Classification of producer characteristics in primate long calls using neural networks

Efstathia Robakis[a)]
Washington University in Saint Louis, One Brookings Drive, St. Louis, Missouri 63130, USA

Mrinalini Watsa
Field Projects International, 7331 Murdoch Avenue, St. Louis, Missouri 63119, USA

Gideon Erkenswick
Field Projects International, 7331 Murdoch Avenue, St. Louis, Missouri 63119, USA

(Received 24 January 2018; revised 3 June 2018; accepted 28 June 2018; published online 25 July 2018)

Primate long calls are high-amplitude vocalizations that can be critical in maintaining intragroup contact and intergroup spacing, and can encode abundant information about a call's producer, such as age, sex, and individual identity. Long calls of the wild emperor (*Saguinus imperator*) and saddleback (*Leontocebus weddelli*) tamarins were tested for these identity signals using artificial neural networks, machine-learning models that reduce subjectivity in vocalization classification. To assess whether modelling could be streamlined by using only factors which were responsible for the majority of variation within networks, each series of networks was re-trained after implementing two methods of feature selection. First, networks were trained and run using only the subset of variables whose weights accounted for $\geq$50% of each original network's variation, as identified by the networks themselves. In the second, only variables implemented by decision trees in predicting outcomes were used. Networks predicted dependent variables above chance ($\geq$58.7% for sex, $\geq$69.2 for age class, and $\geq$38.8% for seven to eight individuals), but classification accuracy was not markedly improved by feature selection. Findings are discussed with regard to implications for future studies on identity signaling in vocalizations and streamlining of data analysis.

© *2018 Acoustical Society of America*. https://doi.org/10.1121/1.5046526

[MLD] Pages: 344–353

I. INTRODUCTION

Primate long calls—high amplitude long-distance vocalizations, sometimes also called loud calls or contact calls—serve several functions. They can be used to maintain intergroup spacing and territory defense (Waser, 1977; Mitani, 1985; Rasoloharijaona *et al.*, 2006; da Cunha and Jalles-Filho, 2007; Caselli *et al.*, 2014), and for mate defense and/or assessment (Mitani, 1984; Cowlishaw, 1996; Geissmann, 1999; Bolt, 2013). They also facilitate the maintenance of intragroup cohesion (*New World monkeys:* Snowdon and Hodun, 1981; Spehar and Di Fiore, 2013; Dubreuil *et al.*, 2015; *Old World monkeys and lemurs*: Byrne, 1982; Macedonia, 1986; Cheney *et al.*, 1996; Ramanankirahina *et al.*, 2016; *apes:* Mitani and Nishida, 1993; White *et al.*, 2015). These functions are made possible by the broad range of information encoded therein: calls can convey information regarding external references, such as location (Snowdon and Hodun, 1981), but also intrinsic qualities of the producer, such as sex, age, and identity (Miller and Thomas, 2012; León *et al.*, 2014; Terleph *et al.*, 2015).

Callitrichids, a diverse family of Neotropical primates, use species-specific long calls for inter- and intragroup communication (Epple, 1968; Brown *et al.*, 1979; Jorgensen and French, 1998; Ruiz-Miranda *et al.*, 1999; Lazaro-Perea, 2001; Windfelder, 2001; de la Torre and Snowdon, 2009). Long calls can be used by an individual seeking to reunite with its group after it has become isolated during ranging and for coordination with fellow group members during agonistic intergroup encounters. For arboreal primates such as callitrichids, visual contact with other group members can be hindered by distance or dense tropical forest. Encoding sex and individual identity in a long call can be important for maintaining intragroup spacing, during ranging, and in intergroup encounters (Arnedo *et al.*, 2010; Lemasson and Hausberger, 2011; Bouchet *et al.*, 2012; Dubreuil *et al.*, 2015). Similarly, recognizing age class from an individual's long call is particularly relevant for infants and juveniles, who often vocalize to solicit assistance or food from adults, and who are the most vulnerable when separated from the rest of their group (Epple, 1968; Chen *et al.*, 2009). Transmission of the characteristics of a call's producer in vocalizations therefore have marked effects on behavior and survival, and captive studies have confirmed that sex (Masataka, 1987; Norcross and Newman, 1993, age (Pola and Snowdon, 1975; Pistorio *et al.*, 2006; Chen *et al.*, 2009), and individual (Jorgensen and French, 1998; Weiss *et al.*, 2001; Miller *et al.*, 2010) can be encoded in the acoustic structure of the long calls of several Callitrichine species.

In the present study we assess whether the long calls of two species of sympatric wild callitrichids, emperor (*Saguinus imperator*) and saddleback (*Leontocebus weddelli*) tamarins

---

[a)] Electronic mail: erobakis@wustl.edu can be discriminated by sex, age class, and individual identity. To our knowledge, neither emperor nor saddleback tamarin vocalizations have been tested for this suite of identity signals in the wild. Both species live in cooperatively breeding social systems, which can have marked effects on the evolution of vocal communication (Burkart and van Schaik, 2009; Freeberg *et al.*, 2012). This includes the capacity for transmitting identity signals in vocalizations, which are important to adult-juvenile interactions, intergroup encounters, dispersal, and group cohesion (Marten *et al.*, 1977; Freeberg *et al.*, 2012). We therefore predict that these signals—age, sex, and identity—will be discernable in the vocalizations of both species. A comparison of cooperatively breeding species can control for social factors, shed light on the adaptive pressures shaping signal variation, and in this case, help disentangle more complicated questions regarding how callitrichid behavior affects and is affected by signal variation in the wild.

To test whether vocalizations can be accurately classified according to identity signals, we use artificial neural networks (ANNs), a method of machine learning. ANNs function analogously to the brain's neurons. Inputs, such as bioacoustic measurements, are used to predict outputs, such as signaler characteristics, via a hidden layer of neurons, which are connected to inputs via axons (Fig. 1). When the network is trained, axons are differentially weighted as the network learns, through an iterative process, which model best predicts outputs (Reby *et al.*, 1997). After the network is trained, the result is a web of neurons connected by weighted axons, where inputs can be classified as outputs according to the best model using novel datasets. ANNs can accommodate any number and type of independent variables, but function best with numeric variables that have been scaled and centered. ANNs have primarily been used to classify vocal repertoires and dialects (*Orcinus orca:* Deecke *et al.*, 1999; *Balaena mysticetus:* Potter *et al.*, 1994; *Eulemur macaca:* Pozzi *et al.*, 2010), and here we assess the extension of their applicability to detecting characteristics of signal producers in primates.

An advantage to neural networks is that they run chiefly without supervision. Users set certain parameters that define the space over which the network can explore optimal models, such as the number of neurons in a hidden layer and the number of hidden layers, but the networks themselves identify the best parameters through the iterative process of model training. Though optimal models do not inherently use a subset of the original variable set (i.e., neural networks do not perform feature selection), users can view each variable's weight in the final model. Weights range from 0 to 100, where higher, or "heavier," values indicate higher relative importance in the model, with all variable weights summing to 100. These weights can therefore be used to identify which independent variables are most diagnostic in predicting outcomes in order to optimize models with fewer inputs for future applications. In order to test whether using highly weighted variables improve network classification accuracy, here we re-train each network using only variables that have been identified by the networks themselves as being highly weighted. Choosing input variables from the dataset that minimize statistical noise while preserving important variation is time-consuming but sometimes critical for model optimization (Yang and Pederson, 1997; Guyon and Elisseeff, 2003); however, using weights assigned to variables during neural network training is one of many methods of feature selection (Dash and Liu, 1997). Thus, a second goal of this study is to compare this feature selection method with one that identifies important variables via decision trees (Tirelli and Pessani, 2011; Wu, 2009). Decision trees are supervised learning algorithms that can be used to predict outcomes using both continuous and categorical independent variables. The heterogenous population is presented as a root node, and the population is split into increasingly

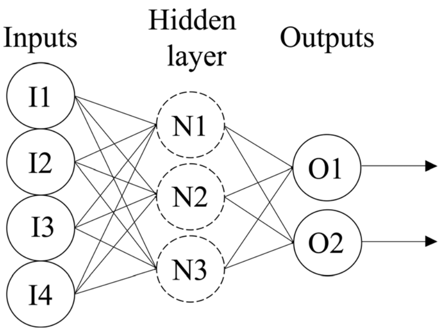

FIG. 1. Schematic of an artificial neural network. The model differentially weights connections between inputs and hidden layer(s) over an iterative process that results in the best method of predicting outputs.

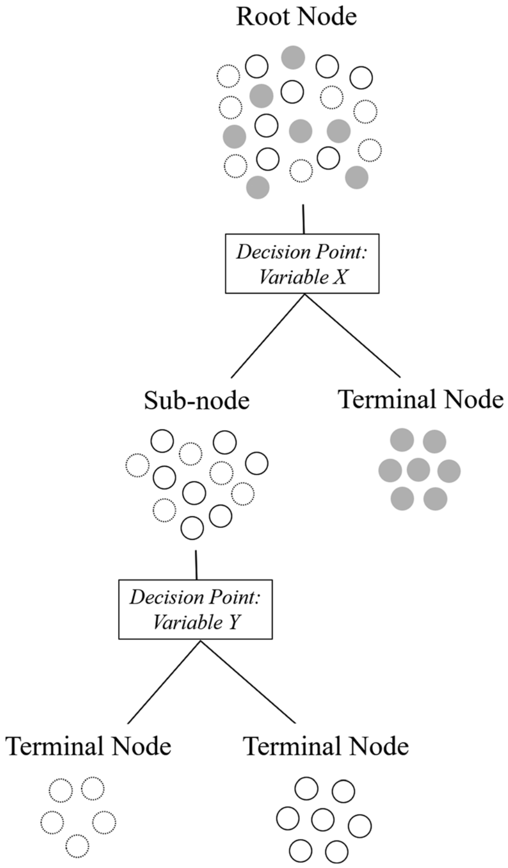

FIG. 2. Schematic of a decision tree. The algorithm progressively uses variables at each decision point to create increasingly homogenous clusters until dependent variables, or terminal nodes, are reached.

homogenous sub-nodes using one independent variable at each "decision point" until outcome variables are isolated (Fig. 2). To assess whether either method of feature selection results in better classification rates of novel datasets, we compare classification accuracies of networks trained on "heavily weighted" variables and networks trained only on variables used in constructing decision trees.

II. METHODS

A. Data collection

Data were collected at the Estación Biologica Rio Los Amigos (EBLA) in the Madre de Dios department of Peru (12°30′–12°36′S and 70°02′–70°09′W) between May and August, 2014–2016. The study populations included five groups each of emperor tamarins (*Saguinus imperator*) (44 unique individuals across three years) and saddleback tamarins (*Leontocebus weddelli*) (51 unique individuals across three years). However, because not all individuals are recaptured annually due to deaths or disappearances, long calls were not recorded from all individuals or groups each year (Table I).

These populations are part of an ongoing mark-recapture program in place since 2009. Monkeys are processed as a group, and age-class and sex are noted for each tamarin. Age classes are assigned as follows: infant = $0–5\frac{3}{4}$ months; juvenile = $5\frac{3}{4}–9\frac{3}{4}$ months; subadult = $9\frac{3}{4}–20$ months; adult = >20 months (Watsa, 2013). Individuals' identities are confirmed across years by the presence of a microchip. Each tamarin is also fitted with a temporary beaded collar, replaced each year, that indicates group, sex, and individual identity, and given unique a bleach pattern on its tail (see Watsa *et al.*, 2015 for a detailed protocol). These visual aids are used to confirm identity during data collection. Data on behavior, including mating and social systems, have been collected on emperor tamarins since 2011, and on saddleback tamarins since 2009 (Watsa, 2013). Vocalizations from untagged individuals were not included in analyses.

Vocalizations were recorded *ad libitum* during 15-min focal follows and during mark-recapture events. Focal follows, which were randomized to balance data collection across individuals and groups over the course of each field season (May–August), were conducted by teams of two researchers. Using a Zoom H5 or Zoom H6 Handy Recorder with accompanying shotgun microphone (Zoom North America, Hauppauge, NY) at the highest available sampling rates (44-kHz/24-bit or 96-kHz/24-bit, respectively), one researcher recorded continuously throughout the focal follow in order to capture vocalizations emitted by the focal individual. When a long call occurred, the identity of the producer was confirmed by speaking into the recorder. If a long call was produced by an identified individual outside of a focal follow, or if the focal individual was out of sight and another individual began calling, observers initiated an *ad*

TABLE I. Individuals in the study populations of *Saguinus imperator* and *Leontocebus weddelli* whose long calls we recorded from 2014 to 2016, and sample sizes of long calls for each age-sex class presented in parentheses. AF = Adult female, AM = adult male, SF = subadult female, SM = subadult male, JF = juvenile female, JM = juvenile male. Adults = >20 months; subadults = $9\frac{3}{4}$–20 months; juveniles = $5\frac{3}{4}$–$9\frac{3}{4}$ months (Watsa, 2013). There were no infants in the study population. Vocalizations from all groups were also recorded during mark-recapture events except where indicated with an asterisk.

Year	Group	AF	AM	SF	SM	JF	JM	Total long call sample size
Saguinus imperator								
2014	SI-3	—	—	—	—	—	1 (2)	2
	SI-4	1 (2)	—	—	—	—		2
2015	SI-1	2 (15)	2 (13)	—	—	—	1 (15)	43
	SI-2*	—	—	—	1 (2)	—	—	2
	SI-3	1 (8)	1 (16)	—	1 (5)	—	—	29
	SI-4	1 (7)	—	—	—	—	—	7
	SI-5*	1 (3)	—	1 (1)	—	—	—	4
2016	SI-1	2 (28)	1 (1)	—	—	—	—	29
	SI-2	1 (9)	1 (7)	—	—	—	—	16
	SI-3	3 (25)	2 (14)	—	—	—	—	39
	SI-4	1 (13)	—	—	1 (4)	—	1 (15)	32
2014–2016	Total long call sample size	110	51	1	11	0	32	**205**
Leontocebus weddelli								
2014	LW-1	1 (2)	—	—	—	—	—	2
	LW-2*	1(1)	1 (7)	—	—	—	—	8
2015	LW-1	2 (12)	2 (8)	—	—	—	1 (5)	25
	LW-2	1 (14)	2 (34)	—	—	1 (2)	—	50
	LW-3	1 (15)	1 (10)	—	1 (2)	1 (4)	—	31
	LW-4	1 (2)	1 (8)	—	—	—	—	10
	LW-5	—	1 (8)	—	—	—	1 (1)	9
2016	LW-1	1 (12)	2 (16)	—	—	—	—	28
	LW-2	1 (2)	2 (23)	—	—	—	—	25
	LW-3	1 (2)	2 (9)	—	—	—	—	11
2014–2016	Total long call sample size	61	123	0	2	7	6	**199**

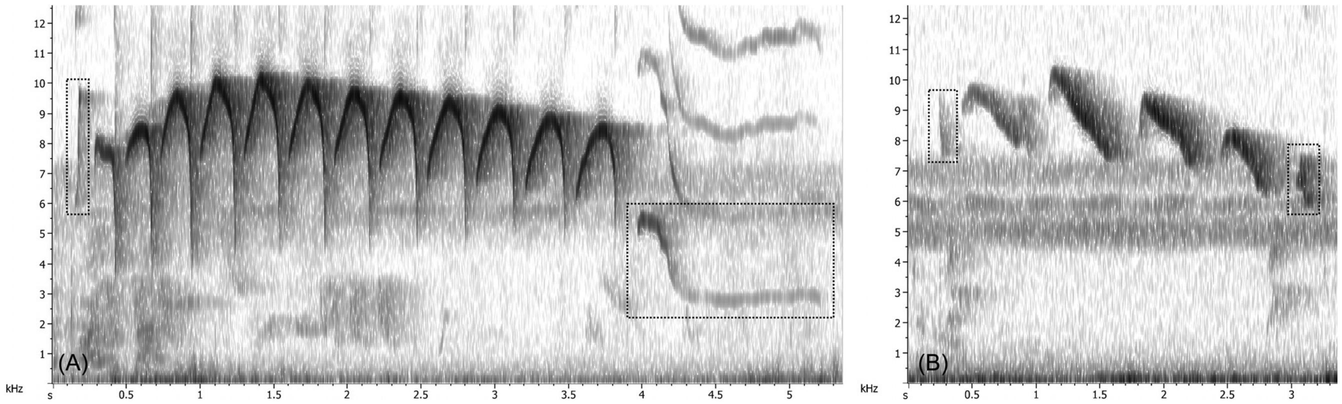

FIG. 3. Spectrograms of long calls of (A) *Saguinus imperator* (14 syllables) and (B) *Leontocebus weddelli* (six syllables), each with one introductory and one terminal syllable. Introductory and terminal syllables are enclosed by dashed lines.

libitum recording of the vocalizer until the individual went out of sight. If an *ad libtum* follow lasted longer than the duration of the 15-min focal follow, the original follow was not resumed, and the next follow on a different animal was begun. Recordings were also made during 13 group mark-recapture events from 2014 to 2016 (Table I). Observers recorded animal vocalizations from 3 to 6 m away.

All data collection was carried out with approval from the Animal Studies Committees of Washington University in St. Louis, the University of Missouri–St. Louis, and the Peruvian Ministry of the Environment (SERFOR), permit number 193-2015-SERFOR-DGGSPFFS.

B. Call measurement

Spectrograms of all long calls were visually inspected in Raven Pro (Hann window, 5.33 mS window size, 1.2 mS hop size, and 2048 DFT to correct for different sampling rates) (Bioacoustics Research Program, 2014), and we discarded those with low signal-to-noise ratios. The long calls of both species comprise a series of disconnected syllables (Fig. 3). As such, we made two sets of measurements for each long call: those based on individual syllables ("syllable set") and those based on the call as a whole ("unit set") (Table II). We used 31 unique variables, 12 of which were measurements automatically generated by Raven Pro. Raven produces certain "robust" measurements based on signal-to-energy ratios rather than the absolute energy contained within a selection, thus reducing errors introduced by inter-observer differences in measurement accuracy and variation in recording conditions (Table II) (Charif *et al.*, 2010). During measurement, each syllable was assigned a quality score from 0 to 3, with 0 indicating that one or more of the measurements was unreliable (for instance, if a syllable was obscured by a high-amplitude background noise) and 3 indicating full confidence in the measurements. Units were then given corresponding scores from 0 to 3, with 0 indicating that none of the syllables had been reliably measured, and 3 indicating that all syllables had been reliably measured. Only syllables and units that had a score of 3 were included in analyses.

C. Analysis with the full set of variables

We implemented artificial neural networks to predict sex, age class, and individual identity (*nnet* package: Venables and Ripley, 2002; *e1071* package: Meyer *et al.*, 2017) (R Core Team, 2017). We scaled all numerical data and assigned dummy variables to categorical data using the *caret* package (Kuhn *et al.*, 2016) before running each ANN. For each ANN, we first trained the model on the complete dataset and ran the best model detected from the training portion on the complete dataset as well. Then, to account for overfitting we trained the model on a randomly selected 66.7% of the data and ran the ANN on the remaining 33.3%. The *nnet* package automatically implements one hidden layer, which is generally sufficient for most models (de Villiers and Barnard, 1992). We tuned the parameters of the model by testing two to ten neurons in increments of two, weight decays of 0.001, 0.01, 0.05, and 0.1, and initial random weights of 0.5 and 1. Because outcomes will vary dependent on inputs, we ran the networks on randomized subsets 10 times and took the mean of the results (Pozzi *et al.*, 2010).

For sex and age class, we trained and ran ANNs using both complete syllable- and unit-sets of measurements, resulting in eight ANNs across both species. For individual identity, we only included individuals for which we had ≥10 long calls: eight emperor tamarin individuals and seven saddleback individuals. We ran one ANN on each syllable and unit set for each species, resulting in four ANNs. In total, there were 24 networks run with the full variable set: 12 trained and run on 100% of the data, and 12 trained on 67.7% and run on 33.3% to account for overfitting.

D. Analysis with selected features

We ran two additional series of neural networks after implementing each of two methods of feature selection. In the first method, we used the *caret* package (Kuhn *et al.*, 2016) to identify which variables were assigned the highest weights in each ANN. We re-ran each network using only the variables whose weights accounted for ≥50% of variation within all ANNs that were trained and run on subset data.

We also used decision trees to identify which variables were most important in predicting dependent variables. Using

TABLE II. Syllable (A) and unit (B) measurements taken on long calls produced by *Saguinus imperator* and *Leontocebus weddelli*. Parameters marked by † are robust signal measurements automatically generated by Raven Pro. Abbreviations for each variable are in parentheses.

Variable	*Description*
A. Syllable set	
Aggregate entropy (AggE) †	Total amount of disorder (unitless)
Average entropy (AvgE) †	Disorder calculated for each time slice, then averaged over the total sound (unitless)
Bandwidth 90% (BW) †	(Frequency 95%)–(Frequency 5%) (Hz)
Center frequency (CF) †	Frequency at which the energy within the selection is divided equally in two (Hz)
Center time (CT) †	Time at which the energy within the selection is divided equally in two (seconds)
Delta time (DT)	Time between the start and end points of the selection (seconds)
Duration 90% (Dur) †	(Time 95%)–(Time 5%) (seconds)
First/last syllable (F/L)	Designation of initial and final syllables of the unit, excluding introductory and terminal syllables
Introductory/terminal syllable (I/T)	Subunits at the beginning and/or end of a unit that were (a) a markedly different shape than the rest of the unit's syllables; and/or (b) had start frequencies that did not follow the overall arc of the unit (Fig. 3). Not present in all units.
Frequency 5%, frequency 95% (F5, F95) †	Frequencies within a selection at 5% and 95% of the energy in the call (Hz)
Maximum frequency (MF) †	Frequency with the maximum amount of energy in the selection (Hz)
Maximum frequency syllable (MFSyll)	Syllable with the maximum frequency (binary designation)
Maximum power (MP) †	Time of maximum amplitude in the selection (seconds)
Maximum power syllable (MPSyll)	Syllable with maximum power (binary designation)
Subunit number (subunit)	Ordinal numbers (1…N) assigned to each discrete syllable in a unit
Time 5%, Time 95% (T5, T95) †	Time of 5% and 95% of the energy in the selection (seconds)
Year	Year (2014–2016) the recording was made
B. Unit set	
Delivery rate (DelRate)	(Number of subunits)/(Total duration).
Duration average (DurAv)	Mean of duration 90% over all syllables
Number of introductory syllables (IntroNo)	Total number of introductory syllables in the unit
Intersyllable interval average (IntAv)	[(Time 5%)–(Time 95% of preceding syllable)]/(Number of subunits)
Maximum frequency syllable number (MFNo)	Subunit number which contains the maximum frequency
Maximum frequency syllable index (MFSI)	(Maximum frequency syllable)/(Total # of subunits)
Maximum power syllable number (MPNo)	Subunit number of the subunit which contains the maximum power
Maximum power syllable index (MPSI)	(Maximum power syllable)/(Total # of subunits)
Number of subunits (SubNo)	Total number of discrete syllables within each unit
Subunits calculated (SubCalc)	(Number of subunits)–(Introductory and terminal syllables)
Number of terminal syllables (TermNo)	Total number of terminal syllables within the unit
Total duration (TD)	(Time 95% of last syllable)–(Time 5% of first syllable)
Year	Year (2014–2016) the recording was made

the *rpart* package, which automatically uses tenfold cross-validation, we employed the default Gini index for splitting, which uses the probability of a sample being correctly assigned to a dependent variable to compute binary outcomes at each decision point (Therneau *et al.*, 2015). We set the minimum bucket, or the smallest number of samples allowed per terminal node, to the smallest sample size of any dependent variable in the dataset. We set the minimum split, or the minimum number of samples that must be in a node before a split is attempted, to the smallest sample size of any dependent variable + 1. For example, if there were 1428 observations for females and 1292 for males in the dataset, the minimum bucket was set to 1292 and the minimum split was 1293. We re-ran each ANN with only the variables used by the decision trees to categorize dependent variables. For all ANNs attempted after feature selection, we trained and ran each one on two- and one-thirds of the dataset, respectively.

III. RESULTS

We analyzed 205 long calls from 20 *Saguinus imperator* individuals and 199 long calls from 18 *Leontocebus weddelli* individuals (Table I). Emperor tamarin long calls comprise an average of 12 ± 2 SD syllables (average total duration: 3.1 s ± 0.7 SD), while saddleback long calls average 6 ± 2 SD syllables (average total duration: 2.6 s ± 0.7 SD). Total units with the numbers of introductory and terminal syllables are presented in Table III.

A. Identity signaling

If networks were randomly assigning calls to each category, we would expect them to do so accurately 50% of the

TABLE III. Number of units with a given number of introductory ("Intro.") and terminal ("Term.") syllables in *Saguinus imperator* (SIMP) and *Leontocebus weddelli* (LWED) long calls.

		Number of syllables							
Spp.	Type	0	1	2	3	4	5	6	7
SIMP	Intro.	133	67	11	0	0	0	0	0
	Term.	205	5	0	1	0	0	0	0
LWED	Intro.	99	67	24	2	4	1	0	2
	Term.	170	20	5	3	0	0	1	0

time for sex (based on two categories: male and female), 12.5% for emperor tamarin individual identity (based on eight individuals in the sample) and 14.3% for saddleback tamarin individual identity (based on seven individuals). There were no tamarins in the trapped population that fell into the "infant" category, and so individuals were divided into juvenile, subadult, and adult age classes; age would therefore be accurately categorized 33% of the time by chance alone. For the complete dataset, accuracy ranged from 83% to 100% for all dependent variables. Networks using all independent variables that were trained on a subset of the data reached ≥57.7% accuracy for sex, ≥69.1% accuracy for age, and ≥38.8% accuracy for individual (Table IV).

B. Feature selection and predictive accuracy

Variable importance for each ANN, or the weight of each independent variable as determined by the network itself, is summarized in Table V. Between five and eight syllable-set variables (average 5.83 ± 0.41 SD) and four to five unit-set variables (average 4.67 ± 0.52 SD) accounted for the top ≥50% of variation within each network (approximately 25%–30% of all variables per set). For the syllable set, decision trees used 14 variables for emperor tamarins and 10 variables for saddleback tamarins. For the unit set, decision trees used nine and 10 variables for emperor and saddleback tamarins, respectively (Table VI). Four variables in the syllable set (aggregate entropy, average entropy, duration 90%, and year) and three in the unit set (duration average, average intersyllable interval, and year) were identified as being important for predictions by both neural networks and decision trees (Table VII).

Using only the variables that were highly weighted within each species, ANNs predicted dependent variables with ≥68% accuracy for the syllable set and ≥58% for the unit set (Table IV). Using the decision trees' variables within each species, ANNs predicted dependent variables with ≥61% accuracy for both the syllable and unit sets. While accuracy between networks was generally similar, those

TABLE IV. Proportion ± SD of long calls accurately classified by each artificial neural network. Descriptions of syllable and unit variable sets are found in Table II. *Full set* = the network was trained and tested on 100% of the data using all variables; *Subset* = the network was trained on 66.7% of the data and tested on 33.3% using all independent variables; *HWV* = the network was trained on 66.7% of the data and tested on 33.3% of the data using only heavily weighted independent variables cumulatively responsible for ≥50% of the networks' variation, as identified by the neural network itself; *DT* = the network was trained on 66.7% of the data and tested on 33.3% of the data using only independent variables used for categorization of dependent variables via decision tree; N = number of samples used in each network; N^{dv} = number of potential outcomes of each dependent variable, with the expected accuracy based on chance alone in parentheses.

Variables	Data	Sex class	Age class	Individual
Saguinus imperator				
Syllable set	*Full set*	0.908	0.924	0.854
	Subset	0.852 ± 0.005	0.864 ± 0.864	0.755 ± 0.011
	HWV	0.812 ± 0.006	0.848 ± 0.005	0.687 ± 0.012
	DT	0.858 ± 0.003	0.877 ± 0.004	0.772 ± 0.014
	N	2720	2720	1840
Unit set	*Full Set*	0.980	0.976	0.993
	Subset	0.577 ± 0.041	0.691 ± 0.031	0.443 ± 0.028
	HWV	0.587 ± 0.020	0.730 ± 0.008	0.937 ± 0.018
	DT	0.612 ± 0.009	0.692 ± 0.027	0.969 ± 0.021
	N	205	205	146
	N^{dv}	2 (0.500)	3 (0.334)	8 (0.125)
Leontocebus weddelli				
Syllable set	*Full Set*	0.914	0.995	0.830
	Subset	0.709 ± 0.014	0.890 ± 0.008	0.580 ± 0.018
	HWV	0.706 ± 0.009	0.892 ± 0.009	0.790 ± 0.012
	DT	0.712 ± 0.0123	0.901 ± 0.005	0.743 ± 0.010
	N	1249	1249	853
Unit set	*Full set*	0.985	100	0.985
	Subset	0.6296 ± 0.034	0.926 ± 0.011	0.388 ± 0.020
	HWV	0.659 ± 0.027	0.904± 0.013	0.974 ± 0.012
	DT	0.613 ± 0.030	0.900 ± 0.009	0.602 ± 0.020
	N	199	199	136
	N^{dv}	2 (0.500)	3 (0.334)	7 (0.143)

TABLE V. Weight of each independent variable for each artificial neural network (ANN) made using syllable (A) and unit (B) measurement sets for *Saguinus imperator* and *Leontocebus weddelli*. Bolded values in each column represent measurements cumulatively responsible for ≥50% of variable importance within each ANN. See Table II for abbreviations and definitions.

	Saguinus imperator			*Leontocebus weddelli*		
Syllable	Sex	Age	Individual	Sex	Age	Individual
A. Syllable set variable importance (%)						
AggE	3.940326	**5.811053**	4.600082	4.729957	1.9685	4.274391
AvgE	**7.702522**	**5.354826**	5.230837	4.315349	3.891026	3.990533
BW	**5.82068**	4.053813	3.174122	3.322188	4.029787	1.805553
CF	5.179038	**6.162214**	**5.327225**	**9.749726**	4.62074	5.275702
CT	2.953494	3.532069	3.02814	1.154513	2.842708	2.441675
DT	3.125817	3.338087	5.222423	**6.545583**	4.015891	**5.300002**
Dur	4.543524	3.848966	**6.83099**	**8.970941**	**5.294947**	3.856782
F5	4.524552	4.115399	3.703291	**9.537488**	4.177183	3.212521
F95	**6.463829**	3.711039	4.419039	5.060321	4.085764	2.112292
F/L	**6.615663**	**10.75836**	**10.45562**	4.18962	**7.4645**	**7.721333**
I/T	**12.09765**	**10.29655**	**13.14657**	4.141274	**13.1059**	**15.89837**
MF	5.000211	3.040267	1.75548	5.027742	2.83285	4.136517
MFSyll	3.086	3.548599	4.894219	2.144069	**5.891673**	4.841391
MP	5.583967	4.562877	3.596555	**12.60861**	4.869816	**5.723543**
MPSyll	3.408961	4.249824	**5.515104**	2.048812	5.164526	4.584367
Subunit	3.475132	1.810198	2.223307	**6.379561**	**5.970645**	4.322923
T5	1.92806	2.61306	2.807181	2.698322	2.507503	1.914813
T95	1.978916	2.456957	2.445457	3.082674	2.710058	1.745228
Year	**12.57165**	**16.73584**	**11.62435**	4.293	**14.5556**	**16.84207**
B. Unit set variable importance (%)						
DelRate	4.308278	5.843962	**8.84766**	**8.10492**	**10.35755**	6.519645
DurAv	**10.19142**	**10.23763**	**9.68606**	6.49499	**11.42508**	5.725365
IntroNo	**8.479541**	**10.9754**	**8.86561**	6.77214	**8.682071**	11.2381
IntAv	7.30987	**11.60157**	**11.9187**	5.00799	3.498439	5.576761
MFSI	5.210473	7.304105	5.43678	6.3165	6.387437	5.985697
MFNo	7.250412	4.489792	6.90265	3.47786	7.252356	**7.241592**
MPSI	**7.400024**	**7.411284**	4.99587	6.27972	6.316438	3.524178
MPNo	**7.861098**	5.811811	5.66026	4.69093	4.206163	4.814498
SubNo	6.372955	5.360565	5.48788	4.66481	5.934309	5.138093
SubCalc	6.245634	4.97299	6.33531	**9.41189**	6.215909	**7.292357**
TD	3.478281	3.265601	2.57518	4.67049	5.662701	4.148481
TermNo	6.106427	5.618074	8.04462	**10.6895**	**7.645154**	**10.37212**
Year	**19.78558**	**17.10722**	**15.2434**	**23.4183**	**16.41639**	**22.42312**

using heavily weighted variables (HWV) were much more accurate in predicting saddleback individual using the unit set of variables (97.4% accuracy) than decision tree (DT) networks (60.2% accuracy). Networks run on the full variable set were not notably different than those using selected features, with the exceptions of emperor tamarin unit set networks and saddleback syllable set networks for predicting individual identity.

IV. DISCUSSION

A. Identity signaling in callitrichids

In keeping with our predictions that tamarin long calls should be discriminable according to identity signals of their producers, artificial neural networks accurately classified calls according to sex, age, and identity well above what would be predicted by chance alone for both species. Moreover, vocalizations were roughly equally discriminable by identity signals in both species. However, much like prior research on callitrichids, we found that spectrotemporal features responsible for variation were not consistent across emperor and saddleback tamarins. In this study, only seven of 31 features were used by both neural networks *and* decision trees. The spectrotemporal variables were based on entropy features and various duration features (Tables VI–VII). Year was also critical in every network across both species and both measurement sets. The influence of year is a departure from prior studies on call discrimination that have not been able to account for resampling of individuals across years (e.g., Steenbeek and Assink, 1998; Fischer *et al.*, 2002; Ey *et al.*, 2007; Erb *et al.*, 2013), but supports earlier research on captive callitrichids indicating that there is individual continuity in vocalizations within, though not necessarily across, years (Jorgensen and French, 1998; Takahashi *et al.*, 2015). Jorgensen and French (1998) tested only dominant adult Wied's marmosets for changes in vocalizations over time, and while all individuals demonstrated a significant difference in at least one spectrotemporal feature across years, it was not always the same feature(s) in each individual. If these differences were the result of development or senescence, we would expect the same features to change in a similar direction across individuals. Pygmy

TABLE VI. Variables used by each decision tree to predict outcomes using syllable and unit sets of variables (definitions are in TABLE II).

	Saguinus imperator			*Leontocebus weddelli*		
Syllable set	Sex	Age	Individual	Sex	Age	Individual
Aggregate entropy (AggE) †		✓	✓		✓	✓
Average entropy (AvgE) †	✓	✓	✓	✓	✓	✓
Bandwidth 90% (BW) †		✓		✓		✓
Center frequency (CF) †				✓		✓
Center time (CT) †	✓	✓				
Delta time (DT)	✓	✓	✓	✓	✓	✓
Duration 90% (Dur) †		✓	✓	✓	✓	
First/last syllable (F/L)						
Frequency 5% †	✓	✓	✓	✓		✓
Frequency 95% †	✓		✓		✓	✓
Introductory/terminal syllable (I/T)				✓		
Maximum frequency (MF) †		✓				✓
Maximum frequency syllable (MFSyll)						
Maximum power (MP) †	✓	✓		✓	✓	✓
Maximum power syllable (MPSyll)						
Subunit number (subunit)		✓	✓			✓
Time 5% †		✓		✓		
Time 95% †		✓				
Year	✓	✓	✓	✓	✓	✓
Unit set	Sex	Age	Individual	Sex	Age	Individual
Delivery rate (DelRate)		✓	✓	✓		✓
Duration average (DurAv)	✓	✓	✓			
Number of introductory syllables (IntroNo)				✓	✓	
Intersyllable interval average (IntAv)	✓	✓	✓	✓		
Maximum frequency syllable number (MFNo)			✓			✓
Maximum frequency syllable index (MFSI)		✓		✓		
Maximum power syllable number (MPNo)						
Maximum power syllable index (MPSI)	✓	✓	✓	✓		
Number of Subunits (SubNo)		✓				
Subunits calculated (SubCalc)	✓					
Number of terminal syllables (TermNo)						✓
Total duration (TD)	✓			✓		✓
Year	✓		✓			✓

TABLE VII. Comparison of variables in the syllable and unit sets of spectrotemporal measurements of long calls that were selected by the neural network and decision trees. Variables selected by both methods for both species are highlighted. Definitions of variables are in Table II. SIMP = *Saguinus imperator;* LWED = *Leontocebus weddelli*.

	Neural network		Decision tree	
Syllable set	SIMP	LWED	SIMP	LWED
Aggregate entropy (AggE) †	✓	✓	✓	✓
Average entropy (AvgE) †	✓	✓	✓	✓
Bandwidth 90% (BW) †	✓		✓	✓
Center frequency (CF) †	✓	✓		✓
Center time (CT) †			✓	
Delta time (DT)		✓	✓	✓
Duration 90% (Dur) †	✓	✓	✓	✓
First/last syllable (F/L)	✓	✓		
Frequency 5% †		✓	✓	✓
Frequency 95% †	✓		✓	✓
Introductory/terminal syllable (I/T)	✓	✓		
Maximum frequency (MF) †			✓	✓
Maximum frequency syllable (MFSyll)		✓		
Maximum power (MP) †		✓	✓	✓
Maximum power syllable (MPSyll)	✓			
Subunit number (Subunit)		✓	✓	✓
Time 5% †			✓	
Time 95% †			✓	
Year	✓	✓	✓	✓

	Neural network		Decision tree	
Unit set	SIMP	LWED	SIMP	LWED
Delivery rate (DelRate)	✓	✓	✓	
Duration average (DurAv)	✓	✓	✓	✓
Number of introductory syllables (IntroNo)	✓	✓		✓
Intersyllable interval average (IntAv)	✓	✓	✓	✓
Maximum frequency syllable number (MFNo)			✓	✓
Maximum frequency syllable index (MFSI)			✓	✓
Maximum power syllable number (MPNo)	✓	✓		
Maximum power syllable index (MPSI)	✓		✓	✓
Number of subunits (SubNo)			✓	
Subunits calculated (SubCalc)		✓		✓
Number of terminal syllables (TermNo)				✓
Total duration (TD)		✓	✓	✓
Year	✓	✓	✓	✓

(Elowson *et al.*, 1992) and common (Takahashi *et al.*, 2015; Takahashi *et al.*, 2016) marmosets similarly both demonstrated ontogenetic changes in vocal behavior and production that could not be explained by physiology alone: changes in call structure, such as syllables per second, and calling rates, or the ratio of infant-typical to adult-typical calls, were not equal across individuals or litters, with certain parameters increasing in some while decreasing in others. The magnitude and direction of change in parameters should have been even across litters if call structure and rate were shaped solely by development. Further, the rate of development of adult-like spectrotemporal features and call rates accelerated with increased vocal responses from adults (Takahashi *et al.*, 2015; Takahashi *et al.*, 2017).

Variables diagnostic in predicting sex are also inconsistent across species. For instance, Norcross and Newman (1993) found that male golden lion tamarins produced long calls of higher average frequency and shorter duration than females, though Benz *et al.* (1990) concluded that males' calls had higher frequencies and demonstrated no significant difference in duration from females' calls. Alternatively, a study on mustached tamarins revealed that average intersyllable intervals in contact calls were longer for females than males (Masataka, 1987). In a captive study of cotton-top tamarins, Scott *et al.* (2006) demonstrated that, between 1983 and 2002, sex differences in rates of long call production reversed, with females producing long calls more often than males in the first sample but less often than males in the second sample. A second test on the same population in 2004 revealed further changes, with some females once again producing more long calls than males, which the authors suggest may reflect responses to changes in external conditions, that is housing pairs (Scott *et al.*, 2006). Variability in group composition and size across years might similarly account for intraindividual changes that are unrelated to development in this population. Regardless of the mechanisms driving vocal change over time, however, the results presented here suggest that, particularly for researchers seeking to classify signaler characteristics from vocalizations, future models should be careful to account for intraindividual changes in vocalizations on the time axis by including a time-based variable in the set of predictors.

B. Predictive accuracy of neural networks

Surprisingly, neither method of feature selection was more effective than simply training them on a subset of the dataset with all variables. This was true for all outcomes except for in cases of predicting individual identity; this may be the result of smaller sample sizes, though even networks run on the full set of variables nevertheless did markedly better than they would have by chance alone. There is currently no consensus on a formula for determining sample size in neural networks; though lower or imbalanced sample sizes can sometimes mean lower predictive accuracy, small or imbalanced datasets do not necessarily preclude the use of ANNs (Deecke *et al.*, 1999; Mazurowski *et al.*, 2008; Pozzi *et al.*, 2010). Here, for instance, a sample size of 136 saddleback vocalizations from seven individuals was sufficient for 38% classification accuracy (versus the expected 14.3%) by a model using 13 unit-set variables. Though the accuracy of those networks improved with fewer variables, this supports the idea that neural networks can function with small sample sizes, particularly since models run using highly weighted variables reached roughly equal accuracy to those networks run using variables chosen by decision trees despite using about half the number of features. Thus, neural networks appear to be efficient at predicting identity signals despite a high ratio of features to samples.

The results presented here indicate that neural networks are powerful tools for the detection of identity signals in emperor and saddleback tamarins. That predictive accuracy is not meaningfully improved after feature selection might be of particular interest to those attempting to use bioacoustic tools to survey primate populations. Long-distance contact vocalizations can be used for passive acoustic monitoring, wherein microphones are placed in a species' home range and passively detect and record vocalizations without human supervision. This can be effective for surveying species presence, abundance, and health in a variety of ecosystems for populations that are remote, unhabituated, or rare (Campbell *et al.*, 2002; Mennill *et al.*, 2006; Hutto and Stutzman, 2009; Blumstein *et al.*, 2011; Digby *et al.*, 2013; Yack *et al.*, 2013; Heinicke *et al.*, 2015; Bryant *et al.*, 2016; Munger *et al.*, 2016). Reduced restrictions on sample sizes make neural networks good options for demographic monitoring of cryptic or remote groups of species for whom the collection of large datasets can be challenging. Similarly, omitting feature selection as a step in data preparation can drastically increase the efficiency of data processing and analysis for researchers.

ACKNOWLEDGMENTS

This research was supported by the American Society of Primatologists, the International Primatological Society, Field Projects International, and Washington University in St. Louis. We thank the staff at EBLA, the Peruvian Ministry of Agriculture, the Amazon Conservation Association, and all of the research assistants that made this research possible. We also thank K. Chiou for guidance on statistical analyses.

Arnedo, L. F., Mendes, F. D. C., and Strier, K. B. (**2010**). "Sex differences in vocal patterns in the northern muriqui (*Brachyteles hypoxanthus*)," Am. J. Primatol. **72**, 122–128.

Benz, J. J., French, J. A., and Leger, D. W. (**1990**). "Sex differences in vocal structure in a callitrichid primate, *Leontopithecus rosalia*," Am. J. Primatol. **21**, 257–264.

Bioacoustics Research Program (**2014**). *Raven Pro: Interactive Sound Analysis Software (Version 1.5)* (The Cornell Lab of Ornithology, Ithaca, NY), available from http://www.birds.cornell.edu/raven (Last viewed 21 December 2017).

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J. L., Krakauer, A. H., Clark, C., Cortopass, K. A., Hanser, S. F., McCowan, B., Ali, A. M., and Kirschel, A. N. G. (**2011**). "Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus," J. Appl. Ecol. **48**, 758–767.

Bolt, L. M. (**2013**). "The function of howling in the ring–tailed lemur (*Lemur catta*)," Int. J. Primatol. **34**, 157–169.

Bouchet, H., Blois-Heulin, C., and Lemasson, A. (**2012**). "Age- and sex-specific patterns of vocal behavior in De Brazza's monkeys (*Cercopithecus neglectus*)," Am. J. Primatol. **74**, 12–28.

Brown, C. H., Beecher, M. D., Moody, D. B., and Stebbins, W. C. (**1979**). "Locatability of vocal signals in Old World monkeys: Design features for the communication of position," J. Comp. Physiol. Psych. **93**, 806–819.

Bryant, J. V., Brule, A., Wong, M. H. G., Hong, X., Zhou, Z., Han, W., Jeffree, T. E., and Turvey, S. T. (**2016**). "Detection of a new Hainan gibbon (*Nomascus hainanus*) group using acoustic call playback," Int. J. Primatol. **37**, 534–547.

Burkart, J. M., and van Schaik, C. P. (**2009**). "Cognitive consequences of cooperative breeding in primates?," Anim. Cognit. **13**, 1–19.

Byrne, R. W. (**1982**). "Primate vocalisations: Structural and functional approaches to understanding," Behaviour **80**, 241–258.

Campbell, G. S., Gisiner, R. C., Helweg, D. A., and Milette, L. L. (**2002**). "Acoustic identification of female Steller sea lions (*Eumetopias jubatus*)," J. Acoust. Soc. Am. **111**, 2920–2928.

Caselli, C. B., Mennill, D. J., Bicca-Marques, J. C., and Setz, E. Z. F. (**2014**). "Vocal behavior of black-fronted titi monkeys (*Callicebus nigrifrons*): Acoustic properties and behavioral contexts of loud calls," Am. J. Primatol. **76**, 788–800.

Charif, R. A., Strickman, L. M., and Waack, A. M. (**2010**). *Raven Pro 1.4 User's Manual* (The Cornell Lab of Ornithology, Ithaca, NY).

Chen, H. C., Kaplan, G., and Rogers, L. J. (**2009**). "Contact calls of common marmosets (*Callithrix jacchus*): Influence of age of caller on antiphonal calling and other vocal responses," Am. J. Primatol. **71**, 165–170.

Cheney, D. L., Seyfarth, R. M., and Palombit, R. (**1996**). "The function and mechanism underlying baboon 'contact' barks," Anim. Behav. **52**, 507–518.

Cowlishaw, G. (**1996**). "Sexual selection and information content in gibbon songs," Ethology **102**, 272–284.

da Cunha, R. G. T., and Jalles-Filho, E. (**2007**). "The roaring of southern brown howler monkeys (*Alouatta guariba clamitans*) as a mechanism of active defence of borders," Folia Primatol. **78**, 259–271.

Dash, M., and Liu, H. (**1997**). "Feature selection for classification," Intell. Data Anal. **1**, 131–156.

Deecke, V. B., Ford, J. K. B., and Spong, P. (**1999**). "Quantifying complex patterns of bioacoustic variation: Use of a neural network to compare killer whale (*Orcinus orca*) dialects," J. Acoust. Soc. Am. **105**, 2499–2507.

de la Torre, S., and Snowdon, C. T. (**2009**). "Dialects in pygmy marmosets? Population variation in call structure," Am. J. Primatol. **71**, 333–342.

de Villiers, J., and Barnard, E. (**1992**). "Backpropagation neural nets with one and two hidden layers," IEEE Trans. Neural Networks **4**, 136–141.

Digby, A., Towsey, M., Bell, B. D., and Teal, P. D. (**2013**). "A practical comparison of manual and autonomous methods for acoustic monitoring," Methods Ecol. Evol. **4**, 675–683.

Dubreuil, C., Notman, H., and Pavelka, M. S. M. (**2015**). "Sex differences in the use of whinny vocalizations in spider monkeys (*Ateles geoffroyi*)," Int. J. Primatol. **36**, 412–428.

Elowson, A. M., Snowdon, C. T., and Sweet, C. J. (**1992**). "Ontogeny of trill and J-call vocalizations in the pygmy marmoset, *Cebuella pygmaea*," Anim. Behav. **43**, 703–715.

Epple, G. (**1968**). "Comparative studies on vocalization in marmoset monkeys (Hapalidae)," Folia Primatol. **8**, 1–40.

Erb, W. M., Hodges, J. K., and Hammerschmidt, K. (**2013**). "Individual, contextual, and age-related acoustic variation in simakobu (*Simias concolor*) loud calls," PLoS One **8**, e83131.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., and Fischer, J. (**2007**). "Age- and sex-related variations in clear calls of *Papio ursinus*," Int. J. Primatol. **28**, 947–960.

Fischer, J., Hammerschmidt, K., Cheney, D. L., and Seyfarth, R. M. (**2002**). "Acoustic features of male baboon loud calls: Influences of context, age, and individuality," J. Acoust. Soc. Am. **111**, 1465–1474.

Freeberg, T. M., Dunbar, R. I. M., and Ord, T. J. (**2012**). "Social complexity as a proximate and ultimate factor in communicative complexity," Philos. Trans. R. Soc., B **367**, 1785–1801.

Geissmann, T. (**1999**). "Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange," Behaviour **136**, 1005–1039.

Guyon, I., and Elisseeff, A. (**2003**). "An introduction to variable and feature selection," J. Mach. Learn. Res. **3**, 1157–1182.

Heinicke, S., Kalan, A. K., Wagner, O. J. J., Mundry, R., Lukashevich, H., and Kühl, H. S. (**2015**). "Assessing the performance of a semi-automated acoustic monitoring system for primates," Methods Ecol., Evol. **6**, 753–763.

Hutto, R. L., and Stutzman, R. J. (**2009**). "Humans versus autonomous recording units: A comparison of point-count results," J. Field Ornithol. **80**, 387–398.

Jorgensen, D. D., and French, J. A. (**1998**). "Individuality but not stability in marmoset long calls," Ethology **104**, 729–742.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., the R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., and Hunt, T. (**2016**). *Caret: Classification and Regression Training. R Package Version 6.0-73*, https://CRAN.R-project.org/package=caret (Last viewed 2 June 2018).

Lazaro-Perea, C. (**2001**). "Intergroup interactions in wild common marmosets, *Callithrix jacchus*: Territorial defence and assessment of neighbours," Anim. Behav. **62**, 11–21.

Lemasson, A., and Hausberger, M. (**2011**). "Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*)," J. Acoust. Soc. Am. **129**, 3341–3352.

León, J. J., Vargas, S. A., Ramírez, M. A., Galvis, N. F., and Stevenson, P. R. (**2014**). "Vocal communication in woolly monkeys (*Lagothrix lagothricha lugens*) in Cueva de los Guacharos National Park, Colombia," in *The*

Woolly Monkey, edited by T. R. Defler and P. R. Stevenson (Springer, New York), pp. 187–205.

Macedonia, J. M. (**1986**). "Individuality in a contact call of the ringtailed lemur (*Lemur catta*)," Am. J. Primatol. **11**, 163–179.

Marten, K., Quine, D., and Marler, P. (**1977**). "Sound transmission and its significance for animal vocalization. II. Tropical forest habitats," Behav. Ecol. Sociobiol. **2**, 291–302.

Masataka, N. (**1987**). "The perception of sex-specificity in long calls of the tamarin (*Saguinus labiatus labiatus*)," Ethology **76**, 56–64.

Mazurowski, M. A., Habas, P. A., Zurada, J. M., Lo, J. Y., Baker, J. A., and Tourassi, G. D. (**2008**). "Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance," Neural Networks **21**, 427–436.

Mennill, D. J., Burt, J. M., Fristrup, K. M., and Vehrencamp, S. L. (**2006**). "Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest," J. Acoust. Soc. Am. **119**, 2832–2839.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. (**2017**). "e1071: Misc functions of the department of statistics, probability theory group (formerly: E1071) TU Wien," R package version 1.6-8, https://CRAN.R-project.org/package=e1071 (Last viewed 2 June 2018).

Miller, C. T., Mandel, K., and Wang, X. (**2010**). "The communicative content of the common marmoset phee call during antiphonal calling," Am. J. Primatol. **72**, 974–980.

Miller, C. T., and Thomas, A. W. (**2012**). "Individual recognition during bouts of antiphonal calling in common marmosets," J. Comp. Physiol. A. **198**, 337–346.

Mitani, J. C. (**1984**). "The behavioral regulation of monogamy in gibbons (*Hylobates muelleri*)," Behav. Ecol. Sociobiol. **15**, 225–229.

Mitani, J. C. (**1985**). "Sexual selection and adult male orangutan long calls," Anim. Behav. **33**, 272–283.

Mitani, J. C., and Nishida, T. (**1993**). "Contexts and social correlates of long-distance calling by male chimpanzees," Anim. Behav. **45**, 735–746.

Munger, L., Lammers, M. O., Cifuentes, M., Würsig, B., Jefferson, T. A., and Hung, S. K. (**2016**). "Indo-Pacific humpback dolphin occurrence north of Lantau Island, Hong Kong, based on year-round passive acoustic monitoring," J. Acoust. Soc. Am. **140**, 2754–2765.

Norcross, J. L., and Newman, J. (**1993**). "Context and gender-specific differences in the acoustic structure of common marmoset (*Callithrix jacchus*) phee calls," Am. J. Primatol. **30**, 37–54.

Pistorio, A. L., Vintch, B., and Wang, X. (**2006**). "Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*)," J. Acoust. Soc. Am. **120**, 1655–1670.

Pola, Y. V., and Snowdon, C. T. (**1975**). "The vocalizations of pygmy marmosets (*Cebuella pygmaea*)," Anim. Behav. **23**, 826–842.

Potter, J. R., Mellinger, D. K., Clark, C. W., Mellinger, D. K., and Clark, C. W. (**1994**). "Marine mammal call discrimination using artificial neural networks," J. Acoust. Soc. Am. **96**, 1255–1262.

Pozzi, L., Gamba, M., and Giacoma, C. (**2010**). "The use of artificial neural networks to classify primate vocalizations: A pilot study on black lemurs," Am. J. Primatol. **72**, 337–348.

R Core Team. (**2017**). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria), http://www.R-project.org/ (Last viewed 21 December 2017).

Ramanankirahina, R., Joly, M., Scheumann, M., and Zimmermann, E. (**2016**). "The role of acoustic signaling for spacing and group coordination in a nocturnal, pair-living primate, the western woolly lemur (*Avahi occidentalis*)," Am. J. Phys. Anthropol. **159**, 466–477.

Rasoloharijaona, S., Randrianambinina, B., Braune, P., and Zimmermann, E. (**2006**). "Loud calling, spacing, and cohesiveness in a nocturnal primate, the Milne Edwards' sportive lemur (*Lepilemur edwardsi*)," Am. J. Phys. Anthropol. **129**, 591–600.

Reby, D., Lek, S., Dimopoulos, I., Joachim, J., Lauga, J., and Aulagnier, S. (**1997**). "Artificial neural networks as a classification method in the behavioural sciences," Behav. Process. **40**, 35–43.

Ruiz-Miranda, C. R., Kleiman, D. G., Dietz, J. M., Moraes, E., Grativol, A. D., Baker, A. J., and Beck, B. B. (**1999**). "Food transfers in wild and reintroduced golden lion tamarins, *Leontopithecus rosalia*," Am. J. Primatol. **48**, 305–320.

Scott, J., Carlson, K., and Snowdon, C. (**2006**). "Labile sex differences in long calling in cotton-top tamarins," Am. J. Primatol. **68**, 153–160.

Snowdon, C. T., and Hodun, A. (**1981**). "Acoustic adaptation in pygmy marmoset contact calls: Locational cues vary with distances between conspecifics," Behav. Ecol. Sociobiol. **9**, 295–300.

Spehar, S. N., and Di Fiore, A. (**2013**). "Loud calls as a mechanism of social coordination in a fission–fusion taxon, the white-bellied spider monkey (*Ateles belzebuth*)," Behav. Ecol. Sociobiol. **67**, 947–961.

Steenbeek, R., and Assink, P. (**1998**). "Individual differences in long-distance calls of male wild Thomas langurs (*Presbytis thomasi*)," Folia Primatol. **69**, 77–80.

Takahashi, D. Y., Fenley, A. R., and Ghazanfar, A. A. (**2016**). "Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys," Philos. T. R. Soc. B **371**, 20150370.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., and Ghazanfar, A. A. (**2015**). "The developmental dynamics of marmoset monkey vocal production," Science **349**, 734–738.

Takahashi, D. Y., Liao, D. A., and Ghazanfar, A. A. (**2017**). "Vocal learning via social reinforcement by infant marmoset monkeys," Curr. Biol. **27**, 1844–1852.

Terleph, T. A., Malaivijitnond, A., and Reichard, U. H. (**2015**). "Lar gibbon (*Hylobates lar*) great call reveals individual caller identity," Am. J. Primatol. **77**, 811–821.

Therneau, T., Atkinson, B., and Ripley, B. (**2015**). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-10, https://CRAN.R-project.org/package=rpart (Last viewed 2 June 2018).

Tirelli, T., and Pessani, D. (**2011**). "Importance of feature selection in decision-tree and artificial-neural-network ecological applications. *Alburnus alburnus alborella*: A practical example," Ecol. Inform. **6**, 309–315.

Venables, W. N., and Ripley, B. D. (**2002**). *Modern Applied Statistics with S*. (Springer, New York).

Waser, P. M. (**1977**). "Individual recognition, intragroup cohesion, and intergroup spacing: Evidence from sound playback to forest monkeys," Behaviour **60**, 28–74.

Watsa, M. (**2013**). "Growing up tamarin: Morphology, reproduction, and population demography of sympatric free-ranging *Saguinus fuscicollis* and *S. imperator*," Ph.D. dissertation, Washington University in St. Louis, St. Louis, MO.

Watsa, M., Erkenswick, G., Halloran, D., Kane, E. K., Poirier, A., Klonoski, K., Cassalett, S., Maciag, E., Mangalea, M. R., Dinsmore, M. P., McCready, H., Boughan, B. K., Parker, C., Hickmott, A., Bazan, I. E. N., and Zuniga, A. (**2015**). "A field protocol for the capture and release of callitrichids," Neotrop. Primates **22**, 59–68.

Weiss, D. J., Garibaldi, B. T., and Hauser, M. D. (**2001**). "The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus*): Acoustic analyses and playback experiments," J. Comp. Psychol. **115**, 258–271.

White, F. J., Waller, M., Boose, K., Merrill, M. Y., and Wood, K. D. (**2015**). "Function of loud calls in wild bonobos," J. Anthropol. Sci. **93**, 89–101.

Windfelder, T. L. (**2001**). "Interspecific communication in mixed-species groups of tamarins: Evidence from playback experiments," Anim. Behav. **61**, 1193–1201.

Wu, D. (**2009**). "Supplier selection: A hybrid model using DEA, decision tree and neural network," Expert Syst. Appl. **36**, 9105–9112.

Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., and Coates, S. (**2013**). "Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat," J. Acoust. Soc. Am. **134**, 2589–2595.

Yang, Y., and Pederson, J. O. (**1997**). "A comparative study on feature selection in text categorization," in *Proceedings of the Fourteenth Annual International Conference on Machine Learning*, pp. 412–420.