimum fundamental frequency (MinF0, on the second discriminant function) had the highest loads in the model (Table II).

Vocalizations of *Eulemur rubriventer* grouped into 14 clusters (Fig. 3; Table I). sDFA showed a correct classification of 73.5 % (cross-validated) when we used the clusters as the grouping variable. Signal duration (on the first discriminant function) and the second formant (F2, on the second discriminant function) had the highest loads in the model (Table II).

Vocalizations of *Eulemur mongoz* grouped into nine clusters (Fig. 4; Table I). sDFA showed a correct classification of 69.2 % (cross-validated) when we used

Table II Stepwise discriminant analysis results for the seven *Eulemur* species

Species	Wilks' L.	P	CCR (%)	1st D. f. (%)	2nd D. f. (%)
<i>E. fulvus</i>	0.003	<0.001	84.2	88.9 (Duration)	11.1 (F1)
<i>E. rufus</i>	0.006	<0.001	94.7	98.2 (Duration)	1.0 (MinF0)
<i>E. rubriventer</i>	0.006	<0.001	73.5	91.7 (Duration)	7.2 (F2)
<i>E. mongoz</i>	0.037	<0.001	69.2	81.4 (Duration)	13.9 (F3)
<i>E. coronatus</i>	0.007	<0.001	83.4	96.6 (Duration)	2.8 (F1)
<i>E. flavifrons</i>	0.011	<0.001	71.4	84.6 (Duration)	14.1 (F1)
<i>E. macaco</i>	0.006	<0.001	82.0	78.2 (Duration)	16.1 (F1)

The table shows the statistical results of the seven stepwise Discriminant Function Analyses (sDFA) using temporal parameters (Duration, P_{tmin} , P_{tmax}), fundamental frequency parameters (MeanF0, MinF0, MaxF0, RangeF0, StartF0, EndF0), and formants (F1, F2, F3). The grouping variable for each sDFA was the cluster membership resulted from the Affinity Propagation clustering analysis. We reported the Wilks' Lambda (Wilks' L.) values, the *P*-values (*P*), the cross-validated correct classification rate (CCR), and the variance explained by the first (1st D. f.) and the second (2nd D. f.) discriminant functions. In parentheses, we also reported the parameters showing the highest load on the discriminant functions.

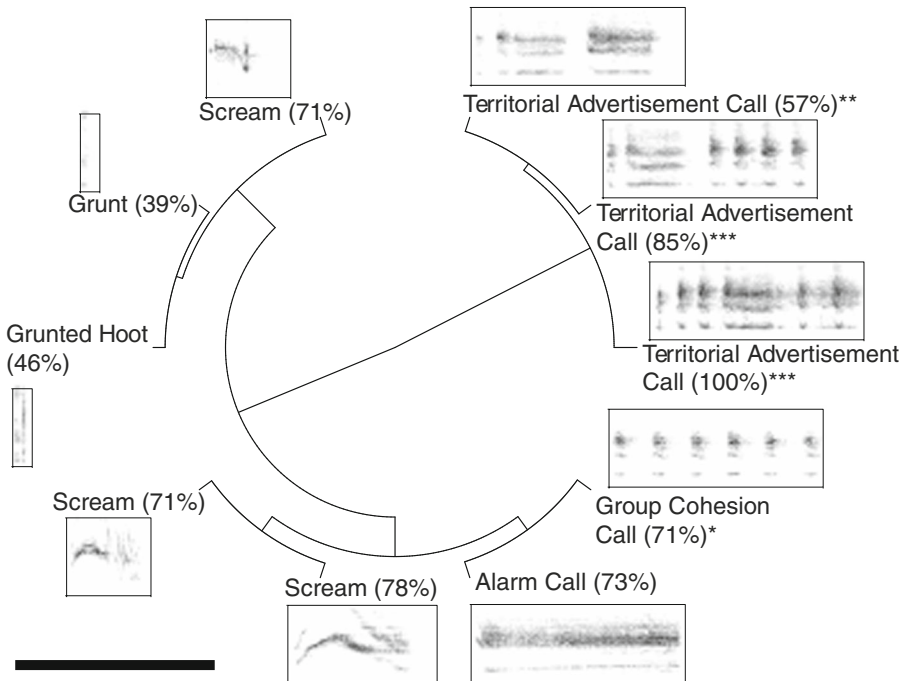


Fig. 2 Polar dendrogram (center) showing how vocalizations of *Eulemur rufus* cluster together (see ESM Appendix S4). For each cluster, we show a spectrogram of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–2.00 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Exceptions are indicated as follows: * for 1.25 s, ** for 1.75 s, *** for 2.00 s. Values in parentheses indicate the percentage of the exemplar's vocalization type in a cluster. Additional information is given in ESM Appendixes S4 to S6.

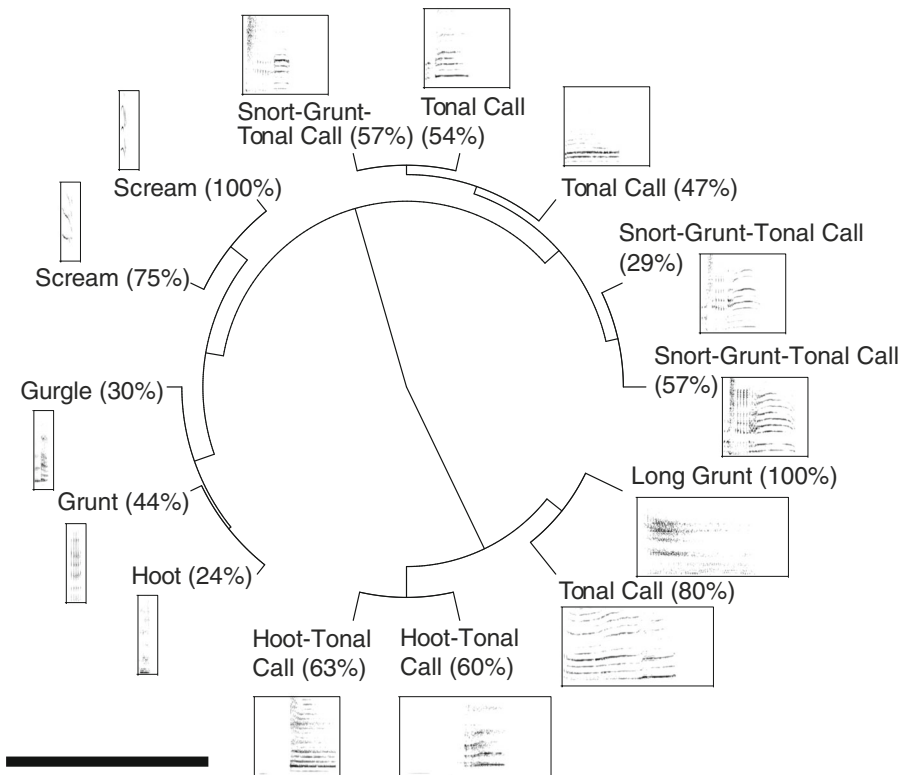


Fig. 3 Polar dendrogram (center) showing how vocalizations of *Eulemur rubriventer* cluster together (see ESM Appendix S4). For each cluster, we show a spectrogram of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–0.75 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Values in parentheses indicate the percentage of the exemplar’s vocalization type in a cluster. Additional information is given in ESM Appendixes S4 to S6.

the clusters as the grouping variable. Signal duration and the third formant (F3) showed the highest loading values on the first and the second discriminant functions respectively (Table II).

Vocalizations of *Eulemur coronatus* grouped into 13 clusters (Fig. 5; Table I). sDFA showed a correct classification of 83.4 % (cross-validated) when we used the clusters as the grouping variable. Signal duration (on the first discriminant function) and the first formant (F1, on the second discriminant function) had the highest loads in the model (Table II).

Vocalizations of *Eulemur flavifrons* grouped into 10 clusters (Fig. 6; Table I). sDFA showed a correct classification of 71.4 % (cross-validated) when we used the clusters as the grouping variable. Signal duration and the first formant had the highest loads on the first two discriminant functions (Table II).

Vocalizations of *Eulemur macaco* grouped into 10 clusters (Fig. 7; Table I). sDFA showed a correct classification of 82.0 % when we used the clusters as the grouping variable. Duration and F1 showed strongest correlation with the first two discriminant functions, respectively (Table II).

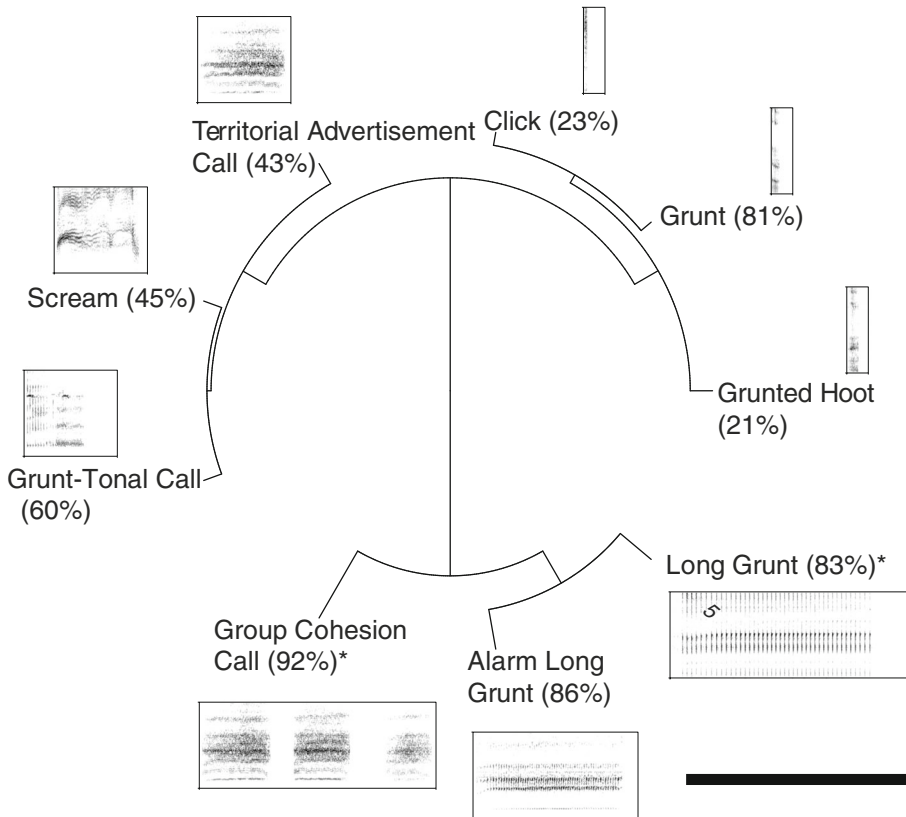


Fig. 4 Polar dendrogram (center) showing how vocalizations of *Eulemur mongoz* cluster together (see ESM Appendix S4). For each cluster, we show a spectrogram of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–1.25 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Exceptions are indicated as * for 1.25 s. Values in parentheses indicate the percentage of the exemplar's vocalization type in a cluster. Additional information is given in ESM Appendixes S4 to S6.

External Cluster Evaluation

The agreement between the *a priori* classification and the grouping identified by the clustering analysis was relatively low across the species, ranging from 0.18 to 0.32 (Table 1).

Discussion

Our approach succeeded in categorizing vocalizations emitted by seven species using dissimilarity indices. Dissimilarity indices have the advantage of being synthetic and convenient but lack the detail of acoustic analysis (Maciej *et al.* 2013; Riondato *et al.* 2013). The discriminant model based on measures of temporal and frequency parameters demonstrated that true lemur calls can be assigned to independently derived clusters identified on the basis of dissimilarity indices with a high rate of correct

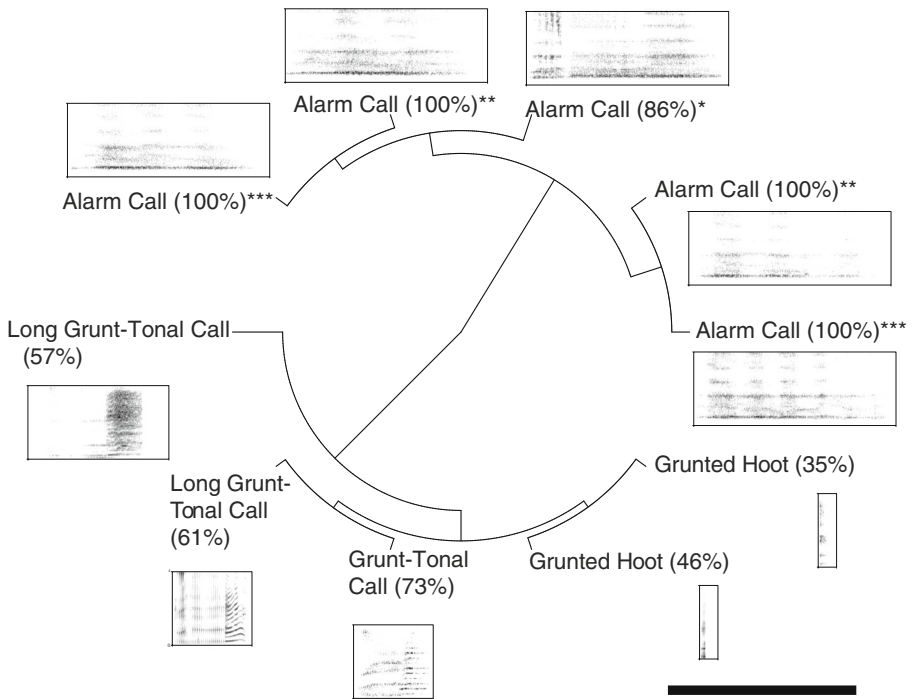


Fig. 6 Polar dendrogram (center) showing how vocalizations of *Eulemur flavifrons* cluster together (see ESM Appendix S4). For each cluster, we show a spectrogram of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–2.50 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Exceptions are indicated as follows: * for 1.25 s, ** for 1.75 s, *** for 2.00 s. Values in parentheses indicate the percentage of the exemplar's vocalization type in a cluster. Additional information is given in ESM S4 to S6.

Our results support the prediction that average group size influences vocal repertoire size in part. Both audio-visual identification and unsupervised classification of vocalization types provide a repertoire size estimate of 14 calls for *Eulemur rubriventer*, an estimate that is surprisingly larger than those observed for other species except *E. coronatus*, which have group sizes of 8.4 (Kappeler and Heymann 1996), whereas *E. rubriventer* has a mean group size of just 3 (Overdorff 1996) or 3.2 (Kappeler and Heymann 1996). *E. mongoz* have a similar average group size of 3.0–3.5 (Kappeler and Heymann 1996; Nadhrou *et al.* 2015) and show a repertoire size of 9 calls. Several authors have suggested a relationship between a species' social organization and its communication, proposing that an egalitarian social structure or stable social groups may favor diversity in communication signals (Mitani 1996). *E. rubriventer* is the only species we studied to have a stable, pair-bonded group structure (Tecot 2008). The other species live in one-male, multifemale groups or multimale, multifemale groups (Fuentes 2002). The social organization in *E. mongoz* varies between populations, and includes both pair bonding and one-male, multifemale groups (Fuentes 2002). The larger distribution of *E. rubriventer* may also influence the diversity of vocal communication, as may the fact that we included only captive *E. rubriventer* in the analysis.

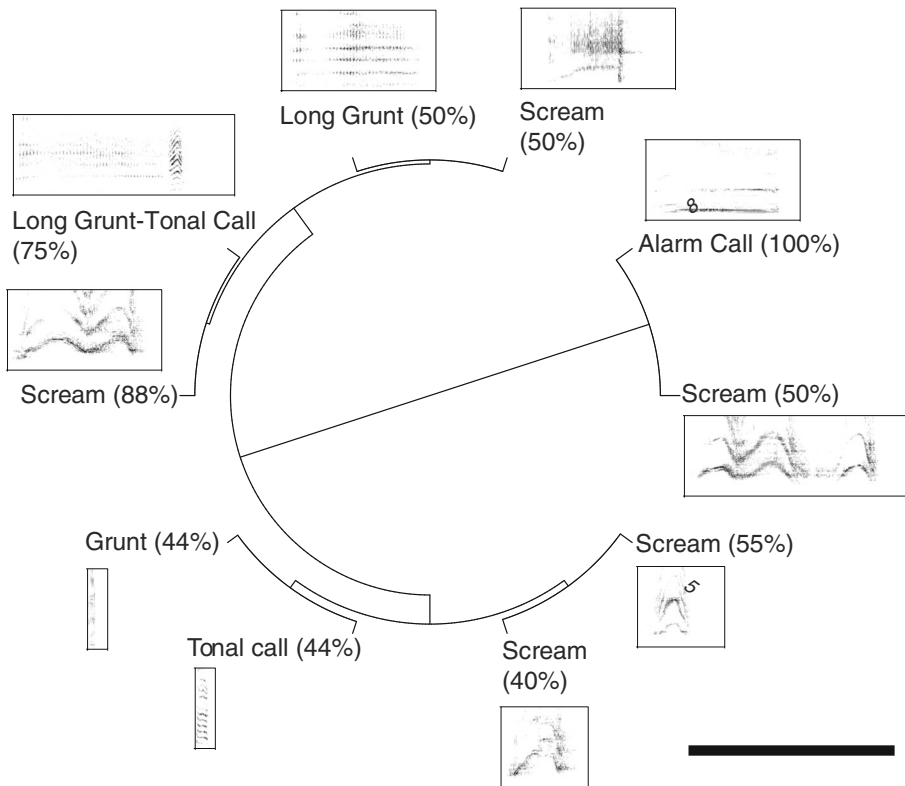


Fig. 7 Polar dendrogram (center) showing how vocalizations of *Eulemur macaco* cluster together (see ESM Appendix S4). For each cluster, we show a spectrogram of the exemplar chosen during the affinity propagation process. All spectrograms are generated in Praat with the following parameters: window length: 0.025 s, time range as shown (0.25–1.00 s); frequency range: 0–10,500 Hz; dynamic range: 35–45 dB. The bar indicates 1 s duration. Values in parentheses indicate the percentage of the exemplar’s vocalization type in a cluster. Additional information is given in ESM Appendices S4 to S6.

However, vocal repertoire appears to be consistent across captive, wild-caught individuals (Colombo, *unpubl. data*), suggesting that other factors may have a stronger effect than the distribution range size. The strong relationships between repertoire size and stable social organization have been proposed for facial expressions (Preuschoft and van Hooff 1995) and the rate of vocal emissions (Mitani 1996), and further studies are needed to clarify whether pair-bonding also “places a selective premium” (Mitani 1996, p. 246) on vocal repertoire size. In support of this proposal, pair-bonding is considered a key factor favoring the convergent evolution of complex singing displays (Geissmann 2000; Torti *et al.* 2013) in the “singing primates” (*Indri indri*, *Tarsius* spp., *Presbytis* spp., and *Hylobates* spp.: Haimoff 1986; *Indri indri*: Bonadonna *et al.* 2014).

We predicted that the unsupervised procedure would recognize a lower number of vocalization types. This was true for *Eulemur fulvus* (11 in the unsupervised analysis vs. 14 in the audio-visual *a priori* assessment), *E. mongoz* (9 vs. 14), *E. rufus* (10 vs. 12) and *E. macaco* (10 vs. 11). The repertoire estimate derived from a previous study of *E. macaco* ($N=13$; Gosset *et al.* 2001) exceeds both that observed during the reassessment process ($N=10$) and the result of the cluster analysis ($N=10$). Although the calls

in our sample may be incomplete, we suspect that this discrepancy arose due to the different criteria used to assess vocalization types in these studies.

Our prediction that the unsupervised procedure would recognize a lower number of vocalization types was not supported in two cases: *Eulemur coronatus* (13 unsupervised vs. 7 audio-visual vocal types) and *E. mongoz* (14 vs. 9). In both cases, the unsupervised procedure recognized more than one type of alarm call. Previous studies of these species estimated a vocal repertoire size of 15 vocalizations for *E. mongoz* (9 validated using sDFA; Nadhrou *et al.* 2015) and 10 vocalizations for *E. coronatus* (all validated using DFA; Gamba and Giacoma 2007). It is clear that different methods led to different estimates, but interesting that, in principle, dynamic time warping allows the identification of vocalization types using a smaller number of calls than sDFA. Whether these differences in vocal repertoire size reflect different arousal states or contexts is an interesting direction for future research.

Cluster vs. *A Priori* Classification

Agreement between the clustering process and the *a priori* criteria was low, with values of the Adjusted Rand Index ranging between 0.18 (in *Eulemur rubriventer*) and 0.32 (in *E. coronatus* and *E. macaco* and *E. rufus*). This supports the prediction that unsupervised clustering of the vocalizations would not find the vocalization types identified in previous studies. However, despite the differences with the *a priori* classification, the clusters obtained using dynamic time warping-generated dissimilarity indices revealed a remarkable potential for grouping calls on the basis of acoustic measurements of different parameters. Among the parameters, duration showed the heaviest loadings on the first discriminant function. Thus, the mismatching between the *a priori* classification and cluster analysis is in line with the suggestion that humans tend to recognize as discrete vocal types sounds that may be grouped into a single type when perceived by other species or classified by quantitative analyses (Hauser 1996).

Both duration and formants contributed to the identification of clusters in almost all the species considered. Formants are known to be crucial for the identification of vocalization types (Gamba 2014; Gamba and Giacoma 2007; Giacoma *et al.* 2011) and have the potential to provide listeners with individual and species-specific cues (Gamba *et al.* 2012a).

Snorts, clicks, and hoots were not selected as cluster representatives and were often grouped with different vocalization types to form fairly dishomogeneous clusters. This result is consistent across the species and is in line with previous data that suggest that low-pitched calls may be part of a graded system more than discrete emissions (Gamba and Giacoma 2007). Identifiable vocalization types are common, but calls with intermediate acoustic structure may also occur and may be either “oversplit” by human listeners or not recognized as discrete by the unsupervised methodology we adopted. Low-pitched calls of *Eulemur* (grunts, clicks, grunted hoots, hoots, snorts, and possibly long grunts) are usually classified as contact calls (Gamba and Giacoma 2005, 2007; Gamba *et al.* 2012a, b; Pflüger and Fichtel 2012; Rendall *et al.* 2000). These low-pitched signals, especially grunts, are the most frequently emitted call type in *Eulemur* (Gamba and Giacoma 2005; Gamba *et al.* 2012a; Pflüger and Fichtel 2012). However, whether acoustic variation in low-pitched signals plays a role in encoding information other than emitter position is still unclear (Pflüger and Fichtel 2012).

The context of call emission is a powerful indicator of their social function and may provide crucial information to the investigation of acoustic structure (Gros-Louis *et al.* 2008; Rendall *et al.* 1999). Future studies are necessary to explore the contextual variation of the vocalization types, how the occurrence of vocal signals relates to their acoustic structure, and how this information can be integrated into unsupervised analyses.

Although there was low agreement between cluster analysis and *a priori* classification, distinct types of grunts and/or grunted hoots emerge in all species. In addition, grunts emitted by *Eulemur coronatus* are identified as three different types. Long grunts, which are reported to denote contexts of disturbance and potential territorial predation, or are emitted during locomotion (Gamba and Giacoma 2005, 2007; Pflüger and Fichtel 2012), occur in *E. mongoz* and *E. fulvus*. Associations between low-pitched calls and tonal calls emerged as distinct clusters (grunt-tonal calls, long grunt-tonal calls) in all species except *E. rufus*, and have been reported for many species (Macedonia and Stanger 1994).

Our findings support the prediction that variation in particular vocal types may mask variation at a lower level, in agreement with a study of Guinea baboon calls (Maciej *et al.* 2013). In baboon calls, variation in screams was stronger than for other vocalization types. In five of six *Eulemur* species, we found that screams represented more than one (usually homogeneous) cluster (*E. flavifrons* did not emit screams in the same situation in which other species emitted them). In *E. fulvus* and *E. rufus*, we identified three clusters of territorial calls, while alarm calls formed three clusters in *E. coronatus* and five clusters in *E. flavifrons*. The fact that cluster analysis identified more than one cluster of alarm calls, screams, and territorial calls indicates variability that has not been reported in previous studies (Gamba and Giacoma 2007; Macedonia and Stanger 1994). These results represent an operationally useful indication for future studies, which may link vocal variation with factors such as level of arousal, social interactions, or audience composition (Clay and Zuberbühler 2012; Fichtel and Hammerschmidt 2002; Slocombe and Zuberbühler 2007; Stoeger *et al.* 2011).

In conclusion, dynamic time warping appears to be a promising method for deepening our knowledge of how lemurs encode information in their vocal signals, and allows the objective identification of vocalization types. We envisage the use of unsupervised classification in different circumstances, including field studies. For example, various researchers report that the classification of calls to be used in playback experiments is particularly challenging. Acoustic analysis may reveal that recorded calls may in fact be different signals (Rendall *et al.* 1999). Researchers can face the problem of classifying calls in different groups when in the field. In these situations, the unsupervised classification of a small number of calls can be very helpful to provide the investigator with an interpretable quantitative analysis, which may result in improved experimental design and aid in the evaluation of the results (Seiler *et al.* 2013).

Acknowledgments This research was supported by the Università di Torino. Grants from the Parco Natura Viva—Centro Tutela Specie Minacciate supported recordings in captivity in the European Zoos and at the PBZT (Antananarivo, Madagascar). We thank Dr. Cesare Avesani Zaborra, Dr. Caterina Spiezo, Gilbert Rakotoarisoa, Jules Medard, Haingoson Randriamialison, Hajanirina Ramino, Fanomezantsoa Andrianirina, Yves Rimpler, Jean-Marc Lermould, Pierre Moisson, Sara De Michelis, Brice Lefaux, Lanto, and Mamatin for their help and logistic support. We also thank Caroline Harcourt, Carlo Comazzi, and Emilio Balletto for their

comments on earlier versions of this manuscript. We thank the editors and two anonymous reviewers for their constructive comments, which helped improve the manuscript.

References

- Anderson, S. E., Amish, D. S., & Margoliash, D. (1996). Template-based automatic recognition of birdsong syllables from continuous recordings. *The Journal of the Acoustical Society of America*, *100*, 1209–1219.
- Bodenhofer, U., Kothmeier, A., & Hochreiter, S. (2011). APCluster: An R package for affinity propagation clustering. *Bioinformatics*, *27*, 2463–2464.
- Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer. Computer program. Version 5.4.04. Retrieved December 28, 2014 from <http://www.praat.org/>
- Bonadonna, G., Torti, V., Randrianarison, R. M., Martinet, N., Gamba, M., & Giacoma, C. (2014). Behavioral correlates of extra-pair copulation in Indri indri. *Primates*, *55*, 119–123.
- Brown, J. C., & Miller, P. J. O. (2007). Automatic classification of killerwhale vocalizations using dynamic time warping. *The Journal of the Acoustical Society of America*, *122*, 1201–1207.
- Chen, W.-P., Cheng, S.-S., Lin, C.-C., Chen, Y. Z., & Lin, W.-C. (2012). Automatic recognition of frog calls using a multi-stage average spectrum. *Computers and Mathematics with Applications*, *64*, 1270–1281.
- Clay, Z., & Zuberbühler, K. (2012). Communication during sex among female bonobos: Effects of dominance, solicitation and audience. *Scientific Reports*, *2*, 291.
- Clemins, P. J., & Johnson, M. T. (2006). Generalized perceptual linear prediction features for animal vocalization analysis. *The Journal of the Acoustical Society of America*, *120*, 527–534.
- Clemins, P. J., Trawicki, M., Adi, K., Tao, J., & Johnson, M. T. (2006). Generalized perceptual feature for vocalization analysis across multiple species. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, *1*, 253–256. Toulouse, France, May 14–19.
- Curtis, D. J. (1997). *The mongoose lemur (Eulemur mongoz): A study in behaviour and ecology*. Ph.D. thesis, University of Zurich.
- Drout, M., & Smith, L. (2013). How to read a dendrogram. National Endowment for the Humanities. Retrieved from lexomics.wheatoncollege.edu (Accessed January 30, 2015).
- Dunbar, R. I. M. (2009). Why only humans have language. In R. Botha & C. Knight (Eds.), *The prehistory of language* (pp. 12–35). Oxford: Oxford University Press.
- Fedurek, P., & Slocombe, K. E. (2011). Primate vocal communication: A useful tool for understanding human speech and language evolution? *Human Biology*, *83*, 153–173.
- Fichtel, C., & Hammerschmidt, K. (2002). Responses of redfronted lemurs to experimentally modified terrestrial alarm calls: Evidence for urgency-based changes in call structure. *Ethology*, *108*, 763–777.
- Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science*, *315*, 972–976.
- Fuentes, A. (2002). Patterns and trends in primate pair bonds. *American Journal of Primatology*, *23*, 953–978.
- Fuller, J. L. (2014). The vocal repertoire of adult male blue monkeys (*Cercopithecus mitis stuhlmanni*): A quantitative analysis of acoustic structure. *American Journal of Primatology*, *76*, 203–216.
- Gamba, M. (2006). *Evoluzione della comunicazione vocale nei lemuri del Madagascar*. Ph.D. dissertation, University of Turin, Italy.
- Gamba, M., & Giacoma, C. (2005). Key issues in the study of primate acoustic signals. *Journal of Anthropological Sciences*, *83*, 61–87.
- Gamba, M. (2014). Vocal tract-related cues across human and nonhuman signals. In A. Pennisi et al. (Eds.), *Reti, saperi, linguaggi* (pp. 49–68). Bologna: Il Mulino.
- Gamba, M., & Giacoma, C. (2007). Quantitative acoustic analysis of the vocal repertoire of the crowned lemur. *Ethology Ecology & Evolution*, *19*, 323–343.
- Gamba, M., Colombo, C., & Giacoma, C. (2012a). Acoustic cues to caller identity in lemurs: A case study. *Journal of Ethology*, *30*, 191–196.
- Gamba, M., Friard, O., & Giacoma, C. (2012b). Vocal tract morphology determines species-specific features in vocal signals of lemurs (*Eulemur*). *International Journal of Primatology*, *33*, 1453–1466.
- García, J., & Reyes García, C. (2003). Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks. *Proceedings of the International Joint Conference on Neural Networks*, *4*, 3140–3145.
- Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 103–123). Cambridge, MA: MIT Press.
- Giacoma, C., Sorrentino, V., Rabarivola, C., & Gamba, M. (2011). Sex differences in the song of *Indri indri*. *International Journal of Primatology*, *31*, 539–551.

- Gosset, D., Fornasieri, I., & Roeder, J. J. (2001). Acoustic structure and contexts of emission of vocal signals by black lemurs. *Evolution of Communication*, 4, 225–251.
- Green, S. (1975). Dialects in Japanese monkeys: Vocal learning and cultural transmission of locale-specific vocal behavior? *Zeitschrift für Tierpsychologie*, 38, 304–314.
- Gros-Louis, J., Perry, S., Fichtel, C., Wikberg, E., Gilkenson, H., Wofsy, S., & Fuentes, A. (2008). Vocal repertoire of white-faced capuchin monkeys (*Cebus capucinus*): Acoustic structure, context and usage. *International Journal of Primatology*, 29, 641–670.
- Haimoff, E. H. (1986). Convergence in the duetting of monogamous Old World primates. *Journal of Human Evolution*, 15, 51–59.
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis*, 52, 258–271.
- Hornik, K. (2013). The R FAQ. Resource document. Retrieved from <http://CRAN.R-project.org/doc/FAQ/R-FAQ.html>.
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2, 193–218.
- Kappeler, P. M., & Heymann, E. W. (1996). Nonconvergence in the evolution of primate life history and socio-ecology. *Biological Journal of the Linnean Society*, 59, 297–326.
- Kogan, J. A., & Margoliash, D. (1997). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. *Journal of the Acoustical Society of America*, 103, 2185–2196.
- Koolagudi, S. G., Rastogi, D., & Rao, K. S. (2012). Identification of language using mel-frequency cepstral coefficients (MFCC). *Procedia Engineering*, 38, 3391–3398.
- Lehner, P. N. (1996). *Handbook of ethological methods* (2nd ed.). New York: Cambridge University Press.
- Lippmann, R. P. (1997). Speech recognition by machines and humans. *Speech Communication*, 22, 1–15.
- Macedonia, J. M., & Stanger, K. F. (1994). Phylogeny of the Lemuridae revisited: Evidence from communication signals. *Folia Primatologica*, 63, 1–43.
- Maciej, P., Ndao, I., Hammerschmidt, K., & Fischer, J. (2013). Vocal communication in a complex multi-level society: Constrained acoustic structure and flexible call usage in Guinea baboons. *Frontiers in Zoology*, 10, 58.
- Maretti, G., Sorrentino, V., Finomana, A., Gamba, M., & Giacoma, C. (2010). Not just a pretty song: An overview of the vocal repertoire of *Indri indri*. *Journal of Anthropological Sciences*, 88, 151–165.
- Maynard Smith, J., & Harper, D. (2003). *Animal signals* (Oxford Series in Ecology and Evolution). Oxford: Oxford University Press.
- McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biological Letters*, 1, 381–385.
- Mercado, E., III, & Handel, S. (2012). Understanding the structure of humpback whale songs. *Journal of the Acoustical Society of America*, 132, 2947–2950.
- Mitani, J. C. (1996). Comparative field studies of African ape vocal behavior. In W. McGrew, L. Marchant, & T. Nishida (Eds.), *Great ape societies* (pp. 241–254). Cambridge, U.K.: Cambridge University Press.
- Mittermeier, R. A., Ganzhorn, J. U., Konstant, W. R., Glander, K., Tattersall, I., Groves, C. P., Rylands, A. B., Hapke, A., Ratsimbazafy, J., Mayor, M. I., Louis, E. E., Jr., Rumpler, Y., Schwitzer, C., & Rasoloarison, R. M. (2008). Lemur diversity in Madagascar. *International Journal of Primatology*, 29, 1607–1656.
- Nadhrou, B., Gamba, M., Andriaholinirina, N. V., Ouledi, A., & Giacoma, C. (2015). The vocal communication of the mongoose lemur (*Eulemur mongoz*): Phonation mechanisms, acoustic features and quantitative analysis. *Ethology Ecology & Evolution*. doi:10.1080/0394937020151039069.
- Nowicki, S., & Nelson, D. A. (1990). Defining natural categories in acoustic signals: Comparison of three methods applied to ‘chick-a-dee’ call notes. *Ethology*, 86, 89–101.
- Oda, R. (2009). Lemur vocal communication and the origin of human language. In T. Matsuzawa (Ed.), *Primate origins of human cognition and behavior* (pp. 115–134). New York: Springer Science+Business Media.
- Overdorff, D. J. (1996). Ecological correlates to activity and habitat use of two prosimian primates: *Eulemur rubriventer* and *Eulemur fulvus rufus* in Madagascar. *American Journal of Primatology*, 40, 327–342.
- Paillette, M., & Petter, J. J. (1978). Vocal repertoire of *Lemur fulvus albifrons*. In D. J. Chivers & J. Herbert (Eds.), *Recent advances in primatology* (pp. 831–834). London: Academic Press.
- Petter, J. J., & Charles-Dominique, P. (1979). Vocal communication in prosimians. In G. A. Doyle & R. D. Martin (Eds.), *The study of prosimian behaviour* (pp. 272–282). New York: Academic Press.
- Pflüger, F. J., & Fichtel, C. (2012). On the function of redfronted lemur’s close calls. *Animal Cognition*, 15, 823–831.

- Pozzi, L., Gamba, M., & Giacoma, C. (2010). The use of artificial neural networks to classify primate vocalizations: A pilot study on black lemurs. *American Journal of Primatology*, *72*, 337–348.
- Preuschoft, S., & van Hooff, J. A. R. A. M. (1995). Homologizing primate facial displays: A critical review of methods. *Folia Primatologica*, *65*, 121–137.
- Range, F., & Fischer, J. (2004). Vocal repertoire of sooty mangabeys (*Cercocebus torquatus atys*) in the Tai National Park. *Ethology*, *110*, 301–321.
- Ranjard, L., & Ross, H. A. (2008). Unsupervised bird song syllable classification using evolving neural networks. *Journal of the Acoustical Society of America*, *123*, 4358–4368.
- Ranjard, L., Anderson, M. G., Rayner, M. J., Payne, R. B., McLean, I., Briskie, J. V., et al. (2010). Bioacoustic distances between the begging calls of brood parasites and their host species: A comparison of metrics and techniques. *Behavioral Ecology and Sociobiology*, *64*, 1915–1926.
- Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, *57*, 583–592.
- Rendall, D., Cheney, D. L., & Seyfarth, R. M. (2000). Proximate factors mediating ‘contact’ calls in adult female baboons and their infants. *Journal of Comparative Psychology*, *114*, 36–46.
- Riondato, I., Giuntini, M., Gamba, M., & Giacoma, C. (2013). Vocalization of red- and grey-shanked douc langur (*Pygathrix nemaeus* and *P. cinerea*). *Vietnamese Journal of Primatology*, *2*, 75–82.
- Seiler, M., Schwitzer, C., Gamba, M., & Holderied, M. W. (2013). Interspecific semantic alarm call recognition in the solitary sahalalaza sportive lemur, *Lepilemur sahalalazensis*. *PLoS ONE*, *8*, e67397.
- Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *PNAS*, *104*(43), 17228–17233.
- Stathopoulos, S., Bishop, J. M., & O’Ryan, C. (2014). Genetic signatures for enhanced olfaction in the African mole-rats. *PLoS ONE*, *9*, e93336.
- Stoeger, A. S., Charlton, B. D., Kratochvil, H., & Fitch, W. T. (2011). Vocal cues indicate level of arousal in infant African elephants. *Journal of the Acoustical Society of America*, *130*, 1700–1710.
- Stowell, D., & Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *Peer J*, *2*, e488. doi:10.7717/peerj.488.
- Tao, J., Johnson, M. T., & Osiejuk, T. S. (2008). Acoustic model adaptation for ortolan bunting (*Emberiza hortulana* L.) song-type classification. *Journal of the Acoustical Society of America*, *123*, 1582–1590.
- Tattersall, I., & Sussman, R. (1998). ‘Little brown lemurs’ of northern Madagascar. *Primatologica*, *69*, 378–388.
- Tecot, S. R. (2008). *Seasonality and predictability: The hormonal and behavioral responses of the red-bellied lemur, Eulemur rubriventer, in southeastern Madagascar*. Ph.D. dissertation, University of Texas.
- Torti, V., Gamba, M., Rabermanajara, Z., & Giacoma, C. (2013). The songs of the indris (Mammalia: Primates: Indridae): Contextual variation in the long-distance calls of a lemur. *Italian Journal of Zoology*, *80*, 596–607.
- Trawicki, M. B., Johnson, M. T., & Osiejuk, T. S. (2005). Automatic song-type classification and speaker identification of Norwegian ortolan bunting (*Emberiza hortulana*) vocalizations. In *IEEE Workshop on Machine Learning for Signal Processing*. 28 Sept. 2005, Mystic, CT, USA, pp. 277–282.
- Wang, K., Zhang, J., Li, D., Zhang, X., & Guo, T. (2007). Adaptive affinity propagation clustering. *Acta Automatica Sinica*, *33*, 1242–1246.