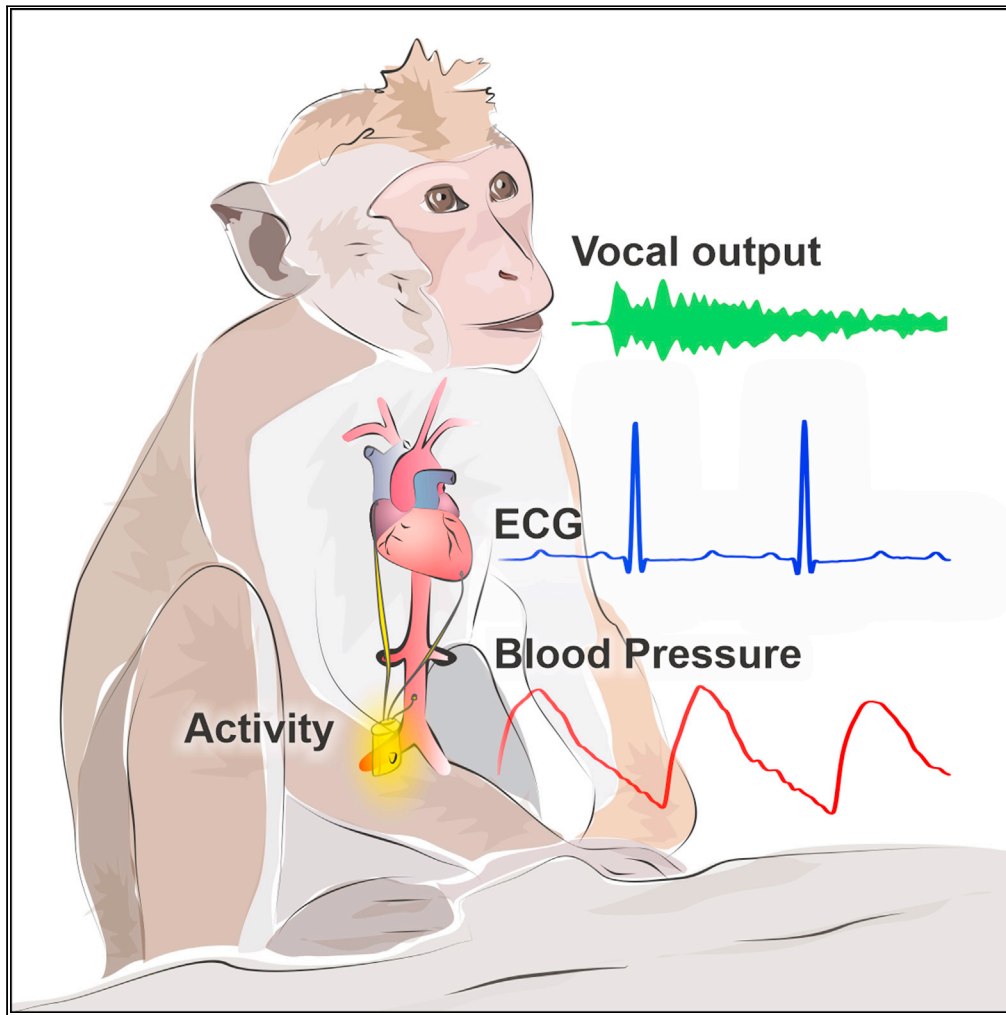


Article

Cardiovascular mechanisms underlying vocal behavior in freely moving macaque monkeys



Cristina Risueno-Segovia, Okan Koç, Pascal Champéroux, Steffen R. Hage

steffen.hage@uni-tuebingen.de (S.R.H.)
cristina.risueno-segovia@cin.uni-tuebingen.de (C.R.-S.)

Highlights

Cardiovascular signals are measured telemetrically in freely moving macaques

A distinct cardiovascular physiological signature is present before vocal onset

Vocal onsets are phase locked to the Mayer wave and heart rate signals

Vocal onsets prediction is performed using machine learning classification models

Risueno-Segovia et al.,
iScience 25, 103688
January 21, 2022 © 2021 The Author(s).
<https://doi.org/10.1016/j.isci.2021.103688>

Article

Cardiovascular mechanisms underlying vocal behavior in freely moving macaque monkeys

Cristina Risueno-Segovia,^{1,2,3,*} Okan Koç,² Pascal Champéroux,⁴ and Steffen R. Hage^{1,2,5,*}

SUMMARY

Communication is a keystone of animal behavior. However, the physiological states underlying natural vocal signaling are still largely unknown. In this study, we investigated the correlation of affective vocal utterances with concomitant cardiorespiratory mechanisms. We telemetrically recorded electrocardiography, blood pressure, and physical activity in six freely moving and interacting cynomolgus monkeys (*Macaca fascicularis*). Our results demonstrate that vocal onsets are strengthened during states of sympathetic activation, and are phase locked to a slower Mayer wave and a faster heart rate signal at ~ 2.5 Hz. Vocalizations are coupled with a distinct peri-vocal physiological signature based on which we were able to predict the onset of vocal output using three machine learning classification models. These findings emphasize the role of cardiorespiratory mechanisms correlated with vocal onsets to optimize arousal levels and minimize energy expenditure during natural vocal production.

INTRODUCTION

Animals continuously communicate with each other (Boë et al., 2019), allowing contact between conspecifics and facilitating development (Gultekin and Hage, 2017; Takahashi et al., 2015) and survival (Perrodin et al., 2011). For instance, vocal communication is important for mating (Nikitopoulos et al., 2004), indicating the presence of predators, and expressing dominance within a group (Balter, 2010). The study of non-human primate vocal systems provides a crucial context for understanding the evolution of human speech (Fitch, 2010; Pomberger et al., 2018; Risueno-Segovia and Hage, 2020) and its associated cardiovascular fluctuations. Despite the deceptive simplicity of natural vocal signaling, it is a complex behavior that carefully coordinates with vital physiological functions, like heartbeat and breathing (Del Negro et al., 2018), with orofacial movements under the influence of emotions (Cameron, 2001) and environmental cues (Barrett and Simmons, 2015).

Mayer waves are spontaneous oscillations of the arterial blood pressure at frequencies lower than respiration, at around 0.1 Hz in humans (Julien, 2006). Rhythmic changes in autonomic nervous system activity, such as changes in Mayer waves, have been proposed to drive vocal production in isolated marmoset monkeys (Borjon et al., 2016). Furthermore, heart rate seems to fluctuate as a function of social distance, with arousal levels decreasing when increasing social interactions (Liao et al., 2018a). These results were then combined in a hypothetical model involving the interaction between the autonomic system and the biomechanics for vocal production during development and adulthood (Zhang and Ghazanfar, 2018, 2020). However, the time frame of the underlying physiological dynamics and their synchronization with vocal affective utterances in freely moving group-housed monkeys remains poorly understood.

A recent study using surface electrodes placed over the chest and back for noninvasive electromyography indicates a potential link between internal states and marmoset vocalizations (Liao et al., 2018a). Here, the animals were separated from the rest of the colony and placed in an experimental setup, heart rate was considered as an indirect measure of arousal and the influence of the physical activity over this variable was disregarded. As a critical advancement over this study, we obtained high-quality physiological measurements of six conscious, freely-moving and naturally interacting cynomolgus monkeys (*Macaca fascicularis*), avoiding stress artifacts that might be induced by handling. To determine the role of cardiorespiratory activity on natural vocal signaling, we performed highly precise internal measurements and

¹Neurobiology of Social Communication, Department of Otolaryngology-Head and Neck Surgery, Hearing Research Centre, University of Tübingen, Medical Center, Elfriede-Aulhorn-Strasse 5, 72076 Tübingen, Germany

²Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Street 25, 72076 Tübingen, Germany

³Graduate School of Neural and Behavioural Sciences-International Max Planck Research School, University of Tübingen, Österberg-Street 3, 72074 Tübingen, Germany

⁴European Research Biology Center, ERBC, Chemin de Montifault, 18800 Baugy, France

⁵Lead contact

*Correspondence: steffen.hage@uni-tuebingen.de (S.R.H.), cristina.risueno-segovia@cin.uni-tuebingen.de (C.R.-S.)
<https://doi.org/10.1016/j.isci.2021.103688>



telemetrically recorded electrocardiography (ECG) and blood pressure (BP) directly from the heart and from the aortic artery, respectively. We calculated the Mayer wave power as a direct measure of arousal. In addition, we collected physical activity with an accelerometer to take into account its influence over the physiological parameters. Implanted radio telemetry allowed for constant real-time monitoring of physiological measurements in naturalistic settings.

We hypothesized that the sensory feedback from the lungs, arteries, larynx, and articulatory systems might carry critical inputs to the primary vocal motor network to adapt the vocal behavior to the current physiological status. In the brainstem, the nucleus of the solitary tract (NST) receives input from these physiological systems that can be transferred to vocal gating (PAG) and vocal patterning structures (parabrachial and reticular regions), which have direct access to all phonatory motoneuron pools. As a consequence, vocalizations during social interactions should be linked to modifications in physiological signals, which can be used to predict the onset of vocal output. Interestingly, our results indicate that vocal onsets are phase locked to both the Mayer wave oscillation and heart rate. Vocalizations were enhanced during states of sympathetic activation and were coupled with a distinct peri-vocal physiological signature. Finally, using three machine learning classification models based on these features from a 2-s window prior to vocal onset, we were able to accurately predict the presence or absence of a call. Furthermore, we propose a neurophysiological network model that links sensory feedback from the cardiorespiratory, laryngeal, and articulatory systems with the vocal pattern generating network, which might explain how distinct arousal states could be integrated into these systems to optimize vocal timing. These findings provide quantitative evidence that cardiovascular variables can reliably predict complex behaviors and raise new questions regarding molecular, cellular, and synaptic mechanisms as key features that may contribute to vocal behavior.

RESULTS

Physiological telemetry

To investigate the physiological states and responses associated with vocalizations uttered during free interactions, we telemetrically recorded physiological signals, including ECG and BP, as well as physical activity, along with the vocal behavior of six freely moving cynomolgus monkeys (*Macaca fascicularis*). The animals were kept in two cages of three males and three females, respectively, with visual and auditory contact between them (Figure 1A). The vocal output and overall behavior were recorded continuously with one microphone and one video camera per cage. Concomitant measures of ECG, BP, and physical activity were collected from wireless telemetry transmitters implanted in the abdominal cavity of the animals (Figure 1B) that monitored the physiological signals that were sent to a nearby receiver (Figure 1A). A total of 486 vocalizations were collected, 229 from the males and 257 from the females (Table S1), including tonal barks, pulsatile grunt bouts, food-related warbles, and screams (Fan et al., 2018; Hauser and Marler, 1993) (Figures 1C and 1D) with a duration of 148.5 ± 4.1 ms (mean \pm SEM; Table S2). Statistical analyses did not reveal significant differences in call duration between individuals or gender (individuals: $p = 0.064$; gender: $p = 0.367$, Kruskal-Wallis test; Figure S1). Therefore, we pooled all vocalizations for further analyses. A high-quality raw signal of the ECG and BP signal sampled at 500 Hz and activity sampled at 1 Hz was obtained from the radio telemeters (Figure 1D). To determine the general dynamics and principles of the arousal states underlying natural vocal production, all vocalizations were included for data analysis regardless of the call type or acoustic features. Focal animal sampling was combined with offline scan sampling from the video recordings to assign specific calls to individual monkeys.

Increases of sympathetic activity have been traditionally considered as measures of autonomic arousal (Weber and Smith, 1990). Therefore, changes in the physiological activity of the autonomic nervous system before, during, and after vocal output should be reflected at the cardiorespiratory level in the heart rate (HR), respiratory amplitude (RA), and respiratory rate (RR) (Figure 2). All of these physiological variables were extracted from the ECG signal (see STAR Methods for details). An acute local increase in HR could be observed in the vocal production signal with respect to its baseline (see STAR Methods for details) (Figure 2A). The greater peri-vocal onset activity points toward a combined vocal motor and non-vocal motor response to either external events or internal physiological states (Figure 2B). A possible explanation for the HR increase might be that calls are associated with higher motor activity, because active movement by itself could induce this effect as well (Obrist, 1981; Smith et al., 1976). Therefore, we defined two control conditions to compare with the signals associated with vocal production to disentangle vocalization-correlated activity from general motor activity: control rest and control active. The control rest condition

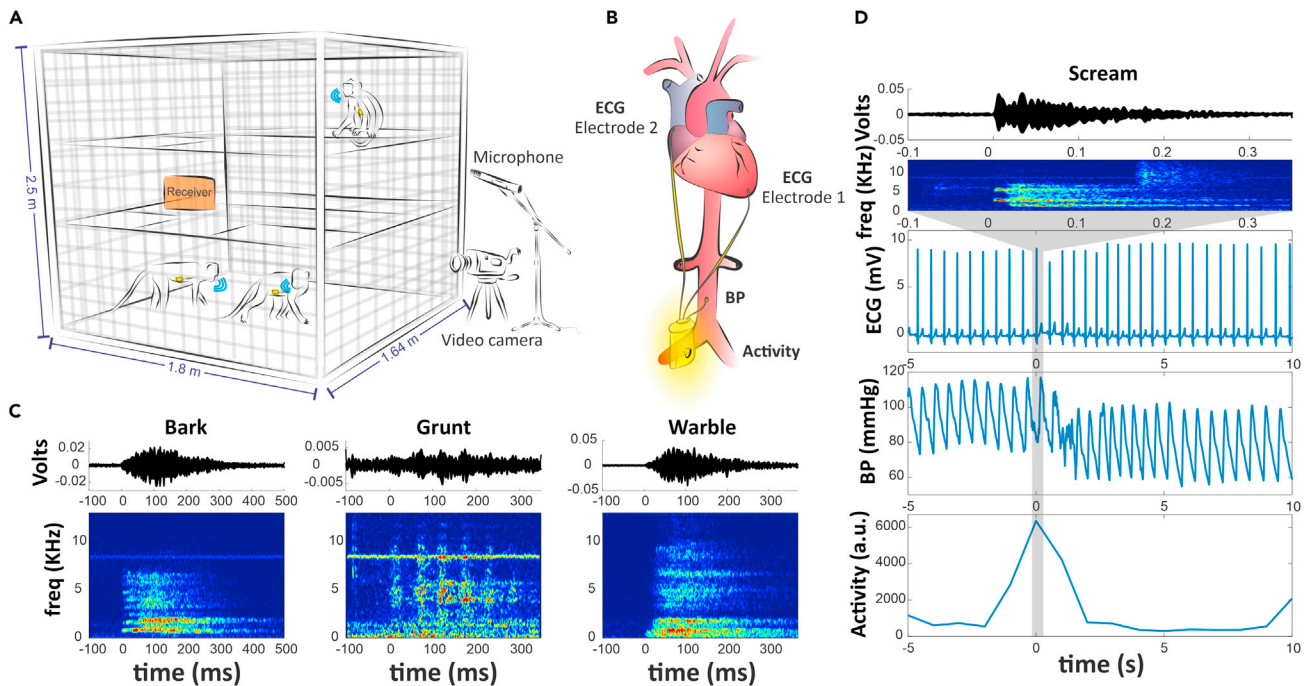


Figure 1. Telemetry recordings of ECG, BP, and physical activity in six freely moving vocalizing monkeys

(A) Schematic representation of one of the recording setups. Implanted wireless transmitters monitored physiological signals while the animals were freely moving in two cages (three males or three females each) with visual and auditory contact between each other. Audio and video recordings were performed simultaneously to collect the vocal output of the animals and identify the caller, respectively. The dimensions of the cages are indicated. (B) Schematic of the implanted radio telemeter and electrodes in the abdominal cavity of the animals to collect ECG, BP, and physical activity data. (C) Oscillograms (top) and spectrograms (bottom) of representative vocalizations: the tonal bark, pulsatile grunt bout, and food-related warble. (D) Vocalization-correlated ECG, BP, and physical activity signals showing the oscillogram (top) and spectrogram (upper middle) of an exemplar vocalization with the accompanying raw ECG (upper middle), BP (lower middle), and physical activity signals (bottom). See also [Tables S1](#) and [S2](#).

consisted of ten periods of at least 5 s in which a specific animal was neither actively moving nor vocalizing. The control active condition consisted of ten periods of at least 5 s in which a specific animal was actively moving but not vocalizing.

The mean HR including all the animals was approximately 15% higher at vocal onset times than at control rest levels and 9% higher than at control active levels ([Figure 2A](#)). The HR of the animals before vocal onset times was at least as high as in control active conditions, indicating that a certain degree of arousal might be required in order to vocalize or might be generally linked to a potential vocal window. This effect could also be observed in the power spectral density of both the ECG and BP signals at ~ 2.5 Hz ([Figures S2A](#) and [S2B](#)).

We hypothesize that the acute increase in HR is directly related to vocal production and cannot be purely explained by physical activity levels or non-vocal active movement. To disambiguate the potential link between the acute increase in HR at peri-vocal onset and activity levels, we divided the vocal production signal into high, intermediate, and low activity and examined the respective HR ([Figure S3](#)). This analysis aims to decipher how much different levels of physical activity might be able to influence the heart rate signal during vocal production. An acute increase in HR locked to vocal onset was observed regardless of physical activity levels. Furthermore, the time to peak HR for the low activity condition was shorter (~ 2 – 3 s) than the time to peak HR for the intermediate and high activity conditions (~ 4 – 5 s). The former might mainly reflect the HR change associated with vocalizations, whereas the latter may represent the summation of both vocal motor and non-vocal motor responses.

Another possible cause of the acute HR variation during vocal production could be respiration. According to the respiratory sinus arrhythmia (RSA), the breathing act triggers an acceleration of the HR during inspiration (because of a decrease in vagal efferent activity) and a decrease in HR during expiration ([Bernardi](#)

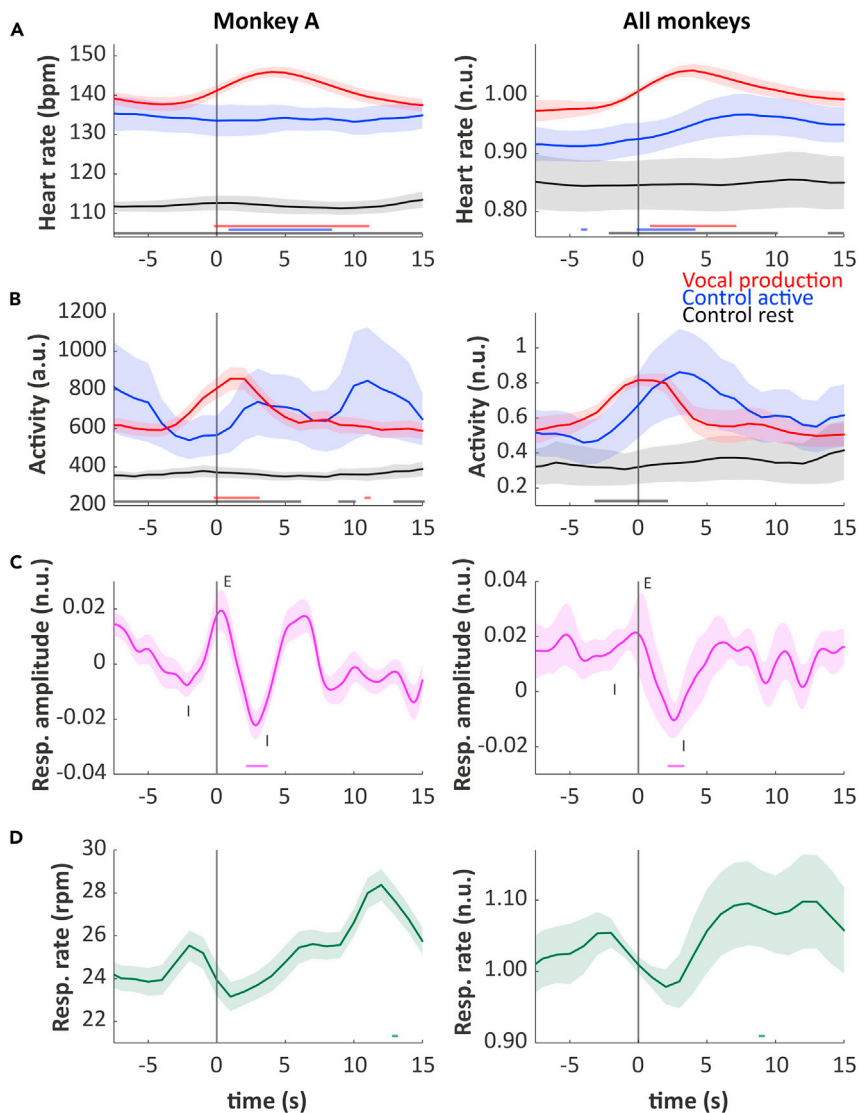


Figure 2. Peri-vocal physiological responses of the autonomic nervous system

(A) There is an increase in heart rate during vocal production compared to its value at baseline and to the control rest and control active conditions.

(B) Vocal production times are related to higher motor activity than that during the control rest condition.

(C) Respiratory amplitude showing acute inspiratory (I) and expiratory (E) phases.

(D) Respiratory rate decreases during vocal production despite the higher activity levels. Physiological signals from an exemplar monkey are depicted on the left and the mean from all monkeys on the right. Resp. amplitude: respiratory amplitude, Resp. rate: respiratory rate, bpm: beats per minute, rpm: respirations per minute, a.u.: arbitrary units, n.u.: normalized units. Vocal onset times are indicated with a vertical line at $t = 0$. The vocal production condition is indicated in red, control active condition in blue, and control rest condition in gray. Significant levels ($p < 0.05$) are color-coded and displayed below the respective signal (Wilcoxon rank-sum test for unmatched groups between vocal production and the control active and control rest conditions in blue and gray, respectively, and signed-rank for matched groups comparing the vocal production signal with its baseline in red). Mean \pm SEM are displayed. See also [Figures S3](#).

[et al., 2001](#); [Berntson et al., 1993](#); [Dick et al., 2014](#); [Eckberg and Sleight, 1992](#); [Grossman and Taylor, 2007](#)).

In this case, the increase in HR before vocal onset could be linked to the inspiratory phase of the respiratory cycle, as opposed to the further increase in HR after vocal onset, which cannot be explained by the expected decrease related to the expiratory phase ([Figure 2C](#)). Note that the inspiratory and expiratory phases of the respiratory cycle could be identified from the raw RA signal without previous alignment of

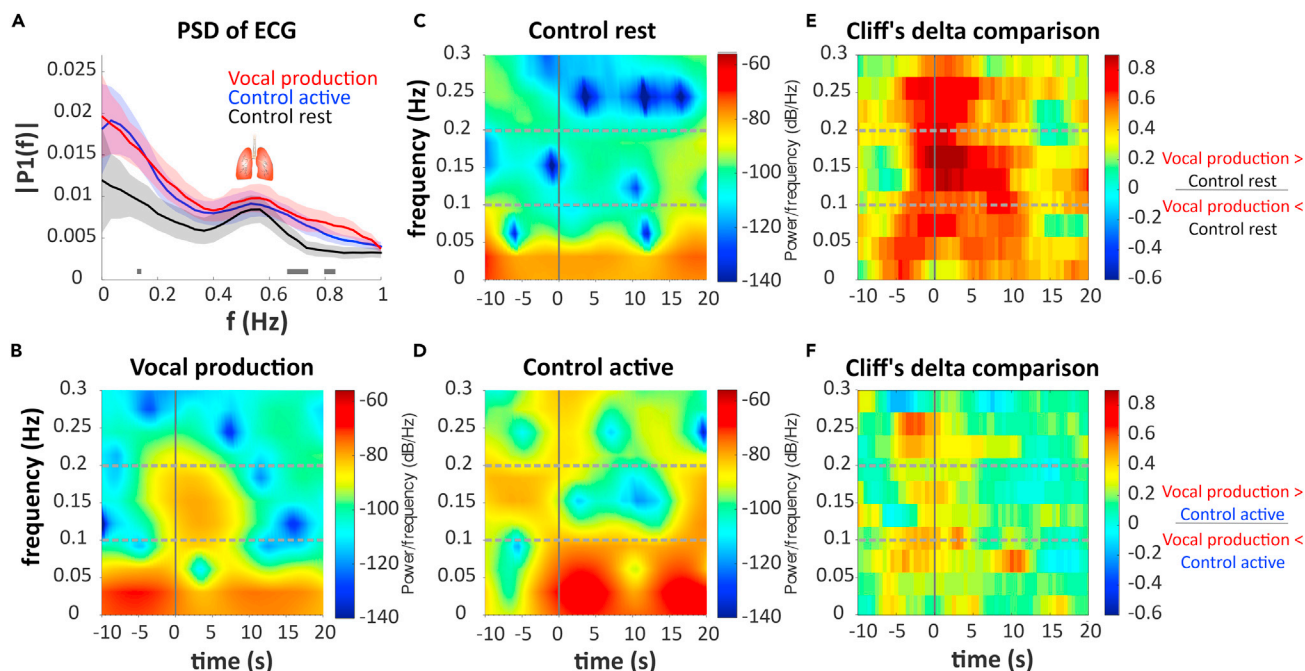


Figure 3. Increase in Mayer wave power linked to vocal production

(A) Mean power spectral densities of the ECG signal from all animals show higher gains at Mayer wave frequencies in the vocal production and control active conditions than in the control rest condition. Vocal production signal is indicated in red, control active in blue, and control rest in gray. Mean \pm SEM are depicted. Significant levels ($p < 0.05$) are color-coded and displayed below the respective signal (Wilcoxon rank-sum test for unmatched groups between vocal production and controls). Schematic lungs illustrate the respiratory rate.

(B) Mean spectrogram with distinct increase in Mayer wave power at 0.1–0.2 Hz during vocal production starting a few seconds before vocal onset.

(C) Mean spectrogram of control rest condition.

(D) Mean spectrogram of control active condition.

(E) Cliff's delta comparison of vocal production and control rest conditions.

(F) Cliff's delta comparison of vocal production and control active conditions. The large effect size boundary was set at 0.474. Vocal onset times are indicated with a vertical line at $t = 0$. Mayer wave band is outlined by dashed horizontal lines in gray (0.1–0.2 Hz). See also [Figures S2](#) and [S4](#).

individual vocal onsets at expiratory phases during data processing. At a small scale, the respiratory rate reflects the act of vocalizing with a higher rate a few seconds before vocal onset and a decrease during vocal production despite the high concurrent physical activity levels ([Figure 2D](#)). In a broader sense, the difference between the respiratory rate during vocal production, control rest, and control active is rather small ([Figures S2A](#) and [S2B](#), [3A](#) and [S4A](#), peak at around 0.5 Hz). Overall, these results show a distinct physiological pattern accompanying vocal output, high HR baseline levels prior to vocal onset ([Figure 2A](#)), acute increase of HR starting a few seconds before vocal onset reaching a maximum peak after approximately two to four seconds ([Figures 2A](#) and [S3B](#)), and high activity levels at peri-vocal onset ([Figure 2B](#)), that cannot fully explain the acute increase in HR ([Figure S3](#)). In addition, vocal onset is coupled to the expiratory phase of the respiratory cycle ([Figure 2C](#)), and there is a decrease in the RR at peri-vocal onset starting a few seconds before vocal onset despite the simultaneous increase in HR and activity ([Figures 2A](#), [2B](#) and [2D](#)). Note that there is an increase in duration and amplitude of the respiratory cycle during vocal production in contrast to consecutive faster cycles with shorter durations and lower amplitudes ([Figures 2C](#) and [2D](#)).

Mayer wave and heartbeat link to vocal production

Spontaneous calls have been proposed to be related to arousal states ([Borjon et al., 2016](#)). Generally, arousal states have been associated with Mayer waves—spontaneous oscillations of arterial pressure occurring in conscious subjects at a frequency lower than respiration (0.1 Hz in humans) ([Julien, 2006](#)). Based on their postulated origin, Mayer waves may be easier to identify in BP signals. However, because of the high noise of the BP signal at very low frequencies we were unable to detect Mayer wave frequencies in our BP dataset ([Figure S2B](#)). Nevertheless, the power spectral densities of the ECG and BP signals at peri-vocal onset were highly similar ([Figures S2A](#) and [S2B](#)) and exhibited a significant correlation ([Figure S3C](#); BP vocal

production: $p = 4.88e-34$, $R = 0.91$, BP control active: $p = 3.16e-51$, $R = 0.97$, BP control rest: $p = 1.99e-58$, $R = 0.98$; ECG vocal production: $p = 2.44e-65$, $R = 0.98$, ECG control active: $p = 5.33e-64$, $R = 0.98$, ECG control rest: $p = 4.88e-55$, $R = 0.97$; Pearson's correlation). Consequently, we used the ECG signal to identify Mayer waves and its correlation with vocal behavior in the subsequent analyses.

The power spectra of the ECG signal exhibit a well-defined peak at breathing frequency (~ 0.5 Hz) reflecting an index of cardiac vagal function (Yasuma and Hayano, 2004) during respiratory-circulatory interactions or RSA (Figures 3A and S4A). In addition, a lower frequency peak at around 0.15 Hz is evident and reflects higher cardiac sympathetic activity during vocal production and in the control active condition than in the control rest condition. This might indicate that higher demands on the sympathetic nervous system derived from motor movements involving higher gains at lower frequencies (Wienecke et al., 2015). Although this analysis reveals that Mayer wave frequencies exhibit a gain around vocal onset, the temporal resolution is quite low in this approach because the signal duration that allows the identification of low frequencies between 0.05 and 0.1 Hz in a power spectrum have to be at least 20 to 10 s, respectively (in this case, 15 s before and after vocal onset) (Gabbiani and Cox, 2010). Therefore, we calculated the spectrogram of the ECG signal during vocal production and in the control conditions by adding time as a third dimension to be able to precisely relate Mayer wave power with vocal onset (Figures 3B, 3C, 3D and S4B, S4C, S4D). The spectrogram of vocal production exhibited a distinct, sharp enhancement in Mayer wave power (0.1–0.2 Hz) upon vocal production, starting a few seconds before vocal onset (Figures 3B and S4B). This activity was absent in the control rest and control active conditions (Figures 3C, 3D, S4C, and S4D). To quantify these differences, we computed Cliff's delta metric. Briefly, the d-metric describes the effect size of group comparisons with values ranging from -1 to 1, with identical groups rendering values around zero (Romano et al., 2006; Weinecke et al., 2020). Cliff's delta matrices revealed higher power in the vocal production condition compared to that in control conditions, starting a few seconds before call onset (Figures 3E, 3F, S4E, and S4F). These results indicate that the Mayer wave oscillations are correlated with the vocal behavior in cynomolgus monkeys.

Next, we bandpass filtered the ECG and BP signals between 0.1 and 0.2 Hz to identify whether vocal onsets were synchronized to the temporal structure of the observed Mayer wave frequencies (Figure 4A). For this purpose, we computed the Hilbert transform to calculate the phase angle at call onset. Our analysis revealed that vocal onsets are significantly phase locked to the observed Mayer wave frequency band, taking place predominantly at local maxima within the oscillatory signals (ECG: $p = 8.19e-11$, $n = 486$, $z = 22.97$; BP: $p = 6.86e-5$, $n = 486$, $z = 9.55$; Rayleigh test; Figures 4C and 4E). Furthermore, we performed a similar analysis for the raw ECG and BP signals (Figure 4B). Surprisingly, aside from being phase locked to Mayer wave cycles, we observed that vocal onsets were also significantly phase locked to the HR cycles of around 400 ms (ECG: $p = 2.19e-23$, $n = 486$, $z = 50.82$; BP: $p = 1.95e-3$, $n = 486$, $z = 6.21$; Rayleigh test; Figures 4D and 4F), which possess a much faster frequency of approximately 2.5 Hz (Figures S2A and S2B). The phase lag between the ECG and BP measurements can be explained by the different recording locations, that are in the heart and in the abdominal aorta near the common iliac arteries, respectively (Figure 1B; see STAR Methods for details). Finally, we also calculated the phase angle of the respiratory amplitude signal at vocal onset and verified our previous results (Figure 2C) indicating that vocal onsets take place predominantly during expiration ($p = 5.74e-5$, $n = 486$, $z = 9.73$; Rayleigh test; Figure S5). The current findings emphasize sympathetic activation preceding vocal production during social interactions and establish a direct link between the amplitude of the ECG and BP signals at specific frequencies and call initiation.

Predictive power of 3 machine learning classification models to indicate the presence or absence of a call

Finally, we trained three machine learning classification models, logistic regression, neural networks, and support vector machines (SVM), to show the predictive power of certain physiological features (Figure 5 and Table S3). We predicted the presence or absence of a call from the mean physiological values 2 s before vocal onset, including heart rate, respiratory rate, respiratory amplitude, blood pressure, and Mayer wave power as custom physiological parameters. We compared the physiological parameters linked to the vocal production signals to the physiological parameters linked to our combined controls: control active and control rest conditions. Hereby, we reserved 11 vocalizations and 22 control data points for the evaluation of the trained classifiers. The rest of the data points were used for the training procedure. We obtained a highest mean predictive accuracy of 85% for monkey A and 86% for all monkeys with the SVM.

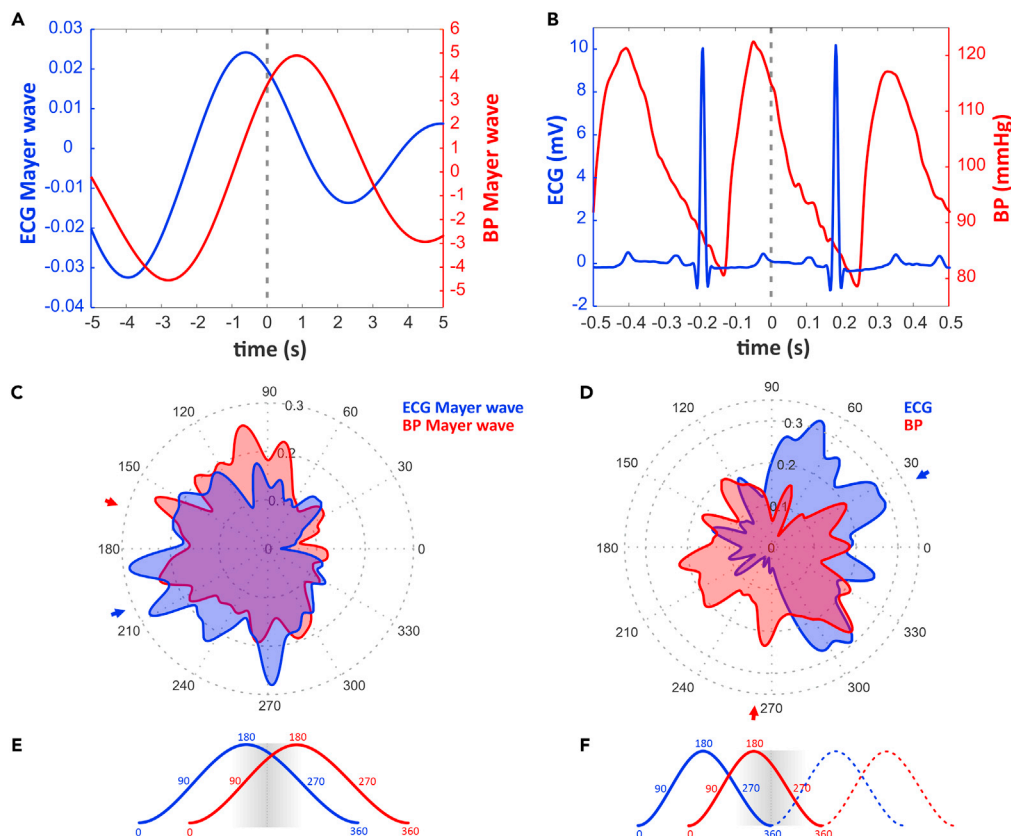


Figure 4. Vocal onsets are phase locked to a slower Mayer wave frequency and to a faster heart rate in the ECG and BP signals

(A and B) (A) Vocal onset example in relation to ECG and BP measurements bandpass filtered at observed Mayer wave frequencies (0.1–0.2 Hz) and to (B) raw ECG and BP measurements. Vertical dashed lines indicate vocal onsets. (C and D) (C) Polar histograms show vocal onset phases within the ECG and BP signals band-pass filtered at observed Mayer wave frequencies (0.1–0.2 Hz) and within (D) raw ECG and BP measurements. The height indicates the relative number of observations within the respective bin in the polar plot. Color-coded arrows indicate the corresponding mean angles. Phase angles of 0° and 360° relate to a minimum and 180° to a maximum amplitude in the oscillatory signals. (E) Schematic representation of the phase angles within ECG and BP signals band-pass filtered at observed Mayer wave frequencies (0.1–0.2 Hz) and (F) raw ECG and BP signals.

Neural networks also had predictive powers significantly above chance, 79% for monkey A and 81% for all monkeys. Logistic regression did not render predictive values above chance. We observed that the non-linear classifiers, SVM and neural networks, are able to better capture these sophisticated physiological features than the linear classifier, logistic regression. Overall, these results point toward a precise physiological signature already present before vocal production.

DISCUSSION

Our findings demonstrate that natural vocalizations in cynomolgus monkeys are coupled with a distinct peri-vocal physiological signature characterized by increases in HR, motor activity, and Mayer wave power. In addition, the spectrogram of vocal production displayed a distinct vocalization-correlated amplification in Mayer wave power at ~0.15 Hz starting a few seconds before calling. Furthermore, vocal onsets were phased locked to the ECG and BP signals at Mayer wave frequencies (0.1–0.2 Hz), as well as to the regular ECG and BP cycles derived from the HR (~2.5 Hz). Therefore, our data suggest a direct influence of cardiac sympathetic activity at the frequency of Mayer waves and heart cycles on vocal behavior. Finally, based on the precise pattern of the physiological measures, we could predict the presence or absence of a call from the mean physiological values within 2 s before vocal onset using three machine learning classification models.

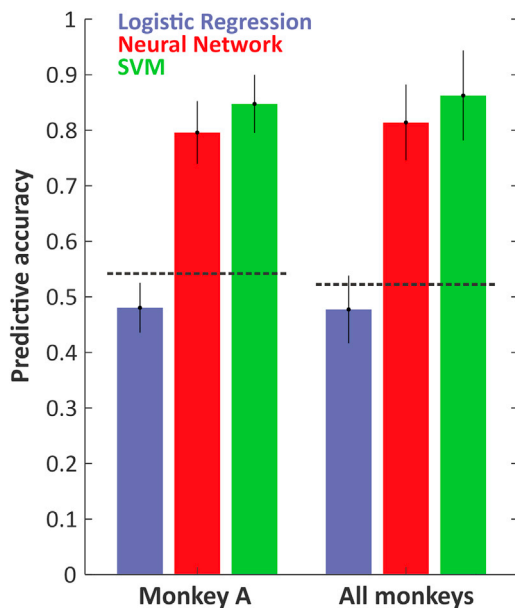


Figure 5. Predictive accuracy of physiological features based on three machine learning classification models

Predictive power of the mean levels of physiological signals (heart rate, respiratory rate, respiratory amplitude, blood pressure, and Mayer wave power) within 2 s before call onset for the presence or absence of a call. Logistic regression is indicated in blue, neural network in red, and support vector machine (SVM) in green. Black dashed lines indicate significance levels. Mean \pm STD are displayed.

Wireless recording techniques enabled us to reliably monitor and investigate the physiological states associated with the vocal production in spontaneously behaving cynomolgus macaques. High-quality internal measurements revealed that physiological states exhibited distinct modulations reflecting the interaction between the cardiovascular system and natural calling. A precise physiological pattern was observed at perivocal onsets including an increase in HR, physical activity, and Mayer wave power. Apart from the acute increase at vocal onset, these variables were also higher than those in control rest and control active situations.

Despite their enigmatic origin, Mayer waves are widely believed to be a measure of arousal levels mediated by changes in sympathetic activity (Julien, 2006). The related baroreflex loop is a homeostatic mechanism that aims to maintain blood pressure levels. Each heartbeat is followed by a transient increase in BP evoking parasympathetic activity to decrease HR (Reilly and Moore, 2003). Perturbations of this system are considered to be one of the main hypotheses for the origin of Mayer waves because they generate oscillations at a resonance frequency of ~ 0.1 Hz (Julien, 2006). The mechanisms behind cardiorespiratory variability have been studied extensively in non-human animal models (Eckberg and Sleight, 1992). The frequency of the Mayer wave is ~ 0.1 Hz for both humans and cats and ~ 0.4 Hz for rats and mice (Julien, 2006). Common marmosets also possess Mayer wave frequencies at ~ 0.1 Hz (Borjon et al., 2016), and according to our results, the Mayer wave frequency of cynomolgus macaques is in the same range, ~ 0.15 Hz. Therefore, animal species with different body sizes as humans, macaques, and marmosets appear to have a similar Mayer wave frequency. Accordingly, Mayer wave frequencies seem to be a species-specific feature rather than a size-specific trait. Although fluctuations in the cardiorespiratory system and the underlying sympathetic nervous system have been predominantly related to respiratory and locomotor activities (Wienecke et al., 2015), their role in vocal signaling is yet largely unknown.

In relation to previous results in primates, physical distance between marmoset monkeys is considered to be a determinant of arousal level, with social distancing increasing arousal (Liao et al., 2018b). The previous study, which measured HR using external electromyography, failed to identify a significant coherence between HR and vocal production at Mayer wave frequency when a marmoset was next to another. With our present work, we were able to measure the Mayer wave in naturally roaming and socially interacting cynomolgus macaques and found that vocalizations were phase locked to it. This indicates that the use of internal wireless measurements might be ideally suited to detect internal arousal rhythms in a social context. Liao et al. observed that the heart rate was 1 s before call onset, generally higher than the mean HR for the session. One interpretation of their findings could be that in their experimental context, external factors might have triggered an increase in HR and consequently also in arousal, being accompanied by a subsequent vocalization. In another marmoset study (Borjon et al., 2016), socially isolated marmosets displayed

an increase in HR 23 s before call onset with a rapid decrease after calling and reduced activity starting ~3 s before calling (Borjon et al., 2016). In contrast, our results show an acute increase in HR starting a few seconds before vocal onset with a maximum peak ~4 s after vocal onset, followed by a cardiac deceleration reaching baseline levels ~10 s after vocal onset, as well as an increase in activity starting ~3 s before calling. These discrepancies might be mainly because of the differences in the experimental model (marmosets vs. cynomolgus monkeys) and approach (external electromyography and activity inferred from video recordings vs. internal telemetry). However, the observed differences could also be related to the different recording contexts (socially isolated monkeys vs. freely moving animals), which might point to different cardiovascular mechanisms underlying different social contexts.

In addition, our wireless telemetry recordings of ECG and BP at 500 Hz provided continuous signals for Fourier analysis methods, as opposed to traditional calculations from discontinuous signals (Cooke et al., 1999; DeBoer et al., 1987). These computations allowed us to observe both the influence of the fast parasympathetic nerves at respiratory rate frequencies and slower sympathetic nerves at Mayer wave frequencies on vocal output. Increased heart rate during speech has been observed previously as well (Peters and Hulstijn, 1984; Reilly and Moore, 2003; Weber and Smith, 1990) and has been related to either a decrease in parasympathetic output or increase in sympathetic output. In humans, changes in the tonic levels of cardiac sympathetic and parasympathetic drive are difficult to segregate and therefore their contribution to the mean HR increase during speech remains speculative (Reilly and Moore, 2003). According to our results, vocal onsets are enhanced during states of sympathetic activation and phase locked to a slow and transient Mayer wave frequency at ~0.15 Hz and to a fast and steady heart rate signal at ~2.5 Hz. The present findings might be indicative of an increase in sympathetic drive, including sympathetic output to the heart. Here, reducing the sympathetic efferent discharge through beta-adrenergic blocking agents such as atenolol could help in future studies to clarify how much sympathetic activity contributes to vocal production. Along similar lines, the absence of RSA linked to vocal signaling reflects low cardiac vagal outflow during the expiratory phase of the respiratory cycle associated with vocal production. Furthermore, similar increases in respiratory duration related to an increase in lung volume during vocal production have been previously noted during speech (Reilly and Moore, 2003). Increased blood pressure has also been linked to speech (Lynch et al., 1980). We did not observe any similar effects.

Sensory feedback from the cardiorespiratory, laryngeal, and articulatory systems into the vocal pattern generator might be of utmost importance to align vocal onsets to the existing physiological state and optimize arousal levels. Based on our physiological measures, we propose a neuro-physiological network that could integrate somatosensory information from the baroreceptors of the lungs as well as the mechanoreceptors of the larynx and supralaryngeal structures into the primary vocal motor network (Figure 6). In this model, the NST is a crucial relay station, because it gets information about the blood pressure from both the carotid sinus baroreceptors via the glossopharyngeal nerve (cranial nerve IX) and the aortic arch baroreceptors via the vagus nerve (cranial nerve X). For example, when the arterial baroreceptors are stretched and stimulated, the baroreceptor reflex is activated, triggering an inhibition of the sympathetic discharge (Lipski and Trzebski, 1975). In addition, NST receives somatosensory information of the current air volume in the lungs from the pulmonary stretch receptors via the vagus nerve. This input seems to be crucial for properly gating call onsets in accordance with the momentary respiratory status (Jürgens, 2002). Furthermore, the laryngeal and articulatory proprioceptive and tactile information from most laryngeal and articulatory muscles are also transferred to NST via the vagus nerve (Altschuler et al., 1989; Hayakawa et al., 2001; Miffilin, 1993; Patrickson et al., 1991; Travers and Norgren, 1995). The NST itself has strong connections to several relay stations of the vocal motor network in the brainstem (Hage and Nieder, 2016) such as the PAG (Bandler and Tork, 1987; Keay et al., 1997; Mantyh, 1982), the ventrolateral parabrachial region as well as the reticular formation encompassing parts of the putative vocal pattern generator (Beckstead et al., 1980; Ezure et al., 1998; Fort et al., 1994; Herbert et al., 1990; Slugg and Light, 1994), which has direct access to all phonatory motoneuron pools (Mantyh, 1983; Meller and Dennis, 1991). Our findings suggest that the somatosensory inputs derived from the cardiorespiratory, phonatory, and articulatory systems are key aspects for gating vocal initiation being essential for the optimization of the cardiorespiratory status during natural vocal production to minimize energy expenditure. Future neurophysiological studies in interacting monkeys will now have to elucidate the role of the NST in primate vocal communication.

Based on the physiological features from a 2-s window before vocal onset, we were able to predict the presence or absence of a call with a mean accuracy of 86% using an SVM classifier in comparison with combined

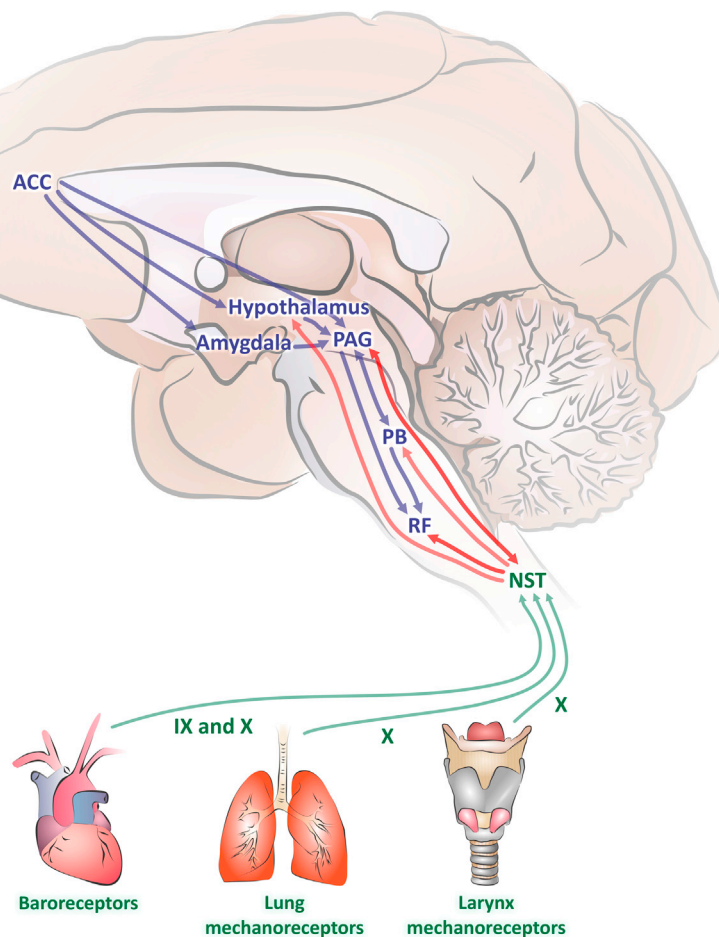


Figure 6. Hypothetical model of an arousal gated vocal neurophysiological network

Simplified circuit summarizing the most relevant structures that integrate the somatosensory feedback derived from the baroreceptors of the lungs as well as the laryngeal mechanoreceptors into the primary vocal motor network (Hage and Nieder, 2016). Blue arrows indicate internal connections within the primary vocal motor network. Green arrows indicate the sensory input to the NST. Red arrows indicate the somatosensory feedback from the cardiorespiratory and laryngeal systems into the primary vocal motor network containing the vocal pattern generator. Abbreviations: ACC, anterior cingulate cortex; PAG, periaqueductal gray; PB, parabrachial nucleus; RF, reticular formation; NST, nucleus of the solitary tract; IX, sensory glossopharyngeal nerve; X, sensory vagus nerve.

controls, control rest, and control active conditions. Our data suggest that physiological parameters can be used as a predictor of vocal motor output, indicating a direct effect of these parameters on vocal behavior. These findings provide a new perspective on how physiological parameters correlate with natural vocal utterances by highlighting a key physiological link between natural vocal production and the cardiovascular system, enabling us to address new questions on volitional vocal control abilities among non-human primates that go beyond arousal-based vocal motor control mechanisms. Recent studies have observed that non-human primates are able to call on command in operant conditioning experiments, indicating the capability of these animals to decouple their vocal utterances from the underlying state of arousal and to use them as a response to abstract sensory cues (Hage and Nieder, 2013; Pomberger et al., 2019). However, the different levels at which the human brain can take control over the speech system are far beyond the control mechanisms underlying vocal production in non-human primates (Hage and Nieder, 2016). We hypothesize that cognitively controlled vocalizations such as human speech might be predicted with a lower accuracy than natural monkey calls that convey a higher emotional content. Accordingly, human speech with a high emotional content should increase the predictive accuracy of speech onset, because the cardiorespiratory system is strongly affected by emotions (Prkachin et al., 1999; Sinha et al., 1992). On the other hand, recent studies in rhesus macaques and marmosets have shown that

monkeys are capable of volitionally vocalizing in response to arbitrary visual cues (Hage and Nieder, 2013; Pomberger et al., 2019). Further studies will now have to elucidate how monkeys are able to volitionally control this motivationally driven behavior, if this has a direct effect on cardiorespiratory activity and whether human speech signals can be predicted by quantitative measures of the cardiorespiratory system.

Limitations of the study

One limitation of our study is that the cardiorespiratory responses we are analyzing are not exclusively correlated to the vocal output, because they are vitally important for all biological processes and could be influenced by, e.g., levels of relaxation, anxiety, or attention. In addition, we grouped all the calls together regardless of acoustic features and, calls with, e.g., longer durations and intensities might be more strongly correlated with the underlying physiology.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Cynomolgus monkeys
- METHOD DETAILS
 - Implantation of telemetry system
 - Data analysis of telemetry
 - Recording of the animal's activity
 - Vocal and video recording setup
 - Machine learning classifiers
 - Analyses of machine learning classifiers
 - Data analysis of physiological signals
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.103688>.

ACKNOWLEDGMENTS

We thank John Holmes for proofreading. This work was supported by the Werner Reichardt Centre for Integrative Neuroscience (CIN) at the Eberhard Karls University of Tübingen (CIN is an Excellence Cluster funded by the Deutsche Forschungsgemeinschaft within the framework of the Excellence Initiative EXC 307).

AUTHOR CONTRIBUTIONS

C.R.-S. and S.R.H. conceived the study, designed the experiments, and interpreted the data; C.R.-S. conducted the experiments, created the visualizations, and drafted the manuscript; C.R.-S. performed the data analyses on the physiological parameters and O.K. performed the data analyses on the machine learning classifiers; P.C. implanted the telemeters and provided the animals; S.R.H. and P.C. acquired funding; S.R.H. supervised the project; all authors commented and approved the final version of the paper.

DECLARATION OF INTERESTS

None of the authors have any competing financial interests.

Received: August 10, 2021

Revised: December 1, 2021

Accepted: December 22, 2021

Published: January 21, 2022

REFERENCES

- Altschuler, S.M., Bao, X., Bieger, D., Hopkins, D.A., and Miselis, R.R. (1989). Viscerotopic representation of the upper alimentary tract in the rat: sensory ganglia and nuclei of the solitary and spinal trigeminal tracts. *J. Comp. Neurol.* 283, 248–268.
- Balter, M. (2010). Animal communication helps reveal roots of language. *Science* 328, 969–971.
- Bandler, R., and Tork, I. (1987). Midbrain periaqueductal grey region in the cat has afferent and efferent connections with solitary tract nuclei. *Neurosci. Lett.* 74, 1–6.
- Barrett, L.F., and Simmons, W.K. (2015). Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 419–429.
- Beckstead, R.M., Morse, J.R., and Norgren, R. (1980). The nucleus of the solitary tract in the monkey: projections to the thalamus and brain stem nuclei. *J. Comp. Neurol.* 190, 259–282.
- Bernardi, L., Porta, C., Gabutti, A., Spicuzza, L., and Sleight, P. (2001). Modulatory effects of respiration. *Auton. Neurosci.* 90, 47–56.
- Berntson, G.G., Cacioppo, J.T., and Quigley, K.S. (1993). Respiratory sinus arrhythmia: autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology* 30, 183–196.
- Boë, L.J., Sawallis, T.R., Fagot, J., Badin, P., Barbier, G., Captier, G., Ménard, L., Heim, J.L., and Schwartz, J.L. (2019). Which way to the dawn of speech?: reanalyzing half a century of debates and data in light of speech science. *Sci. Adv.* 5, eaaw3916.
- Borjon, J.I., Takahashi, D.Y., Cervantes, D.C., and Ghazanfar, A.A. (2016). Arousal dynamics drive vocal production in marmoset monkeys. *J. Neurophysiol.* 116, 753–764.
- Cameron, O.G. (2001). Interoception: the inside story - a model for psychosomatic processes. *Psychosom. Med.* 63, 697–710.
- Champéroux, P., Fesler, P., Judé, S., Richard, S., Le Guennec, J.Y., and Thireau, J. (2018). High-frequency autonomic modulation: a new model for analysis of autonomic cardiac control. *Br. J. Pharmacol.* 175, 3131–3143.
- Combrisson, E., and Jerbi, K. (2015). Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J. Neurosci. Methods* 250, 126–136.
- Cooke, W.H., Hoag, J.B., Crossman, A.A., Kuusela, T.A., Tahvanainen, K.U.O., and Eckberg, D.L. (1999). Human responses to upright tilt: a window on central autonomic integration. *J. Physiol.* 517, 617–628.
- DeBoer, R.W., Karemaker, J.M., and Strackee, J. (1987). Hemodynamic fluctuations and baroreflex sensitivity in humans: a beat-to-beat model. *Am. J. Physiol.* 253, H680–H689.
- Del Negro, C.A., Funk, G.D., and Feldman, J.L. (2018). Breathing matters. *Nat. Rev. Neurosci.* 19, 351–367.
- Dick, T.E., Hsieh, Y.H., Dhingra, R.R., Baekey, D.M., Galán, R.F., Wehrwein, E., and Morris, K.F. (2014). Cardiorespiratory coupling: common rhythms in cardiac, sympathetic, and respiratory activities. *Prog. Brain Res.* 209, 191–205.
- Eckberg, D.L., and Sleight, P. (1992). *Human Baroreflexes in Health and Disease* (Oxford University Press).
- Ezure, K., Tanaka, I., and Miyazaki, M. (1998). Pontine projections of pulmonary slowly adapting receptor relay neurons in the cat. *NeuroReport* 9, 411–414.
- Fan, P., Liu, X., Liu, R., Li, F., Huang, T., Wu, F., Yao, H., and Liu, D. (2018). Vocal repertoire of free-ranging adult golden snub-nosed monkeys (*Rhinopithecus roxellana*). *Am. J. Primatol.* 80, e22869.
- Fitch, W.T. (2010). *The Evolution of Language* (Cambridge University Press).
- Fort, P., Luppi, P.-H., and Jouvet, M. (1994). Afferents to the nucleus reticularis parvicellularis of the cat medulla oblongata: a tract-tracing study with cholera toxin B subunit. *J. Comp. Neurol.* 342, 603–618.
- Gabbiani, F., and Cox, S.J. (2010). *Mathematics for Neuroscientists* (Academic Press).
- Grossman, P., and Taylor, E.W. (2007). Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. *Biol. Psychol.* 74, 263–285.
- Gultekin, Y.B., and Hage, S.R. (2017). Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nat. Commun.* 8, 14046.
- Hage, S.R., and Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* 4, 2409.
- Hage, S.R., and Nieder, A. (2016). Dual neural network model for the evolution of speech and language. *Trends Neurosci.* 39, 813–829.
- Hauser, M.D., and Marler, P. (1993). Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behav. Ecol.* 4, 194–205.
- Hayakawa, T., Takanaga, A., Maeda, S., Seki, M., and Yajima, Y. (2001). Subnuclear distribution of afferents from the oral, pharyngeal and laryngeal regions in the nucleus tractus solitarius of the rat: a study using transganglionic transport of cholera toxin. *Neurosci. Res.* 39, 221–232.
- Herbert, H., Moga, M.M., and Saper, C.B. (1990). Connections of the parabrachial nucleus with the nucleus of the solitary tract and the medullary reticular formation in the rat. *J. Comp. Neurol.* 293, 540–580.
- Julien, C. (2006). The enigma of Mayer waves: facts and models. *Cardiovasc. Res.* 70, 12–21.
- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* 26, 235–258.
- Keay, K.A., Feil, K., Gordon, B.D., Herbert, H., and Bandler, R. (1997). Spinal afferents to functionally distinct periaqueductal gray columns in the rat: an anterograde and retrograde tracing study. *J. Comp. Neurol.* 385, 207–229.
- Liao, D.A., Zhang, Y.S., Cai, L.X., and Ghazanfar, A.A. (2018a). Internal states and extrinsic factors both determine monkey vocal production. *Proc. Natl. Acad. Sci. U S A* 115, 3978–3983.
- Liao, D.A., Zhang, Y.S., Cai, L.X., and Ghazanfar, A.A. (2018b). Internal states and extrinsic factors both determine monkey vocal production. *Proc. Natl. Acad. Sci. U S A* 115, 201722426.
- Lipski, J., and Trzebski, A. (1975). Bulbo-spinal neurons activated by baroreceptor afferents and their possible role in inhibition of preganglionic sympathetic neurons. *Pflügers Archiv Eur. J. Physiol.* 356, 181–192.
- Lynch, J.J., Thomas, S.A., Long, J.M., Malinow, K.L., Chickadonz, G., and Honori Katcher, A. (1980). Human speech and blood pressure. *J. Nervous Ment. Dis.* 168, 526–534.
- Mantyh, P.W. (1982). The ascending input to the midbrain periaqueductal gray of the primate. *J. Comp. Neurol.* 211, 50–64.
- Mantyh, P.W. (1983). Connections of midbrain periaqueductal gray in the monkey. I. Ascending efferent projections. *J. Neurophysiol.* 49, 567–581.
- Meller, S.T., and Dennis, B.J. (1991). Efferent projections of the periaqueductal gray in the rabbit. *Neuroscience* 40, 191–216.
- Mifflin, S.W. (1993). Laryngeal afferent inputs to the nucleus of the solitary tract. *Am. J. Physiol.* 265, R269–R276.
- Nikitopoulos, E., Arnhem, E., Van Hooff, J.A.R.A.M., and Sterck, E.H.M. (2004). Influence of female copulation calls on male sexual behavior in captive *Macaca fascicularis*. *Int. J. Primatol.* 25, 659–677.
- Obrist, P.A. (1981). Cardiovascular psychophysiology: a perspective. *Psychol. Med.* 12, 218.
- Patrickson, J.W., Smith, T.E., and Zhou, S.-S. (1991). Afferent projections of the superior and recurrent laryngeal nerves. *Brain Res.* 539, 169–174.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Perrodin, C., Kayser, C., Logothetis, N.K., and Petkov, C.I. (2011). Voice cells in the primate temporal lobe. *Curr. Biol.* 21, 1408–1415.
- Peters, H.F.M., and Hulstijn, W. (1984). Stuttering and anxiety: the difference between stutterers and nonstutterers in verbal apprehension and physiologic arousal during the anticipation of speech and non-speech tasks. *J. Fluency Disord.* 9, 67–84.

- Pomberger, T., Risueno-Segovia, C., Löschner, J., and Hage, S.R. (2018). Precise motor control enables rapid flexibility in vocal behavior of marmoset monkeys. *Curr. Biol.* *28*, 788–794.
- Pomberger, T., Risueno-Segovia, C., Gultekin, Y.B., Dohmen, D., and Hage, S.R. (2019). Cognitive control of complex motor behavior in marmoset monkeys. *Nat. Commun.* *10*, 3796.
- Prkachin, K.M., Williams-Avery, R.M., Zwaal, C., and Mills, D.E. (1999). Cardiovascular changes during induced emotion. *J. Psychosom. Res.* *47*, 255–267.
- Reilly, K.J., and Moore, C.A. (2003). Respiratory sinus arrhythmia during speech production. *J. Speech Lang. Hearing Res.* *46*, 164–177.
- Risueno-Segovia, C., and Hage, S.R. (2020). Theta synchronization of phonatory and articulatory systems in marmoset monkey vocal production. *Curr. Biol.* *30*, 4276–4283.e3.
- Romano, J., Kromrey, J., Coraggio, J., and Skowronek, J. (2006). Appropriate statistics for ordinal level data: should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys? In *Annual Meeting of the Florida Association of Institutional Research*, pp. 1–3.
- Sinha, R., Lovallo, W.R., and Parsons, O.A. (1992). Cardiovascular differentiation of emotions. *Psychosomatic Med.* *54*, 422–435.
- Slugg, R.M., and Light, A.R. (1994). Spinal cord and trigeminal projections to the pontine parabrachial region in the rat as demonstrated with Phaseolus vulgaris leucoagglutinin. *J. Comp. Neurol.* *339*, 49–61.
- Smith, E.E., Guyton, A.C., Manning, R.D., and White, R.J. (1976). Integrated mechanisms of cardiovascular response and control during exercise in the normal human. *Prog. Cardiovasc. Dis.* *18*, 421–443.
- Takahashi, D.Y., Fenley, A.R., Teramoto, Y., Narayanan, D.Z., Borjon, J.I., Holmes, P., and Ghazanfar, A.A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science* *349*, 734–738.
- Travers, S.P., and Norgren, R. (1995). Organization of orosensory responses in the nucleus of the solitary tract of rat. *J. Neurophysiol.* *73*, 2144–2162.
- Weber, C.M., and Smith, A. (1990). Autonomic correlates of stuttering and speech assessed in a range of experimental tasks. *J. Speech Hearing Res.* *33*, 690–706.
- Weineck, K., García-Rosales, F., and Hechavarría, J.C. (2020). Neural oscillations in the fronto-striatal network predict vocal output in bats. *PLoS Biol.* *18*, e3000658.
- Wienecke, J., Enríquez Denton, M., Stecina, K., Kirkwood, P.A., and Hultborn, H. (2015). Modulation of spontaneous locomotor and respiratory drives to hindlimb motoneurons temporally related to sympathetic drives as revealed by mayer waves. *Front. Neural Circuits.* *9*, 1.
- Yasuma, F., and Hayano, J.I. (2004). Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest* *125*, 683–690.
- Zhang, Y.S., and Ghazanfar, A.A. (2018). Vocal development through morphological computation. *PLoS Biol.* *16*, 1–19.
- Zhang, Y.S., and Ghazanfar, A.A. (2020). A hierarchy of autonomous systems for vocal production. *Trends Neurosci.* *43*, 115–126.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
<i>Macaca fascicularis</i>	European Research Biology Center, Baugy, France	N/A
Software and algorithms		
MATLAB	MathWorks	R2020b
Python	Python Software Foundation	3.8.5
Spike 2	Cambridge Electronic Design	7.19
Logistic regression, neural networks and SVM	Scikit-learn toolbox	0.23.2
Codes and analysis paradigm of the machine learning classifiers	This paper	https://datadryad.org/stash/share/CKeQq0PPIXGBUciuzQv-KpAwcF1T7axtHrKUeXv14UU

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Steffen R. Hage (steffen.hage@uni-tuebingen.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data All data needed to evaluate the conclusions are present in the paper. The raw datasets have not been deposited in a public repository because of further analyses but are available from the lead author upon request.
- Code The codes and analysis paradigm of the machine learning classifiers can be found at <https://datadryad.org/stash/share/CKeQq0PPIXGBUciuzQv-KpAwcF1T7axtHrKUeXv14UU>.
- Other items Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cynomolgus monkeys

We telemetrically recorded ECG, BP, and physical activity in six freely moving and naturally vocalizing adult cynomolgus monkeys (*Macaca fascicularis*), also known as long-tailed macaques, three females and three males, housed at the Center de Recherches Biologiques, Baugy, France (3–7 kg, 28–48 months old). Animals were kept in stainless steel cages in triplets of the same sex for physiological telemetry and vocal recordings, with visual and auditory contact between them. The dimensions of the two cages were 2.5 × 1.64 × 1.8 m each (height × width × depth; [Figure 1A](#)). The facility room was maintained at 20–24°C, 45–65% relative humidity, and with a 12 h:12 h light/dark cycle. The monkeys had *ad libitum* access to water and solid diet, supplemented with fresh fruit, vegetables, bread, and cereals. The experimental procedures were subject to ethical review (ethics committee no CEEA-111) and in agreement with the guidelines of European directive 2010/63/UE on animal welfare.

METHOD DETAILS

Implantation of telemetry system

Radio telemetry transmitters measuring ECG, BP, and activity (TL11M3D70PCTP or TL11M2D700PCT models, Data Sciences International, DSI, Saint Paul, MN, USA) were individually implanted in the

abdominal cavity of the animals after left thoracotomy. Cynomolgus monkeys were administered ketamine ($10 \text{ mg} \cdot \text{kg}^{-1}$, s.c.), buprenorphine ($0.01 \text{ mg} \cdot \text{kg}^{-1}$, s.c.), and meloxicam ($0.1 \text{ mg} \cdot \text{kg}^{-1}$, s.c.). Anesthesia was induced using propofol ($4\text{--}5 \text{ mg} \cdot \text{kg}^{-1}$, i.v.) and 4% halothane and maintained with 1.5% halothane in oxygen. For the ECG recordings, one electrode was placed between the left ventricular parietal and visceral pericardium (or epicardium) near the apex, and the second electrode on the pericardium above the right atrium to approximate a limb lead II ECG (Figure 1B). For the arterial BP recordings, the sensor catheter was introduced into the femoral artery and the catheter tip was placed in the abdominal aorta close to the origin of the common iliac arteries. For the activity recordings, the transmitters included an analog accelerometer measuring the 3D movement of the animal. Analgesic treatment with buprenorphine/meloxicam was continued for a minimum of 2 days to alleviate any postoperative pain. A minimum period of 3 weeks was allowed for recovery from the surgery. No radio telemetry implantation was specifically conducted for this study since the animals had been included in a previous project (Champéroux et al., 2018).

Data analysis of telemetry

The telemetry system could simultaneously measure and collect ECG and BP at a sampling rate of 500 Hz and physical activity at 1 Hz from freely moving animals. Data was telemetrically transmitted to a nearby receiver that converted the telemetry information to a form readily accessible by the DSI software (Figure 1A).

Recording of the animal's activity

PhysioTel Digital implants include an analogical accelerometer which measures 3D values within $\pm 8 \text{ G}$. The more the animal moves in one direction, the higher the value will be on the corresponding axis. The information is transmitted to the Ponemah software (release 6.33, Data Sciences International, St Paul, MN, USA), which calculates changes in acceleration over time, i.e., jerk.

The following equation was used to calculate jerk from the summarization of the three A/D values from the accelerometer channels and subsequent channel values:

$$\text{Jerk meter value} = c \times \sqrt{(X_{i+1} - X_i)^2 + (\gamma_{i+1} - \gamma_i)^2 + (z_{i+1} - z_i)^2}$$

where C is a constant based on the delta time for the accelerometer sampling rate.

$$C = \text{interpolation frequency} \times 3.5347$$

The constant, C , is the sampling rate (e.g., 10 Hz), multiplied by 3.5347 to correspond to activity measurements.

If the acceleration remains constant over time, the activity is 0. However, in practice, the value is never equal to 0 because of the background noise of the accelerometer, which can differ between telemetry implants. The activity values are given in counts by the Ponemah software (arbitrary units).

Vocal and video recording setup

Vocalizations were picked up by a microphone (ME64 microphone with K6 preamplifier, Sennheiser, Germany) located at an approximate distance of 10 cm from the center of the front of the cage and recorded via an analogue/digital interface (sampling rate: 44.1 kHz; Micro 1401 mkII, Cambridge Electronic Design, Cambridge, UK) with the corresponding software (Spike 2 version 7.19, Cambridge Electronic Design, Cambridge, UK). The behavior of the monkeys was constantly recorded using a 4K video camera (HC-VX989, Panasonic Corporation, Japan; sampling frequency: 30 fps). To ensure precise timings, telemetry values and vocal recordings were synchronized. Vocal onset and offset times were detected offline using software (Spike2 ver.7.19, Cambridge Electronic Design). Vocalizations were not classified into different call types. Focal observation and video recordings were used to assign calls to individual monkeys. We only considered vocalizations, which could be clearly assigned to individual monkeys, i.e., by having the callers' face clearly visible during vocal production. Calls that could not be assigned to an individual animal were disregarded. Data acquisition took place over three consecutive days. For the present study, we used recordings made on the second day of recording from 10:00 to 10:45 am for analysis, resulting in a total of 486 calls uttered by six animals (Table S1).

Machine learning classifiers

We trained three machine learning classification models in order to evaluate the predictive power of certain physiological features for the presence or absence of vocalizations. Five custom features were defined and computed from the ECG and BP in order to predict vocalizations: heart rate, respiratory rate, respiratory amplitude, blood pressure and Mayer wave power. We used the standard Scipy and Numpy libraries in Python 3.8.5 for all computations. We tested the performance of the three classifiers based on the mean value of each feature using a 2-s time window prior to call onset. The scikit-learn toolbox (Pedregosa et al., 2011) in Python was used to train three different classifiers: logistic regression, neural networks (multi-layer perceptrons with two hidden layers), and SVM. These three algorithms are the most commonly used classifiers when analyzing real world data (e.g., they are commonly used by Spotify, Evernote or Booking.com; <https://scikit-learn.org/stable/testimonials/testimonials.html>). The first is a linear classifier, whereas the other two are standard nonlinear classification approaches in machine learning. For neural networks, we tested several different multi-layer perceptron architectures and the one with the two hidden layers (of size 10 each) and ReLu (restricted linear unit) output nonlinearities proved to yield the highest predictive accuracy among the others. To determine the learning rate of neural networks we tested various methodologies and decided on the constant learning rate of 0.001. For SVM, we used RBF (radial basis function) kernels and an L2-regularizer parameter with value $C = 3.0$. The regularization parameter makes the SVM less susceptible to outliers and improves its generalization. The kernel coefficient for the RBF kernel was scaled inversely proportional to the number of features and the feature matrix covariance.

Analyses of machine learning classifiers

The vocalizations included in Table S1 were used to train the machine learning classifiers for the vocal production group. Among the control rest conditions, 54 samples were generated by shifting the time window by 1 s during the control rest period. We similarly generated the control active samples ($n = 84, 60, 60, 78,$ and 66 for monkeys 1 to 5). One animal was excluded from the predictive accuracy calculation since high-level low frequency noise in the ECG and BP signals significantly perturbed the features. When testing the predictive accuracy comparing vocal production to a combination of the control conditions, we randomly selected 11 vocalizations and 22 control data out of a total of 226 data points (112 vocalizations, 54 control rest and 60 control active samples) for monkey A and out of a total of 1011 data points (393 vocalizations, 270 control rest and 348 control active samples) for all monkeys. These test data points were reserved for the evaluation of the trained classifiers. The rest of the data points were used for the training procedure. To calculate the variability of the predictive accuracy of the machine learning algorithms we performed a 10-fold cross-validation, i.e., we repeated this procedure ten times.

Data analysis of physiological signals

Analyses of heart rate. The HR was calculated based on the identification of the R peaks, reflecting ventricular contraction, in the ECG trace. The number of R peaks were quantified with a moving average over 2 s in steps of 1 s. The peaks were marked by identifying the local maxima of the input signal with a minimum peak distance of 0.2 and peak height of 2.

Analyses of respiratory amplitude and rate. The respiratory cycles were obtained from the ECG recording by downsampling the signal from 500 to 50 Hz and applying a second-order Butterworth low-pass filter with a cutoff of 4 Hz. The respiratory rate was computed based on the number of peaks in the respiratory signal with a moving average over 2 s in steps of 1 s. The peaks were marked by identifying the local maxima of the input signal with a minimum peak distance of 0.75.

Mean arterial pressure (MAP) calculation. The MAP was computed based on the lower and upper envelopes (diastolic and systolic pressure) of the measured BP signal. The envelopes were determined using spline interpolation over the local maxima separated by at least 300 samples. One-third of the difference between the envelopes was added to the lower envelope to calculate the MAP pressure:

$$\text{MAP pressure} = \text{diastolic pressure} + (\text{systolic pressure} - \text{diastolic pressure})/3.$$

Mayer wave power calculation. The Mayer wave intensity was determined based on the spectrogram of the ECG signal by evaluating its power at 0.15 Hz at vocal onset time.

Data processing of control conditions and physiological signals. Two control conditions were defined for comparisons of the signals associated with vocal production in order to disentangle vocalization-correlated activity from general motor activity: control active and control rest. Both controls were generated during offline focal sampling of the video recordings. They were generated in a chronological order until a minimum of 10 samples were identified per monkey and per condition. Control active consisted of ten periods of at least 5 s in which a specific animal is actively moving but not vocalizing. Control rest consists of ten periods of at least 5 s in which a specific animal is neither actively moving nor vocalizing. The onsets of the two control conditions were aligned to the closest expiratory phase onset to control for the potential influence of the respiratory cycle in the vocal production signal, which was naturally aligned to the beginning of the expiratory phase. Heart rate, respiratory rate, and activity were normalized by dividing the signal by the mean value at vocal onset for each individual animal. The spectrograms of the vocal production, control active, and control rest signals were calculated using 100 s of data, 50 s prior and 50 s after vocal onset. The power spectral densities on the ECG and BP signal include peri-vocal onset data from 15 s prior and 15 s after vocal onset. The raw ECG and BP signals were bandpass filtered between 0.1 and 0.2 Hz prior to the calculation of the local phase angle at Mayer wave frequencies. The local phase angle of ECG and BP at Mayer wave frequencies and at the raw ECG and BP signals (encompassing all frequencies) were estimated over single cycles from the discrete-time analytic signal using the Hilbert transform with the MATLAB function, `ksdensity.m`, with a kernel bandwidth of 0.08. The activity of the vocal production signal was divided into tertiles, each containing a third of the population, in order to categorize vocalizations linked to high activity (over the second tertile threshold), intermediate activity (between the first and second tertiles thresholds) and low activity levels (under the first tertile threshold). For visualization purposes, a smoothing factor of 5 was applied to the HR, respiratory amplitude, respiratory rate, activity signals, and power spectral densities of the ECG and BP.

Cliff's delta metric: The effect size of group comparisons was calculated using Cliff's delta metric (Romano et al., 2006; Weineck et al., 2020). It indicates how often the values in one distribution are larger than the values in the second distribution. Cliff's delta comparison renders values ranging from -1 to 1, with identical groups having values around zero. The considered borders to define large, medium, and small effect sizes were 0.474, 0.333, and 0.147, respectively (Romano et al., 2006; Weineck et al., 2020).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using MATLAB (MathWorks, Natick, MA). The Wilcoxon rank-sum test for unmatched groups was used to reveal differences in HR (with Bonferroni correction), power spectral density of the ECG signal, and activity between vocal production and two control conditions. Signed-rank tests for matched groups were used to examine the difference in HR and RR (with Bonferroni correction), activity, and respiratory amplitude between the vocal production signal and baseline. Baseline signals were defined as the mean value for 4–5 s prior to call onset. Pearson's correlations were used to identify the relationship between BP and ECG power spectral densities between 0.133 and 3 Hz. The Rayleigh test for non-uniformity of circular data was computed for the polar histograms of the articulatory ECG and BP Mayer wave phase, raw ECG and BP signals as well as the respiratory amplitude signal at vocal onset. Significant levels for the predictive accuracy of the machine learning classifiers were estimated based on the binomial cumulative distribution function (Combrisson and Jerbi, 2015). In all performed tests, significance was tested at an $\alpha = 0.05$ level.