Table S1. Number of individuals and calls, and call descriptions by sex for each species in Study 1.

Species	Males			Females		
	Number of Individuals	Call Description	Number of Calls	Number of Individuals	Call Description	Number of Calls
Alouatta caraya	2	Long distance calls (roars); during group encounters	**3**			
Ateles geoffroyi	9	Feeding	3	9	Moving	1
		Resting	3		Feeding	4
		Moving	1		Resting	1
		Whinny	1		Whinny	7
		Play call	1		Bark	1
		Long-distance yell	1			
			Total: 10			**Total: 14**
Brachyteles hypoxanthus	5	Neigh	**30**	5	Neigh	**27**
Callithrix geoffroyi		Song	2			
		Song and cry	1			
			Total: 3			
Callithrix jacchus	4	Food calls	**21**	2	Food calls	**6**
Callithrix kuhlii					Cry	1
					Song and cry	1
						Total: 2
Cebus apella	≥ 2	Fear; "Response to seeing a leather catch glove"	1	≥ 2	Fear; "Response to seeing a leather catch glove"	1
		Unknown	2		Unknown	17
			Total: 3			**Total: 18**
Cebus capucinus	6	Food-associated calls (while foraging/feeding)	**50**	13	Food-associated calls (while foraging/feeding)	**181**
Cercocebus atys	10	Grunt	**27**	34	Copulation call	16
					Grunt (foraging)	72
						Total: 88
Cercopithecus campbelli	3	Boom	5	2	Rapid, Repetitive and Ascendant	**16**
		Hok	5			

		Hokoo	5			
		Krak	5			
		Krakoo	5			
		Wakoo	5			
			Total: 30			
Cercopithecus mitis	≥ 7	Pyow	7	≥ 2	Chirp	8
		Ant	1		Grunt	4
		Boom	7		Long-grunt	2
		Ka	1		Long-grunt + trill	2
		Katrain	2		Growls	3
		Katrain + pyow	1		Trill	1
					Trill + grunt	2
			Total: 19			**Total: 22**
Cercopithecus mona		Grunt (copulation call; copulating adult pair at edge of mixed-sex group)	**3**		Whine (copulation call; copulating adult pair at edge of mixed-sex group)	**6**
Colobus guereza	2	Roar	**2**			
Erythrocebus patas	1	Moo	**2**	≥ 14	Moo	12
					Moo (allomothering)	31
					Moo (inter-group encounter)	17
						Total: 60
Gorilla gorilla	8	Grumble	68	7	Hum	12
		Cough-grunt/cough throat-clear	2		Hum-grumble	5
		Hoot	4		Grumble	22
		Scream	2		Hum-sing	1
		Whimper	1		Cough/cough-grunt/throat-clear	3
		Train grunt	4		Dog whine/sing	1
		Copulatory grunt	1		Whine	2
		Grumble-hum	1			
		Unknown	1			
			Total: 84			**Total: 46**
Homo sapiens	30	Hujambo ("Hello")	15	30	Hujambo ("Hello")	15

		Rainbow Passage (first paragraph)	15 **Total: 30**		Rainbow Passage (first paragraph)	15 **Total: 30**
Hylobates klossii				2	Unknown	**2**
Hylobates lar	7	Solo Unknown	5 2 **Total: 7**	≥ 6	Female great call (as part of male-female duets)	**9**
Hylobates moloch	2	Unknown	**2**	2	Unknown	**2**
Hylobates muelleri	2	Unknown	**2**			
Hylobates pileatus	2	Unknown	**2**			
Macaca fascicularis				5	Female copulation call	**150**
Macaca fuscata				2	Coo call (contact call)	**48**
Macaca mulatta	4	Grunt or girney	**26**	9	Grunt or girney Coo **Total: 75**	62 13
Nasalis larvatus	13	Honk Very short bray?	27 2 **Total: 29**	≥ 9	Honk Honk (angry) Quack **Total: 30**	14 7 9
Nomascus concolor	2	Unknown	**2**			
Pan troglodytes	≥ 12	Pant-hoot Rough-grunt Scream	3 23 1 **Total: 27**	≥ 16	Pant-hoot Rough-grunt Scream Pant-grunt Grunt Pant-bark Whimper **Total: 31**	5 15 5 2 1 2 1
Papio hamadryas					Unknown	**7**
Papio ursinus	7	Grunt	**155**	8	Grunt	**126**
Pongo pygmaeus	6	Grumph, alarm/disturbance	**34**	10	Grumph, alarm/disturbance	**48**
Rhinopithecus bieti		Moan Unknown	1 5	≥ 3	Wika (part of call) Unknown	6 6

			Total: 6			**Total: 12**
Rhinopithecus roxellana		Unknown	**4**		Unknown	**2**
Saguinus oedipus	3	Unknown	**13**	5	Long call Unknown	2 8 **Total: 10**
Saimiri sciureus	2	Unknown	**14**		Unknown	**13**
			Total: 640			**Total: 1081**

Table S2. Mean F_0 (Hz) and body weight (g) by sex, as well as habitat, social system, and mating system for species in Study 1.

Species	Fem. F_0	Male F_0	Fem. Wt.	Male Wt.	Habitat	Social System	Mating System
A. caraya (1-3)		77	4661	7057	Arb	UM, MF	Pr
A. geoffroyi (1, 2)	1326	1692	7315	7939	Arb	MF	Pr
B. hypoxanthus (1)	1715	1731	8400	9600	Arb	MF	Pr
C. geoffroyi (1)		7353	280	307	Arb	PL, UM, MF	M
C. jacchus (1-4)	6478	7435	271	267	Arb	PL, UM, MF	M
C. kuhlii (1-4)	4853		271	267	Arb	PL, UM, MF	M
C. apella (1-5)	1881	1771	2483	3374	Arb, Ter	MF	Po
C. capucinus (1, 3, 4)	709	731	2495	3626	Arb	MF	Pr
C. atys (1)	180	77	6075	10553	Arb, Ter	MF	Pr
C. campbelli (1)	1153	163	3150	4050	Arb, Ter	UM	Po
C. mitis (1, 3)	1453	216	3963	6411	Arb, Ter	UM, MF	Po
C. mona (1, 3, 4)	733	334	2825	4566	Arb	UM	Po
C. guereza (1, 3, 4)		76	8584	11006	Arb	PL, UM, MF	Po
E. patas (1-3)	273	150	6200	11300	Ter	UM	Po
G. gorilla (1, 2, 4, 5)	86	63	81053	162390	Ter	UM, MF	Po
H. sapiens (4, 6)	225	115	50579*	66378	Ter	UM, MF, PL	Variable
H. klossii (1, 3, 4)	756	866	5903	5680	Arb	PL	M
H. lar (1, 2, 4)	860	845	5317	6007	Arb	PL	M
H. moloch (1-3)	973	929	6083	6157	Arb	PL	M
H. muelleri (1)	872	806	5667	6027	Arb	PL	M
H. pileatus (4)		778	5440	5550	Arb	PL	M
M. fascicularis (1-3, 5)	179		3430	5307	Arb, Ter	MF	Po
M. fuscata (1-4)	763		8803	11880	Arb, Ter	MF	Pr
M. mulatta (1-4)	269	160	5706	8466	Arb, Ter	MF	Pr
N. larvatus (1, 4)	175	96	9851	19466	Arb	UM	Po
N. concolor (3, 4, 7)		1215	6710	7363	Arb	PL	M
P. troglodytes (1-4)	466	28	39922	48332	Arb, Ter	MF	Pr
P. hamadryas (1-3)	77		11933	21067	Arb, Ter	UM, MF	Po
P. ursinus (1-3)	107	56	14530	24330	Ter	UM, MF	Po
P. pygmaeus (1-4)	51	41	36605	77158	Arb, Ter	S, MF	Po
R. bieti (1)	552	445	8733	17600	Arb, Ter	UM	Po
R. roxellana (1)	815	508	8439	16344	Arb, Ter	UM	Po
S. oedipus (1-4)	1768	1943	456	456	Arb	PL, UM, MF	M

S. sciureus (1-4)	5859	5123	658	835	Arb	MF	Pr

Body weight data from sources cited after species names. Habitat classifications of arboreality (Arb) and terrestriality (Ter) based on data obtained from(1). Social system classifications (UM = unimale, MF = multimale-multifemale, PL = pair-living) obtained from supplementary information in ref. (8). Mating system classifications (Po = polygynous, Pr = promiscuous, M = monogamous) based on information obtained from (9). *Adjusted from 56884 to equalize male and female body fat percentage (see Supplementary Methods).

Table S3. Descriptive Statistics for Studies 2 and 3.

	Women		Men	
	Mean	SD	Mean	SD
Study 2				
Age	20.03	1.59	20.10	1.70
Duration (s)	5.41	0.94	5.48	0.97
F_o (Hz)	204.02	18.33	112.31	15.03
F_o-SD	36.30	10.24	15.65	5.21
F_1	520.66	39.11	445.25	31.21
F_2	1776.15	79.22	1514.64	68.40
F_3	2696.46	95.26	2393.42	83.89
F_4	3858.28	139.32	3390.18	120.87
Harmonics	15.55	1.59	12.54	1.46
Jitter	0.0191	0.0041	0.0241	0.0065
JitterRap	0.0086	0.0023	0.0098	0.0040
JitterPpq5	0.0086	0.0019	0.0110	0.0029
JitterDdq	0.0259	0.0070	0.0295	0.0119
Shimmer	0.0701	0.0099	0.0836	0.0110
ShimmerApq3	0.0247	0.0058	0.0288	0.0060
ShimmerApq5	0.0340	0.0062	0.0440	0.0073
ShimmerApq11	0.0649	0.0110	0.0856	0.0168
ShimmerDda	0.0742	0.0175	0.0864	0.0179
Voice Breaks	14.74	2.75	11.96	2.59
Study 3: Sample 1				
Age (years)	19.36	1.52	19.92	1.98
F_o (Hz)	199.99	20.61	106.91	12.34
T (pg/mL)	66.45	24.54	163.51	63.27
C (µg/mL)	27.12	20.94	25.54	12.14
Hours since waking	4.89	2.25	4.97	2.47
Study 3: Sample 2				
Age (years)			19.91	1.22
F_o (Hz)			106.11	11.05
T (pg/mL)			169.03	63.13
C (µg/mL)			42.65	19.84
Hours since waking			2.15	1.18

Table S4. Results of PCA on (natural log-transformed) jitter, shimmer, harmonics, and voice breaks for Study 2 women and men.

	Women		**Men**	
	PC 1	**PC 2**	**PC 1**	**PC 2**
	EV=7.2, 65.5%	EV=1.8, 16.7%	EV=6.4, 58.2%	EV=2.4, 21.8%
Ln(voice breaks)	0.042	0.453	-0.008	0.528
Ln(jitter)	0.325	**0.909**	0.253	**0.937**
Ln(jitterRap)	0.276	**0.944**	0.205	**0.951**
Ln(jitterPpq5)	0.346	**0.849**	0.347	**0.851**
Ln(jitterDdq)	0.276	**0.944**	0.205	**0.951**
Ln(shimmer)	**0.932**	0.280	**0.927**	0.198
Ln(shimmerApq3)	**0.783**	0.496	**0.875**	0.339
Ln(shimmerApq5)	**0.946**	0.240	**0.954**	0.148
Ln(shimmerApq11)	**0.867**	-0.021	**0.772**	-0.077
Ln(shimmerDda)	**0.783**	0.496	**0.875**	0.339
Harmonics	**-0.773**	-0.394	**-0.761**	-0.261

EV = Eigenvalue. Percentage refers to amount of variance explained by each component. Bolded acoustic parameters in each column were standardized and summed (excepting harmonics-noise-ratio, which was standardized and subtracted).

Table S5. Multiple regression models predicting attractiveness and dominance in Study 2 men and women (all VIF < 2.2).

Men		*F*	*df*	R^2	*β*	*t*	*p*
Dominance	Model	10.98	7,166	0.32			<0.0001
	Mean F_0				-0.43	-4.56	<0.0001
	F_0-SD				0.07	0.78	0.438
	Formant position				-0.31	-4.53	<0.0001
	Jitter				-0.15	-1.77	0.079
	Shimmer-harmonics				0.08	1.06	0.290
	Voice breaks				-0.20	-2.43	0.016
	Duration				-0.10	-1.34	0.181
ST attractiveness	Model	8.98	7,167	0.27			<0.0001
	Mean F_0				-0.36	-3.72	<0.001
	F_0-SD				0.04	0.42	0.678
	Formant position				-0.03	-0.39	0.696
	Jitter				-0.18	-2.01	0.046
	Shimmer-harmonics				-0.03	-0.34	0.732
	Voice breaks				-0.03	-0.39	0.695
	Duration				-0.41	-5.26	<0.0001
LT attractiveness	Model	7.53	7,167	0.24			<0.0001
	Mean F_0				-0.33	-3.31	0.001
	F_0-SD				0.06	0.68	0.500
	Formant position				0.01	0.18	0.857
	Jitter				-0.15	-1.70	0.091
	Shimmer-harmonics				-0.04	-0.51	0.608
	Voice breaks				-0.00	-0.02	0.983
	Duration				-0.43	-5.36	<0.0001
Women							
ST attractiveness	Model	15.70	7,250	0.31			<0.0001
	Mean F_0				-0.03	-0.45	0.650
	F_0-SD				0.08	1.22	0.223
	Formant position				0.44	7.69	<0.0001
	Jitter				-0.04	-0.49	0.625
	Shimmer-harmonics				-0.02	-0.31	0.758
	Voice breaks				0.10	1.61	0.110
	Duration				-0.33	-5.61	<0.0001
LT attractiveness	Model	14.01	7,249	0.28			<0.0001
	Mean F_0				-0.03	-0.45	0.652
	F_0-SD				0.05	0.80	0.423
	Formant position				0.36	6.24	<0.0001
	Jitter				-0.04	-0.46	0.643
	Shimmer-harmonics				-0.01	-0.19	0.850
	Voice breaks				0.13	1.96	0.051
	Duration				-0.41	-6.72	<0.0001

Table S6. Multiple regression models predicting F_0 (all VIF < 1.6) in Study 3.

		F	*df*	R^2	*β*	*t*	*p*
Sample 1 men	Model	3.00	3,58	0.13			0.038
	T				-0.17	-1.34	0.186
	C				0.24	1.78	0.080
	T × C				0.33	2.53	0.014
	Model	3.58	4,53	0.21			0.012
	T				-0.09	-0.68	0.500
	C				0.37	2.66	0.010
	T × C				0.37	2.81	0.007
	Time since waking				0.26	1.88	0.065
Sample 1 women	Model	0.51	3,49	0.03			0.680
	T				-0.01	-0.04	0.968
	C				0.07	0.48	0.637
	T × C				0.16	1.13	0.263
	Model	0.60	4,48	0.05			0.664
	T				-0.02	-0.11	0.916
	C				0.10	0.67	0.508
	T × C				0.15	1.04	0.304
	Time since waking				0.14	0.94	0.351
Sample 2 men	Model	2.95	3,54	0.14			0.041
	T				-0.27	-1.91	0.062
	C				0.15	1.04	0.303
	T × C				0.28	2.21	0.031
	Model	2.17	4,53	0.14			0.085
	T				-0.27	-1.89	0.064
	C				0.15	0.91	0.369
	T × C				0.28	2.19	0.033
	Time since waking				-0.003	-0.02	0.986

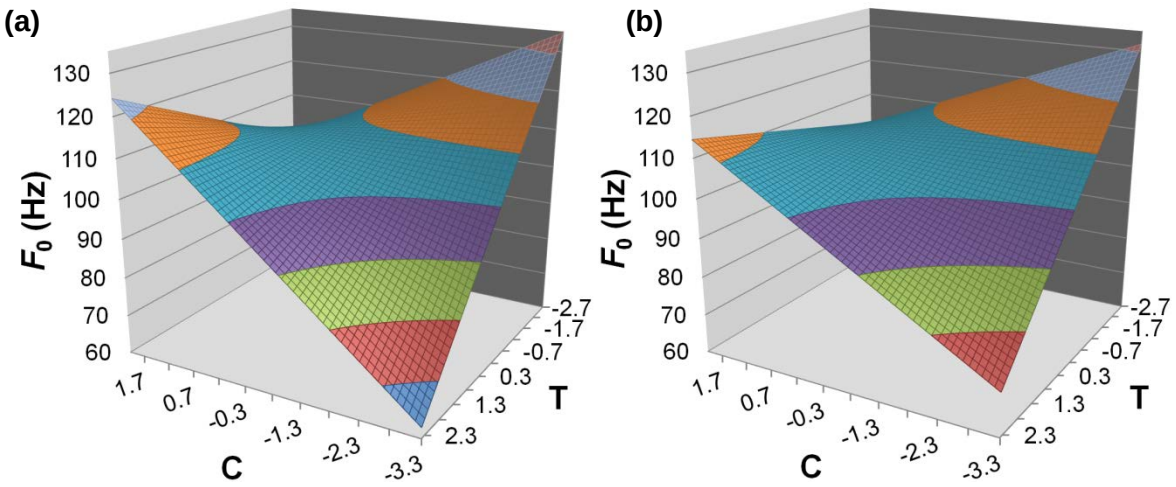

Fig. S1. Relationships of vocal fundamental frequency (F_0) with cortisol (C) and testosterone (T) in men from (a) Sample 1 and (b) Sample 2. Hormone concentrations (unadjusted for time of day) are natural log-transformed, then standardized to reduce collinearity with interaction terms. In both samples, cortisol and testosterone negatively interacted, such that testosterone was significantly negatively related to F_0 only in men with low cortisol levels (see also Fig. S2, Table S5). Colors represent 10-Hz contour intervals.

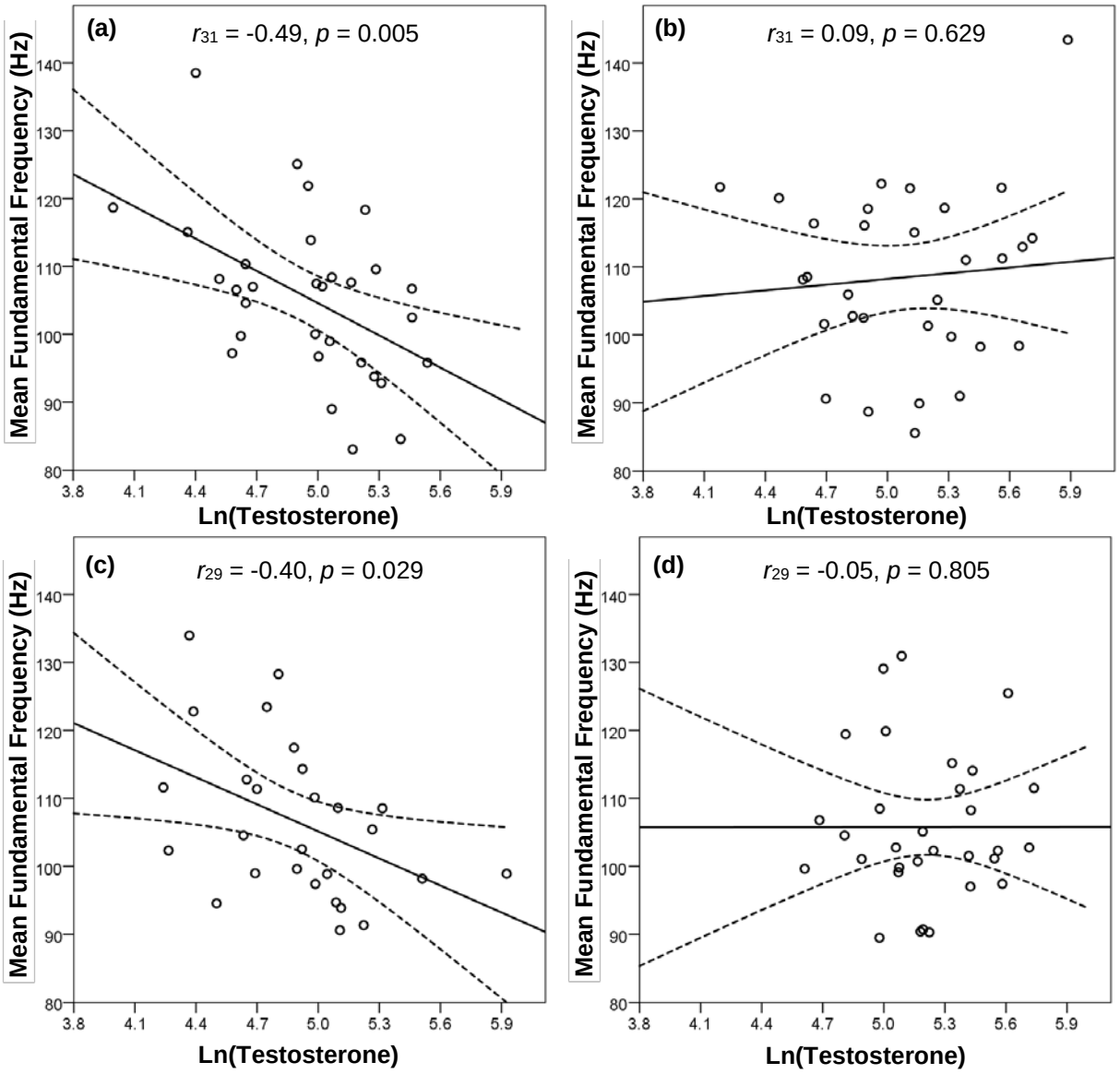

Fig. S2. Correlations from Study 3 between mean fundamental frequency and testosterone in men with cortisol below the median in Sample 1 (a) and Sample 2 (c), and in men with cortisol levels above the median in Sample 1 (b) and Sample 2 (d). In both samples, testosterone and fundamental frequency correlated only in men with low cortisol. Hormone levels are unadjusted for time of day. Least-squares regression lines with 95% CI are shown.

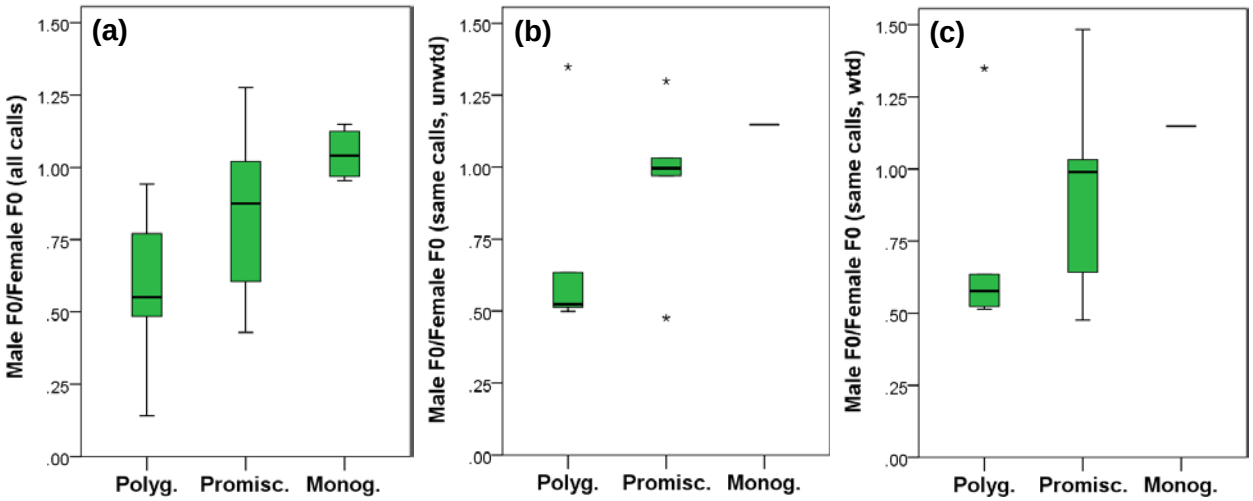

Fig. S3. Sexual dimorphism in F_0 varied as a function of mating system and showed highly similar patterns whether we analyzed all calls (a) or only calls produced by both sexes in our data (b and c). In (b), F_0 was averaged across all vocalizations of each call type in each sex, and then the call type means were averaged for each sex (effectively weighting common call types equally). In (c), F_0 was averaged across all vocalizations for each sex, provided that the vocalization was of a call type that both sexes produced in our data (effectively weighting each common call type by the number of vocalizations). Bold horizontal bars represent medians, boxes represent third and first quartiles, and error bars represent 95% CI.

SI Results

Additional Analyses for Study 1

It is possible that we obtained the observed relationships between F_0 dimorphism and mating system not because males tend to evolve lower F_0 for a given call type in polygynous mating systems, but because call types differ in F_0, and the frequency with which different calls are produced is more sexually differentiated in polygynous species. To explore this possibility, we selected only call types for which we had at least one male and one female call in a particular species, according to the descriptive information provided by the donors of the acoustic files. For each of these call types, we calculated the average male F_0 and average female F_0. We then computed the average male and average female F_0 across all common call types. Finally, we computed the ratio of male-to-female F_0 in two ways: (a) by weighting each call type by the number of vocalizations of that type, and (b) by weighting all call types equally, to eliminate any sex differences in the frequency with which certain calls are given.

Although the total number of calls (n = 1015) and number of species (n = 12) were greatly reduced, precluding phylogenetic generalized least squares (PGLS) regression, F_0 dimorphism across all call types correlated well with F_0 dimorphism from only common call types (call types weighted: $r_{12} = 0.81$, $p < 0.002$; call types unweighted: $r_{12} = 0.88$, $p < 0.0002$). In addition, we observed very similar patterns of F_0 dimorphism across mating systems regardless of whether all calls or only common call types were used (Fig. S3).

SI Materials and Methods
Study 1

Data collection. In addition to utilizing our own data, we contacted approximately 250 primate researchers to request recordings that they may have produced in the course of their own research, as well as accompanying information on call type and any subspecies designations.

We measured all recordings as uncompressed .WAV or .AIFF files. Most were recent in origin and therefore made digitally. Some were older and received on magnetic cassette tapes, which were then converted to digital format. These analog recordings were transferred from cassettes to a computer using the acoustic software program Audacity, after which they were saved as .WAV files and burned in Redbook format to compact discs. The recordings remained un-manipulated except for minimization of white noise where necessary. While recordings of all but one species in our sample were provided in lossless formats, it is possible that some recordings given to us secondhand were previously compressed and then uncompressed. Using compressed file formats (e.g., Digital Compact Cassette, MiniDisc, MP3, MP4, VQF, Liquid Audio) may result in a loss of relevant acoustic information (10). However, common compression algorithms are unlikely to affect F_0, as it is in the audible acoustic range, and these algorithms are designed to remove less salient sound components. As compression techniques aim to preserve sounds perceivable by humans, it is possible that frequencies outside this range may have been affected. However, given that primatologists prefer to record in lossless .WAV and .AIFF formats, it is unlikely that compressed formats were used. In addition, the lone species in our sample (*Erythrocebus patas*) whose files we received in lossy format will have been minimally affected by compression since this species' frequency range is similar to that of humans.

We received thousands of recordings from which 1721 files were selected for use in this study (Table S1). In determining which calls were most appropriate for analysis, we eliminated those with overlapping calls from multiple animals. Moreover, care was taken to ensure that each recording used was produced by a single individual of known species, sex, and adult status. In addition, any vocalizations with substantial background noise (e.g., rustling leaves, vehicles) were carefully scrutinized to determine whether a high-quality F_0 measurement could be made. Despite these stringent criteria, capturing sounds outside of a controlled recording booth presents difficulties that can be negotiated only imperfectly, and human judgment was used to determine the suitability of each recording. Recordings were further edited to remove noises, breaks or other interruptions, and segments of insufficient quality. Each edited file was renamed with a code number in order to blind the measurer to sex and species. Adult status was, for the purposes of this study, defined with respect to developmental stage rather than behavior. For certain species in our sample (e.g., *Ateles geoffroyi*, *Gorilla gorilla*), we received recordings of adults, sub-adults, and juveniles. For such species, vocalizations from both adults and sub-adults (i.e., individuals who are no longer juveniles though not yet fully adult) were used in the acoustic analysis if sub-adults were sexually mature, because sexually mature individuals would be capable of engaging in competition for mating opportunities and therefore relevant to sexual selection.

Although the acoustic properties of primate calls vary across call types and contexts (13), we chose to utilize measurements across all available call types (but see also Supplementary Results) rather than, for example, selecting only calls believed to be analogous across species, or only those calls shared between males and females. We did this for several reasons. First, averaging across call types should maximize the ability to capture information about the physical properties of the sound source (e.g., vocal fold length and thickness), especially if

some call types may provide more information than others. If F_0 is driven by underlying anatomy—as it is fundamentally—then it should be manifest similarly across call types, even if not specifically adapted to be, because the same anatomy supports the production of all calls. Although it is possible that some other special mechanism of vocal fold action is in play for some calls (e.g., loud calls compared to close calls), this is not a given and certainly cannot be assumed a priori. Second, it is unclear whether call types are truly analogous across species, which complicates comparisons of only a particular call type or set of types. Third, the repertoire shared between males and females can be very limited in some primate species, occasionally down to one call type, as in the orangutan. Finally, given the size of our sample—the largest ever compiled for this type of research, it should also not be specifically biased in any particular way.

Acoustic measurement. Files were measured using the acoustic analysis software Praat version 5.3. The F_0 of every waveform was measured for each file from two segments of 20 cycles (glottal openings and closings) in length where the cycles were clearly discernible. Shorter segments were used when 20 uninterrupted cycles were not found. Of the first-segment measurements, 1102 of 1721 (64%) contained fewer than 20 cycles, and of the files containing another discernible segment (n = 1354), 781 (58%) contained fewer than 20 cycles. Cycles were counted by inspection of the raw waveform, and then divided by the duration of each interval to arrive at the fundamental frequency of the segment. The same procedure was applied to a second, non-overlapping segment of the waveform where one was available (78% of files), and the two measures were averaged to obtain an unweighted mean F_0 value for each file. Each of the two measurements made per file was performed blind to the other measurement. Each mean F_0 value was then averaged with all other mean F_0 values per sex for each species to arrive at separate pitch averages for males and females of all 34 species in our sample (Table S2). Between-segment reliability was high for the 1354 files for which there were two discernible segments (Cronbach's α = 0.973). In addition, the first measurements of a randomly chosen 11% (185 of 1721) of the files were re-measured to determine intra-measurer reliability of the same segment, which was very high (Cronbach's α = 1.000).

Additional data. Body size, habitat, and mating system for each species were obtained chiefly from ref. (1) and other sources when needed (Table S2). Unweighted means were calculated for body weight per sex by averaging values obtained from these sources. Data on mating system were obtained from ref. (9).

Humans are uniquely sexually dimorphic in adiposity among primates (11). To correct for women's greater adiposity when exploring relationships with body mass and sexual dimorphism in body mass, we calculated the average woman's mass assuming an identical body fat percentage to the average man. To do so, we first averaged body fat percentage across foragers, farmers, and industrial societies for men and women separately (6). This gave a female body fat percentage of 28.700, and a male body fat percentage of 17.333. To adjust female mass, we then calculated female fat-free mass (mass - [mass × 0.287] = 44277.7 g) and then added to female fat-free mass (FFM) a fat mass (FM) that would render an equivalent body fat percentage to that of males by solving for FM in the following equation: FM / (FFM + FM) = 0.17333. We thus added 9283.7 g (FM) to 44277.7 g (FFM) to render an adjusted female body mass of 53561.4 g.

Data Analysis. We conducted phylogenetically-informed analyses using a consensus phylogeny for all species in our sample from the 10kTrees website (12) and assessed correlated evolution among our variables with PGLS in the R statistical environment (13). PGLS offers a flexible

model of trait evolution that departs from strict Brownian motion by allowing the λ scaling parameter to vary as a measure of the strength of the phylogenetic signal (14).

Study 2

Two hundred fifty-eight female (mean age ± SD = 20.0 ± 1.6 y) and 175 male (20.1 ± 1.7 y) undergraduate students from Michigan State University participated in this study. Reported ethnicities were 91.5% White, 3.2% Asian, 1.8% Hispanic or Latino, 1.4% Black or African American, 0.5% American Indian or Alaska Native, 0.2% Native Hawaiian or other Pacific Islander, and 1.4% "other". Participants were recorded reading the first six sentences of a standard voice passage ("Rainbow Passage"; 15) in an anechoic, soundproof booth using a Shure SM58 vocal cardioid dynamic microphone (frequency response: 50-15,000 Hz) positioned at approximately 30 degrees and 9.5 cm from the speaker's mouth, and connected to a desktop computer via a Sound Devices USBPre 2 preamplifier. Voices were recorded using Goldwave software in mono at a sampling rate of 44,100 Hz and 16-bit quantization, and saved as uncompressed .WAV files.

We extracted the first sentence of each recording in order to minimize rater fatigue, and adjusted mean amplitude of each to 71.5 ± 2.4 dB to equilibrate audibility, and thus improve the reliability of ratings. Recordings were rated by 558 women (19.1 ± 2.4 y) and 568 men (19.4 ± 1.8 y) at The Pennsylvania State University. Each rater assessed one of 30 stimulus sets comprising approximately 25 voice recordings. Recordings were randomly allocated to a set, with the proviso that only one recording per participant be included in each set. Attractiveness for short- and long-term relationships were rated separately by opposite-sex raters, and dominance was rated for male recordings by male raters, using 7-point Likert scales (7 = very attractive/dominant). The order in which participants completed the rating tasks (e.g., short- vs. long-term first) was random across participants, as was the order in which stimuli were presented. Each stimulus set was rated by ≥ 15 raters (mean = 18.9). The first 15 ratings obtained of each voice stimulus were averaged to produce composite ratings of short- and long-term attractiveness for each recording, and dominance for each male recording. The remaining ratings were discarded. As the ethnic and cultural composition of raters was similar to that of the speakers (both drawn from large universities in the Eastern and Mid-Western US, respectively), and F_0 has been shown to have similar effects across diverse populations (16), we did not use demographic data to exclude speakers or raters from analysis.

Each recording (mean duration ± SD: women = 5.4 ± 0.9 s, men = 5.5 ± 1.0 s) was analyzed using Praat version 5.3. Pitch floors were set to 75 Hz and 100 Hz, and pitch ceilings were 300 Hz and 500 Hz, for men and women, respectively, in accordance with the programmers' recommendations (17). Otherwise, default settings were used. We measured mean F_0, standard deviation in F_0 across the utterance (F_0-SD), duration, number of voice breaks, harmonics-to-noise ratio, four measures of jitter (cycle-to-cycle variation in F_0), and five measures of shimmer (cycle-to-cycle variation in amplitude) using the 'voice report' function in Praat (Table S3). We also measured the first four formant (resonant) frequencies (F_1- F_4, Table S3). Lower, more closely spaced formants correspond with a deeper vocal timbre. Formants were measured at each glottal pulse and averaged across measurements(18). Formant measurements obtained by this method correlate highly ($0.93 \leq r \leq 0.98$) with measurements obtained by measuring and averaging across individual vowels(18). We then computed formant position (P_f), defined as the average standardized formant value for the first four formants(18). The following between-sexes means and SDs were used to standardize formants: F_1 = 482.6 ± 49.8 Hz, F_2 = 1643.2 ± 145.7 Hz, mean F_3 = 2544.7 ± 173.9 Hz and mean F_4 = 3618.8 ± 266.8 Hz.

We also performed separate male and female principal components analyses (with varimax rotation and Kaiser normalization for interpretability) on natural log-transformed measures of jitter, shimmer, harmonics-to-noise ratio ("harmonics"), and voice breaks. In both sexes, all shimmer variables loaded positively (0.77 ≤ loadings ≤ 0.95) and harmonics loaded negatively (-0.76 ≤ loadings ≤ -0.77) onto the first component (Table S6). Consequently, for the purposes of data reduction, we subtracted standardized harmonics from standardized and summed shimmer variables. Likewise, in both sexes, all jitter variables loaded positively (0.85 ≤ loadings ≤ 0.95) onto the second component. Consequently, jitter variables were standardized and summed.

Study 3

Participants were students at The Pennsylvania State University. Fifty-three normally-cycling women (mean age ± SD = 19.4 ± 1.6 y) and 62 men (19.9 ± 2.0 y) were recruited through the psychology department subject pool and received either course credit or US$10 (Sample 1). Reported ethnicities were 68.9% White, 10.6% Black or African American, 8.3% Chinese, 4.5% Hispanic or Latino, 2.3% Asian Indian, 1.5% Filipino, 1.5% Korean, 0.8% Other Asian, and 1.5% Other. Fifty-eight men (19.9 ± 1.2 y; Sample 2) of primarily White ethnicity were recruited from two social fraternities participating in a larger study on male mating competition(19, 20) and were paid US$15. Experiments were undertaken with the understanding and written consent of each subject, and approval of the university ethics committee.

Voices were recorded in an anechoic recording booth in a quiet room (Sample 1) or in a quiet room (Sample 2), with a Shure SM58 vocal cardioid microphone, which was kept approximately 9.5 cm from the participants' mouths by a curved wire projection from the microphone stand. Voices were recorded in mono with a sampling frequency of 44,100 Hz as participants spoke the first six sentences of the Rainbow Passage(15). Recordings were analyzed using Praat version 5.3 for mean F_0 (Table S3). Pitch floors were set to 75 Hz and 100 Hz, and pitch ceilings were 300 Hz and 500 Hz, for men and women, respectively. Otherwise, default settings were used.

To ensure that participants were not taking supplements that might affect hormone concentrations, each participant was asked about his most recent caffeine consumption, current medication, and tobacco use. Participants rinsed their mouths with water 10 minutes before providing each of two saliva samples of 1-2 ml via passive drool. The time between saliva samples was 32.3 ± 10.4 min for Sample 1 women, 30.4 ± 8.6 min for Sample 1 men, and 19.3 ± 6.8 min for Sample 2 men. From each saliva sample, 0.5 ml was aliquotted into a third tube to better capture average hormone levels at the time of participation, rather than peaks or troughs in pulsatile secretion patterns. The combined saliva sample was shaken and then frozen at -20°C until analysis by the Johns Hopkins Center for Interdisciplinary Salivary Bioscience Research (Baltimore, MD) using Salimetrics® kits. Fifty-one women and all men provided sufficient saliva for both cortisol and testosterone assays. Samples were analyzed in duplicate via enzyme immunoassay. Duplicates correlated highly for both cortisol (Sample 1 women: r_{50} = 1.00, $p < 0.0001$; Sample 1 men: $r_{62} = 0.99$, $p < 0.0001$; Sample 2 men: $r_{58} = 0.99$, $p < 0.0001$) and testosterone (Sample 1 women: $r_{53} = 0.99$, $p < 0.0001$; Sample 1 men: $r_{62} = 0.98$, $p < 0.0001$; Sample 2 men: $r_{58} = 0.97$, $p < 0.0001$). Duplicates were consequently averaged, except for three cases in which a female participant had no duplicate and the single assay was used (Table S3). Hormone values were natural log-transformed to correct skew prior to analysis. For cortisol assays, sensitivity is < 0.003 µg/dL, and average intra-assay coefficient of variation is 3.5%. For testosterone assays, sensitivity is < 1.0 pg/mL, and average intra-assay coefficient of variation is 4.6%.

References

1. Rowe N, Myers M, editors. All the world's primates. Charlestown, RI: Primate Conservation, Inc.; 2011.
2. Leutenegger W, Cheverud J. 1982 Correlates of sexual dimorphism in primates: ecological and size variables. *Int. J. Primatol.* **3**:387-402.
3. Weckerly FW. 1998 Sexual-size dimorphism: Influence of mass and mating systems in the most dimorphic mammals. *J. Mammal.* **79**:33-52.
4. Smith RJ, Jungers WL. 1997 Body mass in comparative primatology. *J. Hum. Evol.* **32**:523-59.
5. Lucas PW, Corlett RT, Luke DA. 1986 Sexual dimorphism of tooth size in anthropoids. *Hum. Evol.* **1**:23-39.
6. Pontzer H, Raichlen DA, Wood BM, Mabulla AZ, Racette SB, Marlowe FW. 2012 Hunter-gatherer energetics and human obesity. *PloS One.* **7**:e40503.
7. Mootnick AR, Fan P-F. 2011 A comparative study of crested gibbons (Nomascus). *Am. J. Primatol.* **73**:135-54.
8. Shultz S, Opie C, Atkinson QD. 2011 Stepwise evolution of stable sociality in primates. *Nature.* **479**:219-U96.
9. Myers P, Espinosa R, Parr CS, Jones T, Hammond GS, Dewey TA. 2014 [various dates in 2013] The Animal Diversity Web (online). Available from: http://animaldiversity.org.
10. Geissmann T, Parsons S. 2011 Recording primate vocalizations. In: Curtis JMSDJ, editor. Field and laboratory methods in primatology: A practical guide. New York: Cambridge University Press. p. 287-304.
11. Zihlman AL, McFarland RK. 2000 Body mass in lowland gorillas: a quantitative analysis. *Am. J. Phys. Anthropol.* **113**:61-78.
12. Arnold C, Matthews LJ, Nunn CL. 2010 The 10k Trees Website: A New Online Resource for Primate Phylogeny. *Evol. Anthropol.* **19**:114-8.
13. Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse, W. caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.2 ed2013.
14. Munkemuller T, Lavergne S, Bzeznik B, Dray S, Jombart T, Schiffers K, Thuiller, W. 2012 How to measure and test phylogenetic signal. *Meth. Ecol. Evol.* **3**:743-56.
15. Fairbanks G. Voice and articulation drillbook. 2nd ed. New York: Harper & Row; 1960.
16. Puts D, Doll L, Hill A. 2014 Sexual selection on human voices. In: Weekes-Shackelford V, Shackelford T, editors. Evolutionary Perspectives on Human Sexual Psychology and Behavior: Springer, p. 69-86.
17. Boersma P, Weenik D. 2009 [cited 2009 March 21, 2009] Praat: doing phonetics by computer (Version 5.1.03). Available from: http://www.praat.org/.
18. Puts DA, Apicella CL, Cárdenas RA. 2012 Masculine voices signal men's threat potential in forager and industrial societies. *Proc. Roy. Soc.* B. **279**:601-9.
19. Hill AK, Hunt J, Welling LLM, Cárdenas RA, Rotella MA, Wheatley JR, Dawood K, Shriver MD, Puts DA. 2013 Quantifying the strength and form of sexual selection on men's traits. *Evol. Hum. Behav.* **34**:334-41.
20. Doll LM, Hill AK, Cárdenas RA, Welling LLM, Wheatley JR, Puts DA. 2014 How well do men's faces and voices index mate quality and dominance? *Hum. Nat.* **25**:200-12.