

# Vocal repertoire of free-ranging adult golden snub-nosed monkeys (*Rhinopithecus roxellana*)

Penglai Fan<sup>1</sup>  | Xuecong Liu<sup>2</sup>  | Ruoshuang Liu<sup>3</sup> | Fang Li<sup>2</sup>  |  
Tianpeng Huang<sup>4</sup> | Feng Wu<sup>4</sup> | Hui Yao<sup>4</sup> | Dingzhen Liu<sup>1</sup> 

<sup>1</sup>Institute of Ecology, College of Life Sciences, Beijing Normal University, Beijing, China

<sup>2</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

<sup>3</sup>School of Environment, Beijing Normal University, Beijing, China

<sup>4</sup>Shennongjia National Park, Shennongjia, Hubei Province, China

## Correspondence

Dingzhen Liu, Institute of Ecology, College of Life Sciences, Beijing Normal University, #19 Xijiekouwai Street, Haidian District, Beijing 100875, China.

Email: dzliu@bnu.edu.cn

Xuecong Liu, College of Life Sciences, University of Chinese Academy of Sciences, 19A Yuquan Road, Shijingshan District, Beijing 100049, China.

Email: xuecongliu@ucas.ac.cn

## Funding information

National Key R&D Program of China, Grant number: 2016YFC0503202; Scientific Research Grant for Youth Scholars and the Special Grant for Research and Education Integration, Grant number: KJRH2015-016; Hubei Provincial Key Laboratory for Conservation Biology of Snub-nosed Monkeys

Vocal signaling represents a primary mode of communication for most nonhuman primates. A quantitative description of the vocal repertoire is a critical step in in-depth studies of the vocal communication of particular species, and provides the foundation for comparative studies to investigate the selective pressures in the evolution of vocal communication systems. The present study was the first attempt to establish the vocal repertoire of free-ranging adult golden snub-nosed monkeys (*Rhinopithecus roxellana*) based on quantitative methods. During 8 months in Shennongjia National Park, China, we digitally recorded the vocalizations of adult individuals from a provisioned, free-ranging group of *R. roxellana* across a variety of social-ecological contexts. We identified 18 call types, which were easily distinguishable by ear, visual inspection of spectrograms, and quantitative analysis of acoustic parameters measured from recording samples. We found a great sexual asymmetry in the vocal repertoire size (females produced many more call types than males), likely due to the sex differences in body size and social role. We found a variety of call types that occurred during various forms of agonistic and affiliative interactions at close range. We made inference about the functions of particular call types based on the contexts in which they were produced. Studies on the vocal communication in *R. roxellana* are particularly valuable since they provide a case about how nonhuman primates, inhabiting forest habitats and forming complex social systems, use their vocalizations to interact with their social and ecological environments.

## KEYWORDS

complex social system, *Rhinopithecus roxellana*, sexual asymmetry, vocal repertoire

## 1 | INTRODUCTION

Vocal signaling represents a primary mode of communication for nonhuman primates, especially for those living in habitats with limited

visibility (Altmann, 1967). Studies of vocal behavior can reveal important aspects of how animals use their vocalizations to interact with their social and ecological environments. A detailed quantitative description of the vocal repertoire is a critical step in subsequent

Penglai Fan and Xuecong Liu contributed equally.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *American Journal of Primatology* Published by Wiley Periodicals, Inc.

in-depth studies of the vocal communication of particular species, and provides the foundation for comparative studies among populations, species, and taxa to investigate the selective pressures in the evolution of vocal communication systems (Bouchet, Blois-Heulin, & Lemasson, 2013; Ey & Fischer, 2009; Hauser, 1993). In addition, studies of the vocal repertoire can provide common terminology and referents, helping avoid confusion among different studies.

The structure and size of a vocal repertoire appear to be influenced by the factors associated with the characteristics of habitat and social life (Bouchet et al., 2013; Ey & Fischer, 2009). First, the characteristics of the local habitat affect sound propagation (Waser & Brown, 1986). In order that the information contained in vocalizations is effectively transmitted, the physical properties of vocal signals, especially those used over long distances, are expected to be designed to optimize propagation in the environment (the acoustic adaptation hypothesis: Ey & Fischer, 2009; Morton, 1975). Second, it has been argued that vocal repertoires can be distinguished as graded or discrete based on the type of habitat and conspecific interactions (Hammerschmidt & Fischer, 1998; Marler, 1976). Specifically, discrete signal systems without intermediates between call types should evolve if animals live in closed habitats like dense forests or vocalizations function across long distances in order to reduce misunderstanding, because vocal signals must operate without complementary visual or contextual cues in these situations. In contrast, graded signal systems with continuous acoustic variation between call types should be favored when animals inhabit open habitats like savannahs or interact with conspecifics at close range. It has been suggested that the acoustic variation in the vocal signals used during "face-to-face" interactions is more likely to reflect the motivational states of the callers rather than habitat characteristics (the motivation-structural rules: Morton, 1977). Third, social complexity has been hypothesized to co-evolve with vocal complexity (the social complexity hypothesis: Bouchet et al., 2013; Gustison, le Roux, & Bergman, 2012). Animals with more complex social systems, such as those consisting of more interacting individuals, more diverse interactions, and/or more social structural levels, are expected to need more diverse signals to convey a wider range of information about individual identity, behavioral and environmental contexts, and/or emotional and motivational states.

Within a vocal repertoire, the production and use of vocal signals are affected by several factors including sex, age, body size, and social role (Bouchet, Blois-Heulin, & Lemasson, 2012; Bouchet, Pellier, Blois-Heulin, & Lemasson, 2010; Pfefferle & Fischer, 2006). The effects of these factors may not be mutually exclusive. For example, sex-specific discrepancies in terms of call production are common in nonhuman primates (Bouchet et al., 2010; Briseño-Jaramillo, Biquand, Estrada, & Lemasson, 2017; Hohmann, 1991). These differences are often attributed to the sex differences in social role, and in some species, the sex that is more social produces more call types than the other (Bouchet et al., 2010, 2012). In addition, some acoustic parameters (e.g., fundamental frequency and formant dispersion) depend on the size of vocal tract/folds, which is positively related to body size (Hauser, 1993; Pfefferle & Fischer, 2006). Thus, in species with sexual dimorphism, sex differences in call production may also be attributed to the sex differences in body size. For example, it has been suggested that the fundamental frequency is negatively

correlated with body size (Pfefferle & Fischer, 2006), and thus the larger body size may limit one sex (usually males) to produce some high pitched calls that the other sex can emit (Bouchet et al., 2012).

The golden snub-nosed monkey (*Rhinopithecus roxellana*), a colobine endemic to China, inhabits temperate forests in mountainous areas at high altitudes of 1,000–4,100 m (Kirkpatrick & Grueter, 2010; Li, Pan, & Oxnard, 2002). It exhibits pronounced sexual dimorphism in body size; body weights are approximately 15.0 and 9.5 kg for adult males and adult females in captivity, respectively (Davison, 1982; Jablonski & Pan, 1995). This primate is primarily arboreal, but sometimes descends to the ground for foraging (Li, 2007). It is well known for its multilevel social structure in which several one-male multi-female units (OMUs) and one (occasionally more than one) all-male unit (AMU) form a large cohesive group up to several hundred individuals (Qi et al., 2014). The social units of a group maintain a close association and coordinate their activities, while each unit is spatially and socially distinct: the individuals of the same unit usually stay much closer to each other than to those of other units, and most social interactions occur among the individuals within units (Wang et al., 2013; Zhang, Li, Qi, MacIntosh, & Watanabe, 2012).

Vocal signals should be a particularly important tool for communication in *R. roxellana*, living in forests with poor visibility, and further, its complex social system makes vocal research more interesting. However, there have been only few studies focusing on the vocal communication of this primate, and all existing studies are very preliminary, in which limited numbers of vocalizations were recorded and subjectively classified. Tenaza, Fitch, and Lindburg (1988) reported four kinds of vocalizations that frequently occurred, that is, *shrills*, *bawls*, *chucks*, and *whines*, as well as some other vocalizations emitted infrequently, for example, a variety of *grunts*, from four bisexual pairs of adult *R. roxellana* in captivity. The authors also stated that *shrills* and *bawls* actually comprised a variety of forms of vocalizations. Li, Chen, Luo, and Xie (1993) presented five major categories of vocalizations from the Qinling population of wild *R. roxellana*, and termed these vocalizations based on contextual information, that is, *amazement calls*, *alarm calls*, *warning calls*, *peaceful calls*, and *contacting calls*. Ren et al. (2000) recorded the vocalizations from the Shennongjia population of wild *R. roxellana*, but they did not conduct any acoustic analysis and classified these vocalizations into four broad categories only by human auditory sense.

The purpose of the present study was to establish the vocal repertoire of free-ranging adult *R. roxellana* by determining call types based on quantitative analyses of acoustic properties. In addition, the social-ecological contexts were recorded for each call type. We predicted that *R. roxellana* would have a large vocal repertoire because of the need to mediate various social interactions within a complex multilevel society. The results will be helpful to understand the vocal communication of *R. roxellana* in addition to accumulate literature.

## 2 | METHODS

### 2.1 | Study site and subjects

We carried out this study in a provisioned, free-ranging group of *R. roxellana* in the Dalongtan area of Shennongjia National Park, Hubei

Province, China (Dalongtan Conservation Station: 31°29'65"N, 110°17'93"E, 2,170 m) (Yao et al., 2011). The topography within this area is extremely rugged with an elevational range of 2,000–2,700 m. The climate is strongly seasonal. The average monthly temperature is highest in July (ca. 17.1 °C) and lowest in January (ca. –3.5 °C). The annual rainfall is approximately 1,800 mm with the rainy season from July to September. Snowfalls last from November to March. The vegetation is characterized by deciduous broadleaf and evergreen conifer mixed forests.

To facilitate ecotourism and research, the study group has been provisioned and habituated since January 2006. The animals are provisioned two or three times per day with lichens, pine seeds, apples, carrots, oranges, and peaches (Yao et al., 2011), among which lichens and pine seeds are their most important natural foods (Liu, Stanford, Yang, Yao, & Li, 2013). When not provisioned, they forage freely within the area of approximately 9 km<sup>2</sup> around the provisioning site. Close proximity (0.5–10 m) allowed us to identify all individuals except infants, including their age/sex classes and unit memberships, based on their physical characteristics such as body size, hair coloration, scar, genitalia, face shape, and canines (Yao et al., 2011). The unit memberships of infants could also be determined based on their maternal dependence. During the study period, the group contained 7 adult males (≥7 years old), 19 adult females (≥5 years old), and 43–45 juveniles (males: 1.5–7 years old; females: 1.5–5 years old) and infants (≤1.5 years old) of both sexes, forming five OMUs and one AMU. Specifically, there were one adult male and two to six adult females in each OMU, and two adult males in the AMU. These adult individuals were selected as our study subjects. Juveniles and infants were excluded from this study due to the potential influences of developmental factors on vocalizations (Snowdon & Elowson, 2001) and their lesser importance in the maintenance of social structure.

The protocol of this study was approved by the Animal Care Committee of the Beijing Normal University, and conformed to the regulatory requirements of Shennongjia National Park and adhered to the American Society of Primatologists Principles for the Ethical Treatment of Primates.

## 2.2 | Vocalization recording and acoustic analyses

We recorded vocalizations from April to November 2016 using a combination of focal and ad libitum sampling primarily owing to the special multilevel social structure of the monkeys. Specifically, we first selected one social unit as our focal fellow on an observation day (08:00–18:00), and then rotated to another on next day. Attempts were made to rotate observations evenly among the six social units. For each focal unit per day, vocalizations of adult individuals were recorded ad libitum outside of the provisioning times and when there were not excessive human disturbances. Occasionally, calls of adult individuals from non-focal units were also recorded opportunistically to increase the total amount of vocalization samples. Vocalizations were recorded using a Tascam DR44-WL digital recorder at a 44.1 kHz (16 bits) sampling rate, connected to a Sennheiser ME66 directional microphone. The vocalization data were uploaded to a laptop computer for storage and analysis.

Before acoustic analyses, we excluded the recordings from unidentified callers. We generated narrow-band spectrograms for

the selected recordings using Praat 5.3.72 package (Gaussian window shape, view range = 0–20 kHz, window length = 0.03 s, dynamic range = 50 dB) (Boersma & Weenink, University of Amsterdam, the Netherlands). We pre-classified vocalizations based on auditory sense and visual inspection of spectrograms. For the designation of call types, we used descriptive terms that represented the characteristic properties of spectrograms, with exceptions in which the calls occurred only in the mating context and then we used terms implying functional significance. We attempted to use the same terms for the call types identified in the study of captive *R. roxellana* (Tenaza et al., 1988).

For further acoustic analyses, we excluded the recordings with excessive background noise such as water and bird sounds, and those overlapped by other calls. The vocalizations in which the recording distances were >10 m and where the orientations of the callers were opposite to the recording equipment were further excluded to minimize the effect of signal degradation. According to many previous studies of other primates, the distances of ≤10 m should be an appropriate cutoff to obtain high quality recordings and measure acoustic parameters, especially in terms of those related to intensity (e.g., *Macaca sylvanus*: Hammerschmidt & Fischer, 1998; *Papio papio*: Maciej, Ndao, Hammerschmidt, & Fischer, 2013; *Gorilla gorilla*: Salmi & Doran-Sheehy, 2014; *Mandrillus sphinx*: Levréro et al., 2015). For each selected recording, we used Praat to measure 16 temporal, spectral, and intensity parameters: duration, mean  $f_0$ , SD  $f_0$ , start  $f_0$ , end  $f_0$ , min  $f_0$ , max  $f_0$ , range  $f_0$ , %T\_min  $f_0$ , %T\_max  $f_0$ , meanAMP, minAMP, maxAMP, rangeAMP, HNR, and Jitter (parameter definitions and extraction/calculation methods: Table 1) (Charlton, Zhihe, & Snyder, 2009a, 2009b). If vocalizations were uttered in bouts, we considered each call separately for analysis.

## 2.3 | Context observations

For each recorded vocalization, we noted the concurrent contextual information by speaking into a lapel microphone using the second audio channel of the recording equipment, complementarily by videotaping with a Sony Digital Camera (HDR-XR 260) (by J. Yang, one of field assistants). Such information included date, time, the caller, its unit membership and behavior, its distance away from and orientation relative to the recording equipment, and if possible the potential receiver, its unit membership and behavioral response, and the external event that could potentially elicit the call emission. A vocalization was considered to be spontaneous or initial if it was not preceded by other calls within 5 s, and a behavioral response was recorded if it occurred within 5 s from an initial vocalization. All behavioral contexts, live and from videotape, were coded by P. Fan, the same person who recorded and pre-classified vocalizations. If possible, we calculated the proportion of recorded vocalizations by context and callers' sex per call type. Vocalization contexts were classified into several broad social-ecological categories, including traveling, foraging, feeding, resting, greeting/responding, agonistic interactions, affiliative interactions, mating, and environmental disturbances (detailed descriptions of contexts: Table 2).

**TABLE 1** Definitions of acoustic parameters measured from the vocalizations of free-ranging adult *R. roxellana*

Parameters	Definitions (units)
Duration	Duration of the entire call(s)
Mean $f_0$	Mean frequency of the fundamental frequency contour (Hz)
SD $f_0$	Standard deviation of frequency values of the fundamental frequency contour (Hz)
Start $f_0$	Frequency at the start of the fundamental frequency contour (Hz)
End $f_0$	Frequency at the end of the fundamental frequency contour (Hz)
Min $f_0$	Minimum frequency of the fundamental frequency contour (Hz)
Max $f_0$	Maximum frequency of the fundamental frequency contour (Hz)
Range $f_0$	Range of the fundamental frequency (Hz)
%T_min $f_0$	Percentage of the duration from start $f_0$ to min $f_0$ out of the entire fundamental frequency contour (%)
%T_max $f_0$	Percentage of the duration from start $f_0$ to max $f_0$ out of the entire fundamental frequency contour (%)
MeanAMP	Mean intensity (amplitude) of the entire call (dB)
MinAMP	Minimum intensity (amplitude) of the entire call (dB)
MaxAMP	Maximum intensity (amplitude) of the entire call (dB)
RangeAMP	Range of intensity (amplitude) of the entire call (dB)
HNR	Harmonics to noise ratio: periodic distribution of energy within the call (dB)
Jitter	Cycle-to-cycle variability in $f_0$ frequency across the call (%)

Praat was used; (Sound: To Pitch [cc]) command for the extraction of duration, mean  $f_0$ , SD  $f_0$ , start  $f_0$ , end  $f_0$ , min  $f_0$ , max  $f_0$ ; (Sound: To Intensity) command for the extraction of meanAMP, minAMP, maxAMP; (To Harmonicity) command for the extraction of HNR; (Jitter [local]) command for the extraction of Jitter; For %T\_min  $f_0$  and %T\_max  $f_0$ , the durations from start  $f_0$  to min  $f_0$  and max  $f_0$  were first extracted using (Sound: To Pitch [cc]) command and divided by the duration of the entire call, respectively; Range  $f_0$  and rangeAMP were directly calculated by max  $f_0$  - min  $f_0$  and maxAMP - minAMP, respectively.

## 2.4 | Statistical analyses

We conducted a direct discriminant function analysis (DFA) to examine whether our pre-classified call types were acoustically distinct. DFA identifies the quantitative predictor variables that best describe the differences among groups (Klecka, 1980). Based on the discriminant functions combined by these variables, the procedure assigns each vocalization to its appropriate group (correct) or another (incorrect). For external validation, we used the *leave-one-out* classification method, in which each case was classified by the functions derived from all cases except that one. Because the data set was unbalanced, classification coefficients were adjusted according to the observed group sizes. Six highly correlated parameters (Spearman's test:  $r > 0.4$ ) were excluded from DFA, including mean  $f_0$ , SD  $f_0$ , end  $f_0$ , max  $f_0$ , minAMP, and

maxAMP. The call types with sample sizes  $< 10$  were subsequently excluded because the number of recordings for each call type must be larger than the number of parameters used in DFA (Tabachnick & Fidell, 2001). Some parameters were log (duration, min  $f_0$ , start  $f_0$ , Jitter) or square-root (%T-min  $f_0$ , meanAMP, rangeAMP, HNR) transformed, and then normal distributions of all parameters per call type used for DFA were confirmed by the examination of Q-Q plots and the Kolmogorov-Smirnov tests ( $p > 0.05$ ). Although the covariances were unequal across call types (Box's test:  $p < 0.05$ ) and the variances were unequal for some parameters (Levene's test:  $p < 0.05$ ), it has been shown that DFA is robust to the violation of this homoscedasticity assumption (Klecka, 1980; Lachenbruch, 1975). All statistical analyses were conducted using SPSS 21.0.

## 3 | RESULTS

### 3.1 | Classification of call types

In total, we obtained 1,826 vocalization samples with identified callers from adult animals (659 from males, 1,167 from females) in 650 hr over 105 days during the study period (Table 3). We pre-classified these vocalizations into 17 call types, which were easily distinguishable by human ear and visual inspection of their spectrograms (Figure 1). We also noted another call type, which sounded apparently different from any other one and was uttered by adult females during sexual solicitations. However, we were not able to generate spectrograms for it due to the extremely low intensity. We termed this call type as *female sexual solicitation calls* and included it here for completeness. In summary, we identified 18 different call types in the vocal repertoire of free-ranging adult *R. roxellana*.

There were 311 recording samples that were appropriate for the measurement of acoustic parameters (Tables 3 and 4). The contribution per individual to the data set varied among call types. DFA correctly classified 85.8% of the call samples (*shrills*, *long grunts*, *squeals*, *bawls*, *compound squeaks* not included due to small sample sizes). The correct assignment rate of cross validation analysis was 79.9%, better than expected by chance (Chi-square test:  $X^2 = 51.54$ ,  $df = 11$ ,  $p < 0.001$ ), indicating that our pre-classification of call types was appropriate (correct rate per call type: Table 3). DFA generated three canonical discriminant functions that had eigenvalues  $> 1$  (function 1: 14.8; function 2: 3.8; function 3: 1.4) and explained 91.1% of the variance cumulatively. Function 1 was primarily correlated duration and explained 67.5% of the variance. Function 2 was most strongly associated with range  $f_0$  and explained 17.4% of the variance. Function 3 explained 6.2% of the variance and was mainly related to meanAMP and Jitter. *Rattles* (50.0%) and *compound squeals* (47.7%) had the lowest classification rates. *Rattles* were most often misclassified as *three syllabled barks* (18.8%), and *compound squeals* as *modulated tonal screams* (21.1%), reflecting the acoustic similarities between the original and respective misclassified call types in the most significant parameters (Table 4).

**TABLE 2** Context description and number (percentage) of recorded vocalizations by context and callers' sex per call type from the vocal repertoire of free-ranging adult *R. roxellana*

Call type	Context category	Context description	Number (%) of vocalizations		
			Males	Females	Both sexes
Coo	Traveling	Unit/group ranging directionally	184 (51)	99 (57)	283 (53)
	Foraging	Moving around to search for food	86 (24)	47 (27)	133 (25)
	Resting	Not involved in location changes or activities e.g., feeding, mating, social interactions	68 (19)	22 (13)	90 (17)
Shrill	Environmental disturbances	Human voices	22 (6)	5 (3)	27 (5)
	Greeting/responding	Greeting to the adult male of the same OMU rejoining		45 (36)	
	Greeting/responding	OMU members greeting to each other after waking up from sleep		42 (34)	
Whine	Feeding	OMU members chewing/manipulating food items using mouth or hand(s) together		37 (30)	
	Feeding	Chewing/manipulating food items using mouth or hand(s)	29 (39)	3 (100)	32 (41)
	Greeting/responding	Responding to shrills from adult females of the same OMU	23 (31)		23 (30)
Long grunt	Greeting/responding	Responding to distress calls from juveniles of the same OMU	22 (30)		22 (29)
	Agonistic interactions	Mild threats with gazes toward individuals of the same OMU	4 (100)	16 (100)	20 (100)
	Agonistic interactions	Moderate threats with neck stretching and facial threat expressions toward individuals of the same and different units	91 (82)	140 (100)	231 (92)
Growl	Environmental disturbances	Presence of humans	20 (18)		20 (8)
	Agonistic interactions	Intense threats with mobbing behaviors (e.g., standing on foot) toward adult females of the same and different OMUs		94 (100)	
	Agonistic interactions	Intense threats with physical contact (e.g., grasping, biting) toward individuals of the same and different units	14 (100)	22 (100)	36 (100)
Rattle	Agonistic interactions	Suddenly attacked by the adult male from a different OMU		45 (100)	
	Agonistic interactions	Intense agonistic interactions between the adult male of the same OMU and those of different OMUs		135 (100)	
	Agonistic interactions	Calling toward own infants during male-male intense agonistic interactions and other unsafe situations		30 (100)	
Squeak	Agonistic interactions	Submitted with behaviors (e.g., crouching, avoiding) to moderate threats from individuals of the same OMU		58 (100)	
	Affiliative interactions	Reconciliation after intra-OMU female-female and inter-OMU male-male intense agonistic interactions	2 (100)	6 (100)	8 (100)
	Affiliative interactions	Reconciliation after intra-OMU female-female intense agonistic interactions		28 (100)	
Modulated tonal scream	Traveling	Lagging behind during unit/group traveling		16 (73)	
	Greeting/responding	Adult females located at group periphery detecting and calling toward the AMU		6 (27)	
	Environmental disturbances	Presence of humans with food supply		42 (100)	
Chuck	Environmental disturbances	Presence of snakes	9 (39)	4 (2)	13 (5)
	Environmental disturbances	Presence of humans	6 (26)	30 (13)	36 (14)
	Environmental disturbances	Breaking sounds of branches		40 (17)	40 (16)
Mounting grunt	Environmental disturbances	Uncertain disturbances	8 (35)	155 (68)	163 (65)
	Mating	During the period immediately before dismounting	71 (100)		

OMU, one-male multi-female unit; AMU, all male unit.

**TABLE 3** Numbers (N) of all recorded vocalizations and the samples used to measure acoustic parameters, and correct classification rates based on discriminant function analysis (DFA) per call type from the vocal repertoire of free-ranging adult *R. roxellana*

Call type	N of all vocalizations									DFA Correct rate (%)
	N of vocalizations		N of individuals		N of vocalizations per individual (range)		N of vocalizations used to measure acoustic parameters			
	Males	Females	Males	Females	Males	Females	N of vocalizations	N of individuals	N of vocalizations per individuals (range)	
Coo	360	173	7	11	31-94	4-31	63	14	1-12	92.1
Shrill*		124		9		5-36	9	5	1-3	
Whine	74	3	5	2	3-21	1-2	20	4	2-8	90.0
Long grunt*	4	16	1	4	4	2-6	7	3	2-3	
Grunt	111	140	5	6	9-44	4-81	41	7	1-14	100.0
Growl		94		13		2-15	15	7	1-4	100.0
Bark	14	22	3	6	3-7	1-7	11	7	1-3	72.7
Rattle		45		9		2-10	16	11	1-4	50.0
Squeak		135		9		6-28	11	4	1-5	81.8
Squeal*		30		6		2-10	9	4	1-3	
Compound squeal		58		8		2-18	19	6	1-9	47.7
Bawl*	2	6	1	3	2	1-3	3	3	1	
Compound squeak*		28		6		3-12	9	3	1-6	
Three syllabled bark		22		3		2-17	13	2	1-12	100.0
Modulated tonal scream		42		10		1-22	21	10	1-8	66.7
Chuck	23	229	2	16	7-16	3-33	32	14	1-10	93.8
Mounting grunt	71		3		13-21		12	3	1-7	100.0
OVERALL	659	1,167					311			85.8

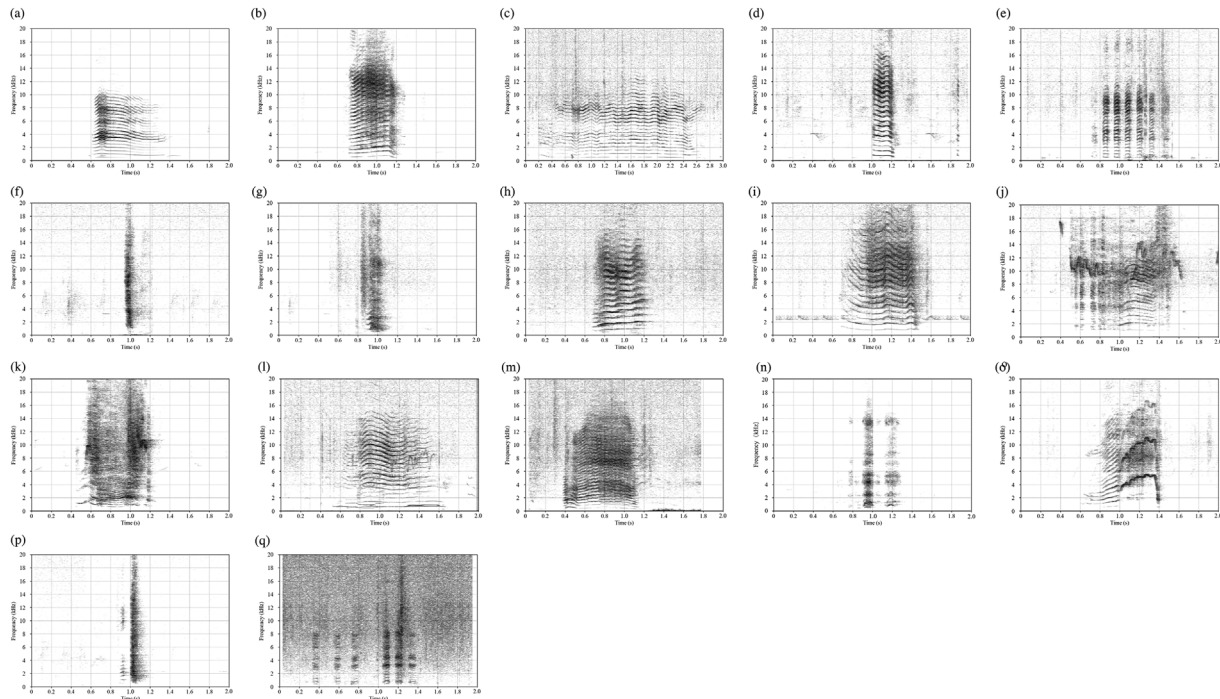
\*Not included in DFA due to small sample sizes.

### 3.2 | Description of call types

Some call types occurred in multiple contexts, especially *coos*, *shrills*, *whines*, and *chucks*, while the others were produced in single contexts, particularly those associated with social interactions at close range and mating contexts (Table 2). Most of all identified call types (10 of 18) occurred during various forms of agonistic and affiliative interactions at close range. In addition, there were sex differences in call production. Specifically, females produced more call types than males, that is, there were more female-specific than male-specific call types (female-specific: 10 call types; male-specific: 1 call type). Most of female-specific call types are produced during social interactions at close range. Below were detailed descriptions of acoustic structure and contexts for each call type.

#### 3.2.1 | Coo

Coos are tonal and characterized by a relatively long duration ( $0.734 \pm 0.194$  s), a low  $f_0$  (mean  $f_0$ :  $545 \pm 96$  Hz), and rich harmonics with a general trend of slow decrease in frequency. Almost all harmonics are below 10 kHz and thus the frequency range is narrower than those of most other call types. Coos were the vocalizations frequently heard from all adult individuals during unit/group traveling (53%), foraging (25%), and resting (17%). In these contexts, coos were often responded to by the same vocalizations, whereas the individuals who responded vocally and the social units they belonged to could not be determined in most cases due to the widely dispersed distribution of the animals. Based on the cases when we could do so, coos emitted by an individual from an OMU were responded by others from both the same and other OMUs, and coos uttered by an individual from the AMU were responded



**FIGURE 1** Representative spectrograms of vocalizations from free-ranging adult *R. roxellana*. (a) *Coo*, (b) *Shrill*, (c) *Whine*, (d) *Long grunt*, (e) *Grunt bout*, (f) *Growl*, (g) *Bark*, (h) *Rattle*, (i) *Squeak*, (j) *Squeal*, (k) *Compound squeal*, (l) *Bawl*, (m) *Compound squeak*, (n) *Three syllabled bark*, (o) *Modulated tonal scream*, (p) *Chuck*, and (q) *Mounting grunt bout*

to by another member from the AMU. The exchanges of *coos* among individuals between OMUs and the AMU were not observed. In addition, *coos* were also observed to be emitted by adult individuals from all social units in response to the voices of staff (5%).

### 3.2.2 | Shrill

*Shrills* are basically tonal and comprise abundant harmonics superimposed by slight noisy elements. These vocalizations begin with a slow increase in frequency continuously until the end with a slight decrease. *Shrills*, with a medium duration ( $0.614 \pm 0.169$  s) and a relatively low  $f_0$  (mean  $f_0$ :  $650 \pm 100$  Hz), were uttered by adult females when the adult males of their own OMUs rejoined (36%), and when they woke up from sleep at noon (34%) and were feeding together (30%). Adult females of an OMU usually emitted *shrills* in a high degree of synchronization, and the chorus was sometimes ended by a *whine* from the adult male of the same OMU.

### 3.2.3 | Whine

*Whines* have a clear harmonic structure characterized by a longest duration ( $1.567 \pm 0.404$  s), a highest HNR ( $10.543 \pm 3.652$  dB), and a relatively low  $f_0$  (mean  $f_0$ :  $490 \pm 44$  Hz) with stable slight vibrations (SD  $f_0$ :  $58 \pm 33$  Hz; range  $f_0$ :  $267 \pm 124$  Hz). As with *coos*, most harmonics are below 10 kHz and the frequency range is narrower than those of most other call types. Adult individuals of both sexes from all social units uttered *whines* spontaneously while feeding (41%). The adult male from an OMU was also observed to emit *whines* in response to

*shrills* of adult females from the same OMU (30%), and the calls, likely expressing anxiety, of juveniles from the same OMU (29%).

### 3.2.4 | Long grunt

*Long grunts* are tonal and rich in harmonics with few frequency modulations. The duration is relatively short ( $0.306 \pm 0.079$  s) and the  $f_0$  is the lowest (mean  $f_0$ :  $453 \pm 195$  Hz). These vocalizations accompanied by threatening gazes were emitted by both sexes from OMUs during intra-unit mild agonistic interactions. When juveniles were threatened by *long grunts* of adults, and adult females were threatened by adult males, they usually interrupted their ongoing behaviors and moved away. When an adult female was threatened by another, she avoided or resisted, which usually led to the escalation of agonistic interactions.

### 3.2.5 | Grunt

Sounding like *long grunts*, *grunts* are also tonal and comprise abundant harmonics with few frequency modulations. However, *grunts* have a higher  $f_0$  than *long grunts* (mean  $f_0$ :  $549 \pm 126$  Hz) and a shorter duration than *long grunts* and any other call type ( $0.092 \pm 0.013$  s). *Grunts* occurred in bouts with regular intervals (range = 2–8, median = 5), which could last up to more than 1.3 s. These vocalizations, accompanied by stretching of the neck and facial expressions of threat such as glaring, were emitted by both sexes mainly during various forms of intra-unit and inter-unit moderate agonistic interactions (92%). The reactions of receivers varied and depended on the specific situations of the interactions: when juveniles were threatened by

**TABLE 4** Physical characteristics (mean  $\pm$  SD) of vocalizations per call type from the vocal repertoire of free-ranging adult *R. roxellana*

Call type	N	Duration	Mean $f_0$	SD $f_0$	Start $f_0$	End $f_0$	Min $f_0$	Max $f_0$	Range $f_0$
Coo	63	0.734 $\pm$ 0.194	545 $\pm$ 96	97 $\pm$ 69	631 $\pm$ 141	586 $\pm$ 205	379 $\pm$ 120	726 $\pm$ 188	347 $\pm$ 223
Shrill	9	0.614 $\pm$ 0.169	650 $\pm$ 100	149 $\pm$ 117	661 $\pm$ 156	552 $\pm$ 150	408 $\pm$ 129	861 $\pm$ 296	453 $\pm$ 350
Whine	20	1.567 $\pm$ 0.404	490 $\pm$ 44	58 $\pm$ 33	475 $\pm$ 155	438 $\pm$ 113	346 $\pm$ 86	613 $\pm$ 102	267 $\pm$ 124
Long grunt	7	0.306 $\pm$ 0.079	453 $\pm$ 195	33 $\pm$ 28	488 $\pm$ 175	432 $\pm$ 221	405 $\pm$ 182	508 $\pm$ 216	102 $\pm$ 74
Grunt	41	0.092 $\pm$ 0.013	549 $\pm$ 126	14 $\pm$ 13	545 $\pm$ 130	547 $\pm$ 131	533 $\pm$ 121	565 $\pm$ 134	32 $\pm$ 30
Growl	15	0.159 $\pm$ 0.062	1,935 $\pm$ 480	118 $\pm$ 79	1,973 $\pm$ 525	1,922 $\pm$ 527	1,778 $\pm$ 429	2,088 $\pm$ 539	310 $\pm$ 235
Bark	11	0.250 $\pm$ 0.070	811 $\pm$ 256	110 $\pm$ 100	803 $\pm$ 303	812 $\pm$ 313	676 $\pm$ 277	964 $\pm$ 289	288 $\pm$ 217
Rattle	16	0.433 $\pm$ 0.109	1,029 $\pm$ 341	305 $\pm$ 312	750 $\pm$ 274	1,217 $\pm$ 760	665 $\pm$ 202	1,648 $\pm$ 807	982 $\pm$ 827
Squeak	11	0.598 $\pm$ 0.144	747 $\pm$ 263	112 $\pm$ 56	768 $\pm$ 349	872 $\pm$ 287	612 $\pm$ 289	980 $\pm$ 277	369 $\pm$ 158
Squeal	9	0.861 $\pm$ 0.236	2,153 $\pm$ 1486	910 $\pm$ 693	1,975 $\pm$ 1682	2,452 $\pm$ 1712	962 $\pm$ 1127	3,582 $\pm$ 1979	2,620 $\pm$ 1750
Compound squeal	19	0.850 $\pm$ 0.329	1,424 $\pm$ 562	615 $\pm$ 402	1,249 $\pm$ 751	1,474 $\pm$ 917	624 $\pm$ 354	2,577 $\pm$ 1046	1,953 $\pm$ 1103
Bawl	3	0.621 $\pm$ 0.419	631 $\pm$ 161	31 $\pm$ 17	644 $\pm$ 130	589 $\pm$ 178	577 $\pm$ 160	693 $\pm$ 205	115 $\pm$ 49
Compound squeak	9	0.935 $\pm$ 0.067	493 $\pm$ 263	133 $\pm$ 161	625 $\pm$ 423	417 $\pm$ 229	297 $\pm$ 123	718 $\pm$ 406	421 $\pm$ 415
Three syllabled bark	13	0.522 $\pm$ 0.088	931 $\pm$ 182	241 $\pm$ 66	679 $\pm$ 90	835 $\pm$ 292	488 $\pm$ 84	1,254 $\pm$ 149	765 $\pm$ 168
Modulated tonal scream	21	0.993 $\pm$ 0.223	1,924 $\pm$ 648	996 $\pm$ 433	1,236 $\pm$ 494	2,444 $\pm$ 1,089	753 $\pm$ 386	3,702 $\pm$ 1,218	2,949 $\pm$ 1,164
Chuck	32	0.280 $\pm$ 0.050	1,393 $\pm$ 279	702 $\pm$ 249	642 $\pm$ 192	2,101 $\pm$ 460	626 $\pm$ 167	2,213 $\pm$ 463	1,588 $\pm$ 511
Mounting grunt	12	0.100 $\pm$ 0.027	514 $\pm$ 126	18 $\pm$ 21	509 $\pm$ 128	513 $\pm$ 129	490 $\pm$ 116	539 $\pm$ 143	50 $\pm$ 59
Call type	N	%T_min $f_0$	%T_max $f_0$	MeanAMP	MinAMP	MaxAMP	RangeAMP	HNR	Jitter
Coo	63	58.2 $\pm$ 28.9	31.8 $\pm$ 33.1	68.8 $\pm$ 12.6	56.2 $\pm$ 13.2	75.1 $\pm$ 12.9	18.8 $\pm$ 6.1	6.740 $\pm$ 2.488	2.468 $\pm$ 1.940
Shrill	9	41.7 $\pm$ 30.1	55.3 $\pm$ 35.4	67.5 $\pm$ 13.6	55.1 $\pm$ 11.4	72.6 $\pm$ 14.2	17.5 $\pm$ 6.7	6.601 $\pm$ 3.326	1.937 $\pm$ 1.292
Whine	20	54.2 $\pm$ 42.7	38.6 $\pm$ 35.0	70.1 $\pm$ 11.4	59.1 $\pm$ 10.2	76.9 $\pm$ 12.8	17.8 $\pm$ 5.4	10.543 $\pm$ 3.652	1.297 $\pm$ 1.217
Long grunt	7	67.3 $\pm$ 22.2	25.8 $\pm$ 27.1	73.2 $\pm$ 10.9	64.1 $\pm$ 14.0	77.2 $\pm$ 9.6	13.1 $\pm$ 5.3	5.897 $\pm$ 2.174	2.435 $\pm$ 2.054
Grunt	41	45.7 $\pm$ 24.0	50.2 $\pm$ 17.6	75.0 $\pm$ 8.1	74.1 $\pm$ 8.0	75.5 $\pm$ 8.2	1.4 $\pm$ 1.7	5.314 $\pm$ 2.835	2.503 $\pm$ 2.416
Growl	15	51.2 $\pm$ 19.9	49.8 $\pm$ 18.5	76.6 $\pm$ 4.1	66.5 $\pm$ 10.7	80.0 $\pm$ 3.7	13.4 $\pm$ 10.6	6.669 $\pm$ 4.540	9.444 $\pm$ 1.698
Bark	11	30.8 $\pm$ 31.1	29.0 $\pm$ 9.2	78.3 $\pm$ 2.8	65.5 $\pm$ 7.7	82.8 $\pm$ 3.0	17.2 $\pm$ 7.6	3.821 $\pm$ 1.987	5.105 $\pm$ 2.476
Rattle	16	21.2 $\pm$ 26.3	55.2 $\pm$ 26.0	77.6 $\pm$ 3.8	63.6 $\pm$ 6.6	83.5 $\pm$ 4.6	20.0 $\pm$ 8.2	6.285 $\pm$ 3.153	3.693 $\pm$ 1.329
Squeak	11	41.7 $\pm$ 24.0	58.5 $\pm$ 32.0	55.4 $\pm$ 3.1	46.1 $\pm$ 2.6	58.5 $\pm$ 2.9	12.4 $\pm$ 1.3	6.556 $\pm$ 1.743	1.004 $\pm$ 0.328
Squeal	9	48.2 $\pm$ 32.0	35.9 $\pm$ 34.2	70.6 $\pm$ 5.9	54.6 $\pm$ 9.3	77.1 $\pm$ 6.8	22.5 $\pm$ 8.2	5.105 $\pm$ 2.261	4.707 $\pm$ 2.351
Compound squeal	19	48.8 $\pm$ 30.0	61.6 $\pm$ 32.2	71.2 $\pm$ 8.2	56.5 $\pm$ 9.4	77.1 $\pm$ 8.2	20.6 $\pm$ 5.7	6.493 $\pm$ 3.334	4.175 $\pm$ 2.096
Bawl	3	79.8 $\pm$ 24.4	18.6 $\pm$ 22.6	56.3 $\pm$ 10.0	48.1 $\pm$ 11.2	61.9 $\pm$ 9.4	13.8 $\pm$ 1.9	7.664 $\pm$ 5.666	14.832 $\pm$ 3.684
Compound squeak	9	73.6 $\pm$ 19.3	24.7 $\pm$ 16.7	69.5 $\pm$ 11.7	60.0 $\pm$ 11.5	73.7 $\pm$ 11.3	13.8 $\pm$ 6.1	6.079 $\pm$ 2.074	1.688 $\pm$ 1.648
Three syllabled bark	13	40.2 $\pm$ 26.5	55.7 $\pm$ 17.8	79.0 $\pm$ 2.0	57.3 $\pm$ 4.8	85.1 $\pm$ 1.1	27.8 $\pm$ 5.0	5.091 $\pm$ 1.484	5.326 $\pm$ 1.435
Modulated tonal scream	21	39.0 $\pm$ 30.4	74.2 $\pm$ 19.0	75.3 $\pm$ 5.9	58.5 $\pm$ 7.3	81.6 $\pm$ 5.4	23.2 $\pm$ 4.2	8.006 $\pm$ 1.783	2.947 $\pm$ 1.971
Chuck	32	12.7 $\pm$ 11.2	72.0 $\pm$ 15.8	72.7 $\pm$ 7.5	57.3 $\pm$ 10.4	79.3 $\pm$ 7.1	22.0 $\pm$ 8.1	4.418 $\pm$ 7.459	7.065 $\pm$ 2.034
Mounting grunt	12	46.0 $\pm$ 26.8	50.9 $\pm$ 21.4	43.9 $\pm$ 11.4	41.8 $\pm$ 9.4	45.0 $\pm$ 12.4	3.3 $\pm$ 4.3	6.322 $\pm$ 3.405	3.861 $\pm$ 2.053



adults of the same social units (OMUs or AMU), they always submitted with behaviors such as crouching or avoiding; when adult females were threatened by adult males of the same OMUs, they submitted (sometimes with *compound squeal* calls) or resisted (usually with *bark* calls); when adult females were threatened by other adult females of the same OMUs, they submitted (sometimes with *compound squeal* calls) or resisted (sometimes with *growl* calls); when adult females were threatened by other adult females of different OMUs, they always resisted (with *growl* calls). The resistance against *grunt* threats often led to the immediate escalation of agonistic interactions. In addition, *grunts* were also observed to be emitted, by adult males of all social units, toward approaching humans (8%).

### 3.2.6 | Growl

*Growls* are harsh and plosive calls characterized by a relatively short duration ( $0.159 \pm 0.062$  s), a high  $f_0$  (mean  $f_0$ :  $1935 \pm 480$  Hz), and high intensity (meanAMP:  $76.6 \pm 4.1$  dB). This type of call accompanied by mobbing behaviors could be uttered singly or in bouts (range = 1–4, median = 2) by adult females involved in female-female ritualized agonistic interactions without physical contact within and between OMUs. The receivers avoided or defended with the same vocalizations accompanied by the same behaviors. Sometimes, *growl* calls appeared to be able to attract some other adult females from the same OMUs of the callers to form an alliance.

### 3.2.7 | Bark

*Barks* are also harsh and loud (meanAMP:  $78.3 \pm 2.8$  dB) calls like *growls*, but the duration is longer ( $0.250 \pm 0.070$  s) and the  $f_0$  is lower (mean  $f_0$ :  $811 \pm 256$  Hz). These vocalizations sounded extremely like dog barks to the human ear. *Barks* were uttered singly or doubly by both sexes involving in various forms of intense agonistic interactions with physical contact (e.g., grasping and biting), including female-female, male-female interactions within OMUs, and male-male interactions between two OMUs and between an OMU and the AMU.

### 3.2.8 | Rattle

When an adult female from an OMU was suddenly attacked by the adult male from another OMU, she uttered *rattle* calls accompanied by facial expressions of fear and by running quickly to the adult male of her own OMU. *Rattles* are tonal and rich in harmonics with a general trend of slow increase in frequency through the entire call, and have a medium duration ( $0.433 \pm 0.109$  s) and a relatively high  $f_0$  (mean  $f_0$ :  $1029 \pm 341$  Hz), and high intensity (meanAMP:  $77.6 \pm 3.8$  dB). These vocalizations usually elicited intense agonistic interactions between the adult male of the caller's OMU and the other who attacked the caller.

### 3.2.9 | Squeak

*Squeaks* are tonal calls with a medium duration ( $0.598 \pm 0.144$  s) and  $f_0$  (mean  $f_0$ :  $747 \pm 263$  Hz), and low intensity (meanAMP:  $55.4 \pm 3.1$  dB).

The harmonic structure has a general trend of slow decrease in frequency with one slight increase in the middle and another near the end. *Squeaks* were uttered by adult females during intense agonistic interactions between the adult males from their own OMUs and those from other OMUs. This type of call appeared to express a high degree of excitement that was almost always transferred from one female to another. This led to a highly synchronized pattern of vocal behavior, which could probably provide vocal support for the adult males of their own OMUs during male-male intense agonistic interactions. Female callers were never observed to be directly involved in such interactions.

### 3.2.10 | Squeal

*Squeals* comprise both tonal and harsh components and are characterized by a relatively long duration ( $0.861 \pm 0.236$  s) and a highest  $f_0$  (mean  $f_0$ :  $2153 \pm 1486$  Hz). The harmonic structure with an upward frequency modulation in middle is superimposed by slight broadband noisy elements. Adult females uttered *squeals* accompanied by moving quickly to their infants, upon realizing that their infants might fall in potential unsafe situations, such as male-male intense agonistic interactions.

### 3.2.11 | Compound squeal

*Compound squeals* comprise both tonal and harsh components, and the harmonic structure with few frequency modulations is superimposed by heavy and broadband noisy elements. This type of call has a relatively long duration ( $0.850 \pm 0.329$  s) and a high  $f_0$  (mean  $f_0$ :  $1424 \pm 562$  Hz). *Compound squeals* accompanied by submissive behaviors (e.g., crouching, avoiding) were emitted by subordinate adult females involving in female-female, male-female moderate agonistic interactions within OMUs. Upon hearing these vocalizations, the dominant party of the interactions usually stopped their threatening behaviors including *grunt* calls.

### 3.2.12 | Bawl

*Bawls* are tonal and characterized by a medium duration ( $0.621 \pm 0.419$  s) and a relatively low  $f_0$  (mean  $f_0$ :  $631 \pm 161$  Hz), and low intensity (meanAMP:  $56.3 \pm 10.0$  dB), and abundant harmonics with well frequency modulations. This type of call was emitted by both sexes immediately after intense agonistic interactions. Adult females uttered *bawls* accompanied by reconciliation behaviors (i.e., hugging each other) after female-female interactions within OMUs, while adult males emitted *bawls* after male-male interactions between OMUs. While calling, adult males laid the face down on the back of the individual in close proximity from their own OMUs.

### 3.2.13 | Compound squeak

*Compound squeaks* are basically tonal, and the harmonic structure is characterized by close frequency bands with few modulations

superimposed by slight noisy components. The duration is relatively long ( $0.935 \pm 0.067$  s) and the  $f_0$  is the second lowest (mean  $f_0$ :  $493 \pm 263$  Hz). These vocalizations were observed to be emitted by adult females during reconciliation behaviors immediately after intra-unit intense agonistic interactions. The winners of the interactions always called initially and the losers usually responded with the same vocalizations.

### 3.2.14 | Three syllabled bark

*Three syllabled barks* are loud calls (meanAMP:  $79.0 \pm 2.0$  dB) with a medium duration ( $0.522 \pm 0.088$  s), composed of three harsh syllables with most energy concentrated on the second and third syllables. These vocalizations were emitted by adult females lagging behind during unit/group traveling (73%) and by those located at the periphery of the group toward the AMU (27%). While no vocal responses were heard from other individuals, the adult females from the callers' OMUs stopped their ongoing behaviors and looked toward the directions of the callers.

### 3.2.15 | Modulated tonal scream

*Modulated tonal screams* are characterized by the second to longest duration ( $0.993 \pm 0.223$  s) and the second to highest  $f_0$  (mean  $f_0$ :  $1924 \pm 648$  Hz). These vocalizations comprise two tonal parts. The harmonic structure in the first part includes close continuously increasing frequency bands, whereas the frequency bands in the second part become much more dispersed and modulated. *Modulated tonal screams* were only observed to be emitted by adult females toward approaching humans with food supply.

### 3.2.16 | Chuck

*Chucks* consist of two harsh syllables with most energy concentrated on the second one. The duration is relatively short ( $0.280 \pm 0.050$  s) and the  $f_0$  is high (mean  $f_0$ :  $1393 \pm 279$  Hz). *Chucks* were uttered by all adult animals in response to sudden environmental disturbances, including the presence of snakes (5%), approaching humans (14%), breaking sounds of branches (16%), and other uncertain disturbances (65%). In the first three contexts, the callers scanned toward the directions of disturbances while calling, and those on the ground climbed up into the trees. Upon hearing *chucks*, the individuals in proximity, no matter which social units they belonged to, responded with the same vocalizations and accompanying behaviors. In the context of uncertain disturbances, neither the callers nor the individuals in proximity were observed to change their ongoing behaviors.

### 3.2.17 | Mounting grunt

Similar to *grunts*, *mounting grunts* have a harmonic structure with few frequency modulations and occurred in bouts (range = 2–6, median = 3) with irregular intervals, which could last up to more than

1 s. However, the duration is slightly longer ( $0.100 \pm 0.027$  s) and the  $f_0$  is slightly lower (mean  $f_0$ :  $514 \pm 126$  Hz) than those of *grunts*. The intensity is much lower than that of *grunts* and any other call type (meanAMP:  $43.9 \pm 11.4$  dB). *Mounting grunts* were observed to be emitted by adult males from OMUs during the period immediately before dismounting.

## 4 | DISCUSSION

To our knowledge, the present study was the first attempt to establish the vocal repertoire of free-ranging adult *R. roxellana* based on quantitative methods. We identified 18 call types based on auditory sense, visual inspection of spectrograms, and quantitative analyses of acoustic structure. It was not claimed that the complete vocal repertoire or all concurrent contexts of particular call types present in the natural habitat were covered due to the relatively short duration and habituation/provisioning effect in this study, however, we believed that the most and essential part was observed. The vocal repertoire of adult *R. roxellana* appeared to be larger than those of many other colobines that typically live in one-male multi-female (5 call types in *Colobus guereza*: Marler, 1972; 14 in *Trachypithecus johnii*: Hohmann, 1991; 8 in *Procolobus versus*: Bene & Zuberbueher, 2009) or multi-male multi-female (14 in *Semnopithecus entellus*: Hohmann, 1991) smaller groups without stratified structures. This comparison suggested a positive association between the size of vocal repertoire and group size/structure, consistent with the findings of some previous comparative studies of Old World primates and supporting the social complexity hypothesis (Bouchet et al., 2013; Gustison et al., 2012). However, the social complexity hypothesis does not appear to work when applied to New World primates (Cleveland & Snowdon, 1982; Snowdon, 2013). The adult cotton-top tamarins (*Saguinus oedipus*), for example, live in small family groups, but can produce up to 38 different types of calls (Cleveland & Snowdon, 1982).

We did not have much difficulty in classifying all vocalizations, but this did not discount the possible acoustic gradation existing among call types. Although some call types were relatively invariant or stereotyped, for example, *coos* and *chucks*, some others, especially those used at close range, appeared to be graded. For example, increasing the intensity of *mounting grunts* may lead to *grunts*, and further, if the duration increases and the  $f_0$  decreases in *grunts*, *long grunts* may result. Although some primates have been considered to have graded vocal repertoires (*Macaca fuscata*: Green, 1975; *Macaca sylvanus*: Hammerschmidt & Fischer, 1998) and some others have discrete vocal repertoires (*Cercopithecus diana*: Zuberbühler, Noë, & Seyfarth, 1997; *Cercopithecus neglectus*: Bouchet et al., 2012), a mixed vocal system with both graded and discrete signals appears to be the norm for most primates, living in either relatively open or closed habitats (reviewed in Green & Marler, 1979; *Papio hamadryas*: Rendall, Notman, & Owren, 2009; *Cercopithecus campbelli*: Lemasson & Hausberger, 2011). The level of gradedness or discreteness is likely to be varied among different call types depending on their specific functions (Bouchet et al., 2013; Lemasson & Hausberger, 2011). In

*C. campbelli*, for example, male alarm call types appear to be discrete, whereas female contact call types exhibit a high degree of variation (Lemasson & Hausberger, 2011; Ouattara, Lemasson, & Zuberbühler, 2009).

The results of the present study revealed a sexual asymmetry in the vocal repertoire size of adult *R. roxellana*: females emitted many more call types than males. Similar findings have already been reported in some species of Old World monkeys (*Cercocebus torquatus*: Bouchet et al., 2010; *C. neglectus*: Bouchet et al., 2012), whereas a high degree of call type sharing between sexes is found in some other species, particularly macaques (*M. sylvanus*: Hammerschmidt & Fischer, 1998; *Macaca thibetana*: Bernstein, Sheeran, Wagner, Li, & Koda, 2016). The great sex discrepancy in terms of call production in adult *R. roxellana* could be attributed to two non-exclusive factors. First, adult males are 1.5–2.0 times heavier in body mass than adult females (Davison, 1982; Jablonski & Pan, 1995), and thus the larger body size may limit male *R. roxellana* to emit some high pitched call types. Indeed, the call types with the highest  $f_0$ , that is, *growls*, *squeals*, *compound squeals*, and *modulated tonal screams*, were all female-specific. Meanwhile, male *R. roxellana* shared a relatively high pitched call type, *chucks* in particular, with females. Thus, the second factor, that is, sex specificity of social role, may also be an important reason for the sexual asymmetry in the vocal repertoire size of adult *R. roxellana*. Previous studies of this species have shown that females play a more important role in the maintenance of the OMU cohesion, and most social interactions occur among females within OMUs (Wang et al., 2013; Zhang et al., 2012). This was consistent with the result of the present study that most female-specific call types occurred during various forms of social interactions. Male *R. roxellana* may be potentially able to produce some call types, but their social roles constrained the expression of these vocalizations.

*R. roxellana* inhabits dense forests with poor visibility and lives in large multilevel groups typically formed by several OMUs and one AMU (Qi et al., 2014). In order to maintain intragroup cohesion and spacing, this primate should have developed vocalizations that allow information to be effectively transmitted in forest habitats where sound degradation is high. According to the acoustic adaptation hypothesis, vocalizations with long durations, harmonic patterns, low  $f_0$ , low mean frequencies, few frequency modulations, and narrow frequency ranges are suitable for long range communication in closed habitats (Ey & Fischer, 2009; Morton, 1975). The present study indicated that stereotyped *coo* calls possess these acoustic characteristics. We estimated the transmission distance of *coos* by one person walking toward the direction away from the monkey group while another person recording vocalizations (two persons keeping contact with wireless interphones). According to our rough estimates ( $N = 8$ ), *coos* could be transmitted over long distances up to at least 0.5 km in the dense forest (i.e., the whole group could be covered; based on our observations, the group spread in diameter was usually <0.3 km). Meanwhile, the exchanges of *coos* were observed to frequently occur among individuals within and between social units (except between OMUs and the AMU) in a variety of contexts, mainly including unit/group traveling and foraging when maintaining constant vocal contact

was particularly important. The acoustic properties and concurrent contexts suggested that *coos* were likely to be the contact calls used for both intra-unit and intragroup cohesion and spacing, especially when visual cues were blocked by long distances. The absence of *coo* exchanges among individuals between the AMU and OMUs was consistent with the observation of previous studies that the AMU is located at the spatial and social periphery of the whole group (Qi et al., 2014). Except *coos*, three other call types with relatively low  $f_0$  and harmonic patterns, that is, *shrills/whines* and *squeaks*, appeared to play important roles in facilitating the cohesion of OMUs under peaceful and agitated (e.g., male–male intense agonistic interactions) states, respectively, as suggested in the study of captive *R. roxellana* (Tenaza et al., 1988). In addition, *three syllabled barks* were likely to be used as “isolated calls” for adult females, and the loudness and harshness of these vocalizations might reflect the anxiety of the callers.

In order to maintain its complex social system, *R. roxellana* should also have evolved a variety of vocalizations for mediating various forms of social interactions at close range (intra-unit and inter-unit, agonistic, and affiliative) (McComb & Semple, 2005), as found in the present study: *long grunts*, *grunts*, *growls*, *barks*, *rattles*, *squeals*, *compound squeals*, *bawls*, and *compound squeaks*. Overall, the acoustic patterns of these vocalizations followed the concept of the motivation-structural rules (Morton, 1977). Specifically, *long grunts*, *grunts*, *growls*, and *barks* with relatively short durations and low  $f_0$  (except for *growls*) appeared to represent a ladder of hostility levels from the lowest to the highest, and the loudness and harshness increased along this ladder (Bernstein et al., 2016; Cleveland & Snowdon, 1982). *Growls* have a relatively high  $f_0$  probably because they also reflected the motivation of fear (Bernstein et al., 2016). The “appeasing calls,” *compound squeals*, have a long duration and a high  $f_0$  directed toward threatening individuals and inhibiting further attacks (Zimmermann, 1985). *Squeals*, with a long duration and a high  $f_0$ , were also likely to reflect the motivation of fear because these vocalizations occurred when the callers’ infants were actually or potentially attacked (Bernstein et al., 2016). *Rattles* were emitted when the callers were being attacked and were used to recruit support from unit members (adult males in particular), and thus the harmonic structure and medium  $f_0$  probably represented a balance between frightened motivation and information propagation (Bernstein et al., 2016; Morton, 1975). *Bawls* and *compound squeaks* were given as “friendly calls” and show harmonic structures with relatively long durations and low frequencies (Zimmermann, 1985).

In conclusion, our study showed that adult *R. roxellana* had a large vocal repertoire with a great sexual asymmetry. The results of this study could serve as a basis for in-depth studies on the vocal behavior of this species. For example, to fulfill their functions, the contact calls, especially *coos*, should be potentially able to convey the information about individual identity. Furthermore, some call types occurred in multiple contexts, for example, *coos* and *chucks*. There may be acoustic differences encoding context-specific information within each of these call types, allowing the receivers to make inferences about the events experienced by the callers, which is quite common in nonhuman primates (Fischer, Hammerschmidt, Cheney, & Seyfarth, 2001; Oda, 1996; Sugiura, 2007). Actually, this was evidenced in this study by the


varied reactions of the receivers in response to *chucks*, apparently serving as alert/alarm signals, provoked by different categories of environmental disturbances. Playback experiments should be used in future studies to investigate the possible acoustic variation and encoded meanings within call types. Studies on the vocal communication in *R. roxellana* are particularly valuable since they provide a case about how nonhuman primates, inhabiting forest habitats, and forming complex social systems, use their vocalizations to interact with their social and ecological environments.


## ACKNOWLEDGMENTS

We thank the field assistants, Jinwen Yang, Guangming Chen, and Shiping Zhu for helping us collect data in the field, and Lina Yi for assisting us in generating spectrograms. We appreciate the anonymous reviewers for great suggestions and comments on previous versions of the manuscript. We appreciate the financial support from the National Key R&D Program of China (2016YFC0503202), Hubei Provincial Key Laboratory for Conservation Biology of Snub-nosed Monkeys, and the Scientific Research Grant for Youth Scholars and the Special Grant for Research & Education Integration (KJRH2015-016) from the University of Chinese Academy of Sciences.

## ORCID

Penglai Fan  <http://orcid.org/0000-0002-3471-8294>

Xuecong Liu  <http://orcid.org/0000-0001-9345-2046>

Fang Li  <http://orcid.org/0000-0002-3757-402X>

Dingzhen Liu  <http://orcid.org/0000-0002-1870-6197>

## REFERENCES

- Altmann, S. A. (1967). The structure of primate social communication. In S. A. Altmann, (Ed.), *Social communication among primates* (pp. 325–336). Chicago: University of Chicago Press.
- Bene, J. K., & Zuberbuehler, K. (2009). Sex differences in the use of vocalizations in wild olive colobus monkeys. *European Journal of Scientific Research*, 25, 266–279.
- Bernstein, S. K., Sheeran, L. K., Wagner, R. S., Li, J., & Koda, H. (2016). The vocal repertoire of Tibetan macaques (*Macaca thibetana*): A quantitative classification. *American Journal of Primatology*, 78, 937–949. <https://doi.org/10.1002/ajp.22564>
- Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2012). Age- and sex-specific patterns of vocal behaviour in De Brazza's monkeys (*Cercopithecus neglectus*). *American Journal of Primatology*, 74, 12–28. <https://doi.org/10.1002/ajp.21002>
- Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, 4, 390. <https://doi.org/10.3389/fpsyg.2013.00390>
- Bouchet, H., Pellier, A., Blois-Heulin, C., & Lemasson, A. (2010). Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *American Journal of Primatology*, 72, 360–375. <https://doi.org/10.1002/ajp.20791>
- Briseño-Jaramillo, M., Biquand, V., Estrada, A., & Lemasson, A. (2017). Vocal repertoire of free-ranging black howler monkeys' (*Alouatta pigra*): call types, contexts, and sex-related contributions. *American Journal of Primatology*, 79, 1–15. <https://doi.org/10.1002/ajp.22630>
- Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009a). The information content of giant panda, *Ailuropoda melanoleuca*, bleats: Acoustic cues to sex, age and size. *Animal Behaviour*, 78, 893–898. <https://doi.org/10.1016/j.anbehav.2009.06.029>
- Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009b). Vocal cues to identity and relatedness in giant pandas (*Ailuropoda melanoleuca*). *Journal of the Acoustical Society of America*, 126, 2721–2732. <https://doi.org/10.1121/1.3224720>
- Cleveland, J., & Snowdon, C. T. (1982). The complex vocal repertoire of the adult cotton-top tamarin (*Saguinus oedipus oedipus*). *Zeitschrift für Tierpsychologie*, 58, 231–270. <https://doi.org/10.1111/j.1439-0310.1982.tb00320.x>
- Davison, G. W. H. (1982). Convergence with terrestrial cercopithecines by the monkey *Rhinopithecus roxellanae*. *Folia Primatologica*, 37, 209–215. <https://doi.org/10.1159/000156033>
- Ey, E., & Fischer, J. (2009). The “acoustic adaptation hypothesis”—A review of the evidence from birds, anurans and mammals. *Bioacoustics: The International Journal of Animal Sound and Its Recording*, 19, 21–48. <https://doi.org/10.1080/09524622.2009.9753613>
- Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2001). Acoustic features of female chacma baboon barks. *Ethology*, 107, 33–54. <https://doi.org/10.1111/j.1439-0310.2001.00630.x>
- Green, S. (1975). Communication by a graded vocal system in Japanese monkeys. In L. A. Rosenblum, (Ed.), *Primate behaviour* (pp. 1–102). New York: Academic Press.
- Green, S., & Marler, P. (1979). The analysis of animal communication. In P. Marler, & J. G. Vandenbergh, (Eds.), *Handbook of behavioral neurobiology: social behavior and communication* (pp. 73–147). New York: Plenum Press.
- Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B*, 367, 1847–1859. <https://doi.org/10.1098/rstb.2011.0218>
- Hammerschmidt, K., & Fischer, J. (1998). The vocal repertoire of barbary macaques: A quantitative analysis of a graded signal system. *Ethology*, 104, 203–216. <https://doi.org/10.1111/j.1439-0310.1998.tb00063.x>
- Hauser, M. D. (1993). The evolution of nonhuman primate vocalizations: Effect of phylogeny, body weight, and social context. *The American Naturalist*, 142, 528–542. <https://doi.org/10.1086/285553>
- Hohmann, G. (1991). Comparative analyses of age- and sex-specific patterns of vocal behaviour in four species of Old World monkeys. *Folia Primatologica*, 56, 133–156. <https://doi.org/10.1159/000156538>
- Jablonski, N. G., & Pan, R. (1995). Sexual dimorphism in the snub-nosed langurs (Colobinae: *Rhinopithecus*). *American Journal of Physical Anthropology*, 96, 251–272. <https://doi.org/10.1002/ajpa.1330960304>
- Kirkpatrick, R. C., & Grueter, C. C. (2010). Snub-nosed monkeys: Multilevel societies across varied environments. *Evolutionary Anthropology*, 19, 98–113. <https://doi.org/10.1002/evan.20259>
- Klecka, W. (1980). *Discriminant analysis*. Beverly Hills, CA: Sage.
- Lachenbruch, P. A. (1975). *Discriminant analysis*. NY: Hafner.
- Lemasson, A., & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Journal of the Acoustical Society of America*, 129, 3341–3352. <https://doi.org/10.1121/1.3569704>
- Lévréro, F., Carrete-Vega, G., Herbert, A., Lawabi, I., Courtiol, A., Willaume, E., ... Charpentier, M. J. E. (2015). Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communications*, 6, 7609. <https://doi.org/10.1038/ncomms8609>
- Li, B., Chen, F., Luo, S., & Xie, W. (1993). Major categories of vocal behaviour in wild Sichuan snub-nosed monkey (*Rhinopithecus roxellana*). *Acta Theriologica Sinica*, 13, 181–187.

- Li, B., Pan, R., & Oxnard, C. E. (2002). Extinction of snub-nosed monkeys in China during the past 400 years. *International Journal of Primatology*, 23, 1227–1244. <https://doi.org/10.1023/A:1021122819845>
- Li, Y. (2007). Terrestriality and tree stratum use in a group of Sichuan snub-nosed monkeys. *Primates*, 48, 197–207. <https://doi.org/10.1007/s10329-006-0035-9>
- Liu, X., Stanford, C. B., Yang, J., Yao, H., & Li, Y. (2013). Foods eaten by the Sichuan snub-nosed monkey (*Rhinopithecus roxellana*) in Shennongjia National Nature Reserve, China, in relation to nutritional chemistry. *American Journal of Primatology*, 75, 860–871. <https://doi.org/10.1002/ajp.22149>
- 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. <https://doi.org/1186/1742-9994-10-58>
- Marler, P. (1972). Vocalizations of East African monkeys II: Black and white colobus. *Behaviour*, 42, 175–197. [https://doi.org/10.1163/156853972\(00266\)](https://doi.org/10.1163/156853972(00266))
- Marler, P. (1976). Social organization, communications and graded signals: The chimpanzee and the gorilla. In P. P. G. Bateson, & R. A. Hinde, (Eds.), *Growing points in ethology* (pp. 239–280). Cambridge: Cambridge University Press.
- McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, 1, 381–385. <https://doi.org/10.1098/rsbl.2005.0366>
- Morton, E. S. (1975). Ecological sources of selection on avian sounds. *The American Naturalist*, 109, 17–34. <https://doi.org/10.1086/282971>
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111, 855–869. <https://doi.org/10.1086/283219>
- Oda, R. (1996). Effects of contextual and social variables on contact call production in free-ranging ring tailed lemurs (*Lemur catta*). *International Journal of Primatology*, 17, 191–205. <https://doi.org/10.1007/BF02735447>
- Quattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys use affixation to alter call meanings. *PLoS ONE*, 4, e7808. <https://doi.org/10.1037/journal.pone.0007808>
- Pfefferle, D., & Fischer, J. (2006). Sounds and size: Identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*. *Animal Behaviour*, 72, 43–51. <https://doi.org/10.1016/j.anbehav.2005.08.021>
- Qi, X., Garber, P. A., Ji, W., Huang, Z., Huang, K., Zhang, P., . . . Li, B. (2014). Satellite telemetry and social modeling offer new insights into the origin of primate multilevel societies. *Nature Communications*, 5, 5296. <https://doi.org/10.1038/ncomms6296>
- Ren, R., Yan, K., Su, Y., Zhou, Y., Li, J., Zhu, Z., . . . Hu, Y. (2000). *The society of golden monkeys (Rhinopithecus roxellana)*. Beijing: Peking University Press.
- Rendall, D., Notman, H., & Owren, M. J. (2009). Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *Journal of the Acoustical Society of America*, 125, 1792–1805. <https://doi.org/10.1121/1.3068453>
- Salmi, R., & Doran-Sheehy, D. M. (2014). The function of loud calls (hoot series) in wild western gorillas (*Gorilla gorilla*). *American Journal of Physical Anthropology*, 155, 379–391. <https://doi.org/10.1002/ajpa.22575>
- Snowdon, C. T. (2013). Language parallels in new world primates. In S. A. Helekar, (Ed.), *Animal models of speech and language disorders* (pp. 241–261). New York: Springer.
- Snowdon, C. T., & Elowson, A. M. (2001). Babbling' in pygmy marmosets: Development after infancy. *Behaviour*, 138, 1235–1248. <https://doi.org/10.1163/15685390152822193>
- Sugiura, H. (2007). Effects of proximity and behavioral contexts on acoustic variation in the coo calls of Japanese macaques. *American Journal of Primatology*, 69, 1412–1424. <https://doi.org/10.1002/ajp.20447>
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.
- Tenaza, R. R., Fitch, H. M., & Lindburg, D. G. (1988). Vocal behaviour of captive Sichuan snub-nosed monkeys (*Rhinopithecus r. roxellana*). *American Journal of Primatology*, 14, 1–9. <https://doi.org/10.1002/ajp.1350140102>
- Wang, X., Wang, C., Qi, X., Guo, S., Zhao, H., & Li, B. (2013). A newly-found pattern of social relationships among adults within one-male units of golden snub-nosed monkeys (*Rhinopithecus roxellana*) in the Qinling Mountains, China. *Integrative Zoology*, 8, 400–409. <https://doi.org/10.1111/1749-4877.12026>
- Waser, P. M., & Brown, C. H. (1986). Habitat acoustics and primate communication. *American Journal of Primatology*, 10, 135–154. <https://doi.org/10.1002/ajp.1350100205>
- Yao, H., Liu, X., Stanford, C. B., Yang, J., Huang, T., Wu, F., & Li, Y. (2011). Male dispersal in a provisioned multilevel group of *Rhinopithecus roxellana* in Shennongjia Nature Reserve, China. *American Journal of Primatology*, 73, 1280–1288. <https://doi.org/10.1002/ajp.21000>
- Zhang, P., Li, B., Qi, X., MacIntosh, A. J. J., & Watanabe, K. (2012). A proximity-based social network of a group of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*). *International Journal of Primatology*, 33, 1081–1095. <https://doi.org/10.1007/s10764-012-9608-1>
- Zimmermann, E. (1985). The vocal repertoire of the adult Senegal bushbaby (*Galago senegalensis senegalensis*). *Behaviour*, 94, 212–233. [https://doi.org/10.1163/156853985\(00190](https://doi.org/10.1163/156853985(00190)
- Zuberbühler, K., Noë, R., & Seyfarth, R. M. (1997). Diana monkey long-distance calls: Messages for conspecifics and predators. *Animal Behaviour*, 53, 589–604. <https://doi.org/10.1006/anbe.1996.0334>

**How to cite this article:** Fan P, Liu X, Liu R, et al. Vocal repertoire of free-ranging adult golden snub-nosed monkeys (*Rhinopithecus roxellana*). *Am J Primatol*. 2018;80:e22869. <https://doi.org/10.1002/ajp.22869>