



# The Vocal Repertoire of Pale Spear-Nosed Bats in a Social Roosting Context

Ella Z. Lattenkamp<sup>1,2\*†</sup>, Stephanie M. Shields<sup>1†</sup>, Michael Schutte<sup>1</sup>, Jassica Richter<sup>1</sup>, Meike Linnenschmidt<sup>1</sup>, Sonja C. Vernes<sup>2,3</sup> and Lutz Wiegreb<sup>1</sup>

<sup>1</sup> AG Wiegreb, Department Biology II, Ludwig-Maximilians University Munich, Martinsried, Germany, <sup>2</sup> Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, <sup>3</sup> Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands

## OPEN ACCESS

### Edited by:

Claudia Fichtel,  
Deutsches Primatenzentrum,  
Germany

### Reviewed by:

Marco Gamba,  
University of Turin, Italy  
Genevieve Spanjer Wright,  
University of Maryland, College Park,  
United States

### \*Correspondence:

Ella Z. Lattenkamp  
ella.lattenkamp@evobio.eu

<sup>†</sup>These authors share first authorship

### Specialty section:

This article was submitted to  
Behavioral and Evolutionary Ecology,  
a section of the journal  
Frontiers in Ecology and Evolution

Received: 11 January 2019

Accepted: 21 March 2019

Published: 12 April 2019

### Citation:

Lattenkamp EZ, Shields SM,  
Schutte M, Richter J,  
Linnenschmidt M, Vernes SC and  
Wiegreb L (2019) The Vocal  
Repertoire of Pale Spear-Nosed Bats  
in a Social Roosting Context.  
*Front. Ecol. Evol.* 7:116.  
doi: 10.3389/fevo.2019.00116

Commonly known for their ability to echolocate, bats also use a wide variety of social vocalizations to communicate with one another. However, the full vocal repertoires of relatively few bat species have been studied thus far. The present study examined the vocal repertoire of the pale spear-nosed bat, *Phyllostomus discolor*, in a social roosting context. Based on visual examination of spectrograms and subsequent quantitative analysis of syllables, eight distinct syllable classes were defined, and their prevalence in different behavioral contexts was examined. Four more syllable classes were observed in low numbers and are described here as well. These results show that *P. discolor* possesses a rich vocal repertoire, which includes vocalizations comparable to previously reported repertoires of other bat species as well as vocalizations previously undescribed. Our data provide detailed information about the temporal and spectral characteristics of syllables emitted by *P. discolor*, allowing for a better understanding of the communicative system and related behaviors of this species. Furthermore, this vocal repertoire will serve as a basis for future research using *P. discolor* as a model organism for vocal communication and vocal learning and it will allow for comparative studies between bat species.

**Keywords:** vocal communication, *Phyllostomus discolor*, syllable classes, vocal repertoire, social behavior

## INTRODUCTION

Bats are highly gregarious mammals that have been extensively studied for their ability to echolocate (i.e., gain spatial information from the echoes of prior emitted ultrasonic calls). However, bats also emit social vocalizations to communicate with conspecifics and some bat species have been shown to possess rich vocal repertoires (e.g., Kanwal et al., 1994; Ma et al., 2006; Bohn et al., 2008), supporting intricate social interactions (Wilkinson, 1995, 2003). Current literature on vocal communication in bats illustrates that social vocalizations can be very complex, are highly important for bat sociality, and often vary notably between species. However, research in this field has only been scratching the surface; there is still much to learn about social communication in bats. Relative to the total number of bat species (being the second richest order of mammals with over 1,300 species), very few species have been studied, and even fewer have had their vocal repertoires described.

Research on social communication in bats generally focuses on studying a specific subset of vocalizations in a species repertoire—such as neonatal calls (Gould, 1975), calls produced during ontogeny (Knörnschild et al., 2006, 2010a), mother-infant calls (Esser and Schmidt, 1989), male song (Davidson and Wilkinson, 2004)—or more commonly on studying only one particular type of vocalization—such as distress calls (Russ et al., 2004; Hechavarria et al., 2016) or aggressive calls (Bastian and Schmidt, 2008). Fewer studies have sought to describe the repertoire of a species more comprehensively, defining several types of syllables emitted often in specific behavioral contexts (e.g., Behr, 2006; Knörnschild et al., 2010b; Wright et al., 2013). Even fewer have investigated the occurrence of syllable combination and temporal emission patterns (e.g., Kanwal et al., 1994; Bohn et al., 2008). These studies have reported a great deal of vocal diversity, ranging from 2 to 22 described vocalization types per species.

The pale spear-nosed bat, *Phyllostomus discolor*, has been in the focus of scientific attention for several years and has been investigated in a variety of psychophysical and neurophysiological studies (e.g., Firzlaff et al., 2006; Hoffmann et al., 2008; Heinrich and Wiegreb, 2013) and, more recently, neurogenetics studies (Rodenas-Cuadrado et al., 2015, 2018). *P. discolor* is a scientifically particularly interesting species as it belongs to the handful of bat species for which evidence of vocal learning (i.e., the ability to produce new or strongly modified vocalizations according to auditory experiences) has been presented (Esser, 1994; Knörnschild, 2014; Lattenkamp et al., 2018). Social vocalizations of *P. discolor* are thus especially intriguing as these bats are a valuable system for the study of vocal learning that will help deepen our understanding of this phenomenon (Lattenkamp and Verne, 2018). However, previous studies of social vocalizations in *P. discolor* have mainly focused on mother-infant communication (Esser and Schmidt, 1989; Esser, 1994; Esser and Schubert, 1998; Luo et al., 2017).

The current study is the first to assess the vocal communicative repertoire of *P. discolor* in an undisturbed social roosting context, which covers about 80% of their daily activity (La Val, 1970). Pairs and groups of three, four, and six pale spear-nosed bats were repeatedly recorded with a high resolution ultrasonic microphone array under anechoic conditions. Following the methodology of Kanwal et al. (1994), vocalizations were initially classified by two independent human raters and the classifications were subsequently statistically verified based on a fixed set of 19 automatically extracted spectral and temporal vocalization parameters. Eight distinct syllable classes were identified, and four additional, infrequently emitted classes were observed, suggesting that *P. discolor* possesses a diverse vocal repertoire. For the eight distinct syllable classes, the behavioral context at the time of emission was analyzed. The combined results present an extensive assessment of the vocal repertoire of the pale spear-nosed bat, *P. discolor*, in a social roosting context.

## MATERIALS AND METHODS

### Terminology

We follow previous literature in defining syllables as continuous vocal emissions surrounded by periods of silence (Kanwal et al.,

1994; Doupe and Kuhl, 1999; Behr and Von Helversen, 2004; Bohn et al., 2008; Gadziola et al., 2012). By this definition, syllables are the smallest, independent acoustic unit of a vocalization. A call can consist of a single or multiple syllables (Gadziola et al., 2012). For clarity, we specifically focused on studying individual syllables rather than the less objective entity of a call. Syllable classes are used to describe groups of statistically different syllables (cf. Gadziola et al., 2012; Hechavarria et al., 2016), which are assigned depending on the outcome of the classification process described below. We follow the definitions of syllable train and phrase used by Kanwal et al. (1994) (cf. “simple phrase” and “combination phrase” used by Ma et al. (2006)). The term syllable train describes a combination of two or more syllables from the same class, while a phrase describes a combination of syllables from at least two different classes. The silent period between any two syllables in a train or phrase is roughly similar and may be longer than the duration of any one syllable (Kanwal et al., 1994).

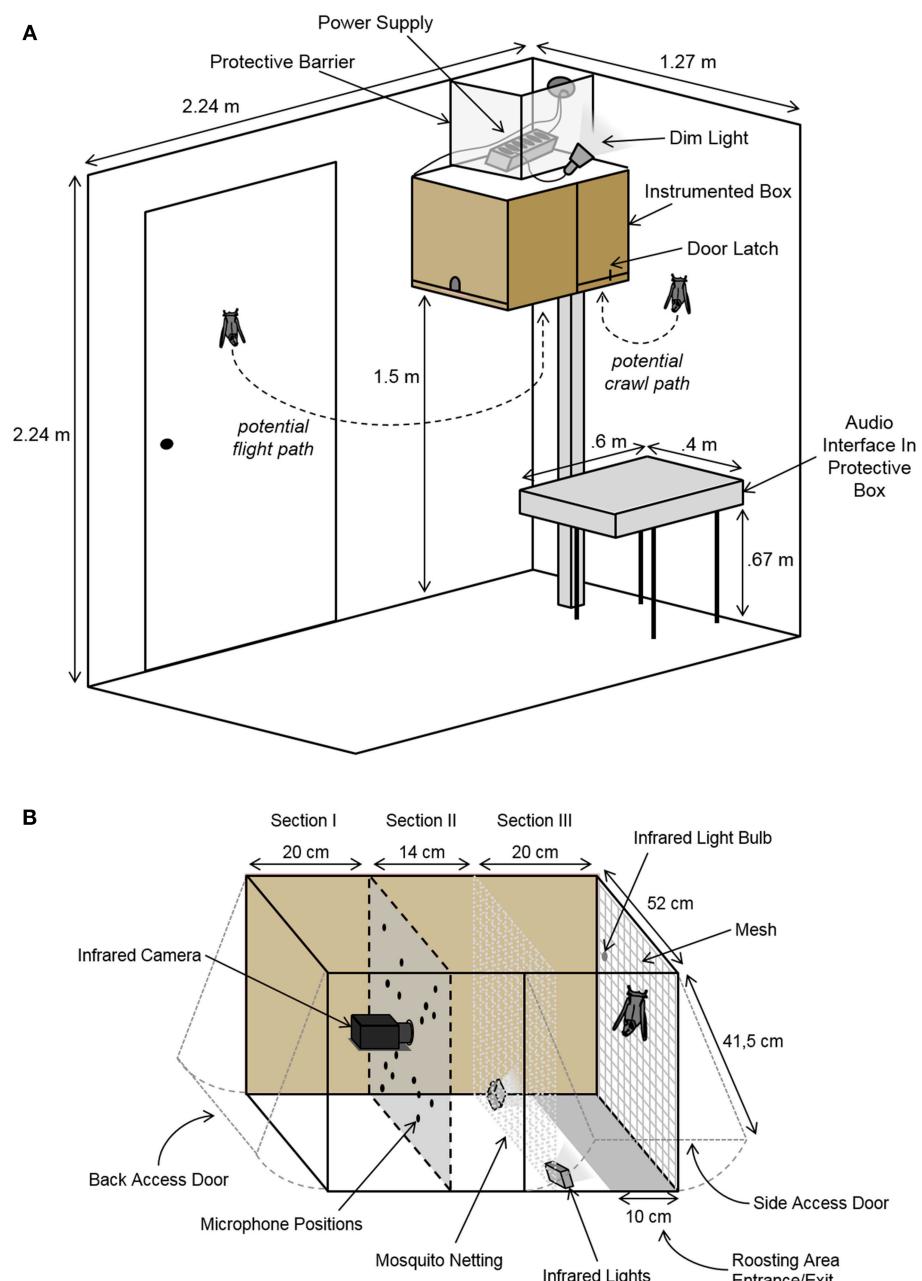
### Animals

Six adult pale spear-nosed bats, *P. discolor*, were recorded in pairs or groups of three, four, and six. Recordings were done between January and March 2018 for 5 days per week. The animals recorded in this experiment originated from a breeding colony at Ludwig Maximilian University of Munich, where they were born and housed together throughout their lives. The sex ratio between the bats was equal. One male and one female were approximately 1 year old, while the other bats were between 6 and 9 years old. The bats were provided with a species specific diet (fruits, supplements, and meal worms) and had *ad libitum* access to water during and outside of the experiment. This experiment was conducted under the principles of laboratory animal care and the regulations of the German Law on Animal Protection. The license to keep and breed *P. discolor* as well as all experimental protocols were approved by the German Regierung von Oberbayern (approval 55.2-1-54-2532-34-2015).

### Recording Setup

The recording setup was mounted in a sound-insulated chamber ( $2.24 \times 1.27 \times 2.24 \text{ m}^3$ ; L × W × H; **Figure 1A**) and consisted of a box containing recording equipment and space for the bats to roost (**Figure 1**). The instrumented box was mounted 1.5 meters above the ground, allowing the bats to fly in and out as they pleased. The ceiling light was only turned on when the experimenter was in the room. Otherwise, the chamber was only dimly illuminated by a small lamp, encouraging the bats to remain in the darker roosting area inside the box. During experimental sessions, the chamber was monitored via an infrared CCD camera (Renkforce CMOS, Conrad Electronic, Hirschau, Germany). Temperature and humidity were monitored from outside the chamber.

Vocalizations and behaviors were recorded with both high temporal and spatial resolution via a custom-built acoustic camera. This acoustic camera consisted of a 16-unit ultrasonic microphone array (custom-made on basis of SPU0410LR5H, Knowles Corporation, Itasca, IL, USA) and a high resolution infrared video camera (Point Gray Research Grasshopper3 GS3-U3-41C6NIR; FLIR Integrated Imaging Solutions, Inc.,



**FIGURE 1 |** Schematic of the setup. **(A)** The sound-insulated chamber containing the instrumented box and the audio interface. The instrumented box was mounted to the wall allowing a free flight path in and out of the box. Dim illumination encouraged the bats to remain in the darker roosting area inside the box. The chamber was monitored via an infrared camera. **(B)** Detailed schematic of the instrumented box containing the acoustic camera. The microphone array and camera faced the freely accessible roosting space. The box was divided into three sections: **(I)** an area containing the equipment and corresponding cords, **(II)** a secured space between the bats' roosting area and the panel with the recording equipment, and **(III)** a section for the bats. The bats' roosting area was illuminated with two infrared lights.

Richmond, BC, Canada) controlled and synchronized via a custom-written MATLAB (R2015a, MathWorks, Cambridge, MA, USA) script. By comparing time-of-arrival differences between all microphones of the array, the acoustic camera allows to determine the exact location of a sound source in the recorded video. The camera and microphones were mounted inside of the instrumented box ( $54 \times 52 \times 41.5 \text{ cm}^3$ ; L × W × H; **Figure 1**),

which was lined with acoustic foam. The bats could enter or exit through a 10 cm wide opening along the bottom of the backside of the box (cf. section Results, **Figure 1B**). Two additional doors with latches allowed the experimenter to access the bats and the equipment independently (**Figure 1**). The back wall of the bats' roosting space was lined with mesh for the bats to hang from and crawl on. Two small infrared lights were mounted in the

lower corners of the bats' area, illuminating the back wall. An additional infrared light bulb was hung from the mesh mounted on the back wall. This infrared light was used to synchronize the recorded video with the recorded audio. Audio data was recorded via a Horus audio interface (Merging Technologies SA, Puidoux, Switzerland) placed next to the instrumented box in the experimental chamber (**Figure 1A**).

## Recording Procedure

The six bats were observed in the recording chamber for 47 sessions (either 1.5 or 3 h long), amounting to a total of 96 h of observation. All 15 possible pair combinations between the six bats were observed for 1.5 h each. On these pair-recording days, the remaining four bats were added into the recording chamber after the first 1.5 h and all six bats were subsequently observed for another 1.5 h. In two additional sessions, first all males and then all females were observed together for 3 h each. Next, all 15 possible combinations of four bats were observed for 3 h as well.

During the recording sessions, the bats were monitored in real-time. The recording of audio and visual data was manually triggered by an experimenter from outside the chamber, when social vocalizations were emitted in the chamber. Ultrasonic vocalizations were made audible for the experimenter via real-time heterodyning of two of the 16 microphone channels and presented via headphones. The data acquisition was controlled via a custom-written MATLAB script, which saved a 10 s audio ring buffer synchronously for all 16 microphones (sampling rate: 192; microphone gain: 18 dB). The corresponding 10 s long video files were recorded synchronously via StreamPix 6 Single-Camera (NorPix, Inc., Montreal, QC, Canada) (frame rate: 100/s; shutter speed: 9.711 ms). The video files were compressed using the Norpix Motion-JPEG Encoder AVI Video Codec.

## Acoustic Analysis

For the acoustic analysis, we detected and extracted all vocalizations surrounded by silence via a custom-written MATLAB script. Syllable detection was based on amplitude peaks identified in the recordings, which were at least 20 dB louder than the background noise and were separated in time from previously detected peaks by at least 5 ms. For each identified syllable, the recording from the microphone that picked up the loudest signal was used for analysis. Nineteen acoustic parameters were extracted or calculated for each detected syllable: (1) Syllable duration and (2) maximum syllable amplitude were calculated. To represent the overall frequency content of the syllable, 5 parameters were calculated: (3) spectral centroid frequency (SCF; i.e., weighted mean of the frequencies contained in a syllable), (4) peak frequency (PF; i.e., the frequency with the most energy content), (5) minimum frequency, (6) maximum frequency, and (7) overall syllable bandwidth. The fundamental frequency ( $f_0$ ) contour of each syllable was detected using the YIN algorithm (de Cheveigné and Kawahara, 2002), and six parameters describing this  $f_0$  contour were then extracted: (8) mean  $f_0$ , (9) minimum  $f_0$ , (10) maximum  $f_0$ , and (11) starting  $f_0$  at the syllable onset. Seven additional parameters describing the  $f_0$  contour were extracted: (12, 13) the coefficients of the best-fitting linear (degree 1) polynomial and (14, 15, 16) quadratic (degree 2) polynomial to

the raw contour of the  $f_0$ . (17, 18) Furthermore, the root-mean-square errors (RMSE) between the fitted polynomials and the  $f_0$  contours were calculated (19). Lastly, the aperiodicity of syllables was also calculated via the YIN algorithm. It represents how noisy a signal is and functions as a proxy for entropic state of the vocalization (i.e., an aperiodicity of  $\geq 0.1$  indicates high entropy). The YIN algorithm first assesses the degree of aperiodicity of a recorded call and then tries to assign a fundamental frequency to those call segments where aperiodicity is low enough to do so. In the analyses of some quite complex syllables (see below), the fundamental frequency estimate may jump very quickly between quite different values.

## Syllable Classification

### Qualitative Categorization

Following Kanwal et al. (1994) and Ma et al. (2006), a preliminary classification key consisting of 20 vocalization classes was generated based on the spectrograms of a subset of recordings and previous literature (Kanwal et al., 1994; Ma et al., 2006). Subsequently, two independent raters visually assessed the spectrograms and waveforms of the extracted syllables based on their duration and frequency information, such as spectral contour, aperiodicity, or suppression of frequencies. The syllables were presented to the raters in four different ways: (1) the waveform of the syllable; (2) the spectrogram of the extracted syllable; (3) the spectrogram of the extracted syllable scaled to a fixed 100 ms window; (4) the spectrogram in a 100 ms context window, which displayed the recording 50 ms before and after the extracted syllable. This way of displaying the data allowed the raters to determine whether the syllable was extracted well or erroneously. Syllables were either sorted into syllable classes defined in the preliminary classification key, or they were marked as unsuitable for analysis due to low quality (e.g., because of spectral smear, syllable overlap, or incorrect extraction). A few vocalizations were marked as not matching any of the syllable classes present in the preliminary key. These potentially novel syllable classes were later reexamined, and two additional syllable classes were suggested as a result.

### Quantitative Categorization

For the quantitative categorization only high quality recordings of social syllables that were classified identically by both raters were used. Only classes containing at least 50 detected syllables were analyzed. The separability of the classes based on the 19 extracted spectro-temporal parameters was verified and refined based on a 5-fold cross validation procedure (Hastie et al., 2009). The dataset was stratified prior to splitting into folds to avoid empty classes and reduce variance (Forman and Scholz, 2010). In each fold, ~80% of the data for each class were employed to fit a linear discriminant analysis (LDA) classifier (Hastie et al., 2009), and this classifier was used to predict the classes of the remaining 20% of the calls. Each call was used in the test dataset exactly once. A mean confusion matrix was computed from the ground-truth labels assigned by the human raters and the labels predicted by the LDA classifier. The confusion matrix was normalized by multiplying each row vector with a constant factor to have row sums of 1. The normalized confusion matrix guided the

refinement of the preliminary labels obtained from the qualitative categorization. As the ultimate goal of the classification process was the development of an automatic classifier, which renders human raters redundant in the future, an algorithmically greedy procedure was used to merge the pair of classes with the highest off-diagonal normalized confusion score. This procedure was done with the input of the human raters, confirming the reasonableness of the merge. The LDA analysis was then rerun on the altered dataset and this algorithm was iterated as long as the human raters agreed that the two candidate classes for merging were non-trivial to separate by their spectrograms. The merging was continued, until a 60% overlap of the human raters and LDA classification was reached.

## Behavioral Video Analysis

We assessed the behavioral context observed during the emission of syllables belonging to the previously established classes. For that reason, an ethogram containing 56 detailed behaviors for *P. discolor* was generated based on personal observations (ML, SS, EL). More specifically, the ethogram encompassed 20 behaviors observed in neutral contexts, 18 in prosocial, and 18 in antagonistic behavioral contexts. This ethogram was used by a naïve rater to score the behaviors observed in the video files. The rater was blinded to the emitted syllables contained in the videos. The behavioral scoring was done in the Behavioral Observation Research Interactive Software (BORIS) (Friard and Gamba, 2016), and the behavior that occurred at the time of syllable emission was extracted.

## RESULTS

Within the 96 h of observation 1,434 recordings were made. The automatic syllable finder identified 57,955 vocalizations in these recordings, which were assessed by the two independent raters. The majority of these vocalizations were excluded from the subsequent quantitative analyses for several reasons: 56% ( $n = 32,551$ ) were excluded, because one or both raters marked them as unsuitable for the classification (due to syllable overlap or low recording quality occurring when vocalizations were emitted outside the instrumented box) or because the two independent raters disagreed on their classification; 2% ( $n = 1,115$ ) of the recorded sounds were excluded as they presented no vocalizations, but rather scratching noises produced by the bats brachiating on the back wall of the box; and 10% ( $n = 5,630$ ) of the data were eventually excluded, because not all 19 spectro-temporal syllable parameters could fully be extracted. The remaining 32% ( $n = 18,658$ ) of the vocalizations represented conservatively selected, high quality syllables classified identically by both independent raters. These syllables were qualitatively and quantitatively assessed as belonging to 13 syllable classes. Of these 13 classes eight were represented by more than 50 syllables and thus evaluated as commonly occurring in this social roosting context ( $n = 6,162$ ) and four classes were represented by <50 syllables and are thus reported as rarely occurring ( $n = 81$ ). The largest class ( $n = 12,416$ ) was comprised of calls with a suppressed fundamental frequency (SF class) and is reported separately below.

For the 19 extracted spectro-temporal parameters, the 25th, 50th, and 75th percentiles (i.e., first, second, and third quartiles) are reported below to represent data distribution. These values are presented as follows: Q50 [Q25 Q75]. Additionally, all quartiles for each parameter are listed in **Supplementary Table S1** for each common syllable class and in **Supplementary Table S2** for each rare syllable class and the suppressed fundamental frequency class. An example of all commonly occurring syllables is given in **Figure 2**, while the variation within these classes is illustrated in the Supplementary Material (**Supplementary Figure S1**).

## Common Syllable Classes

### High Entropy (HE) Vocalizations

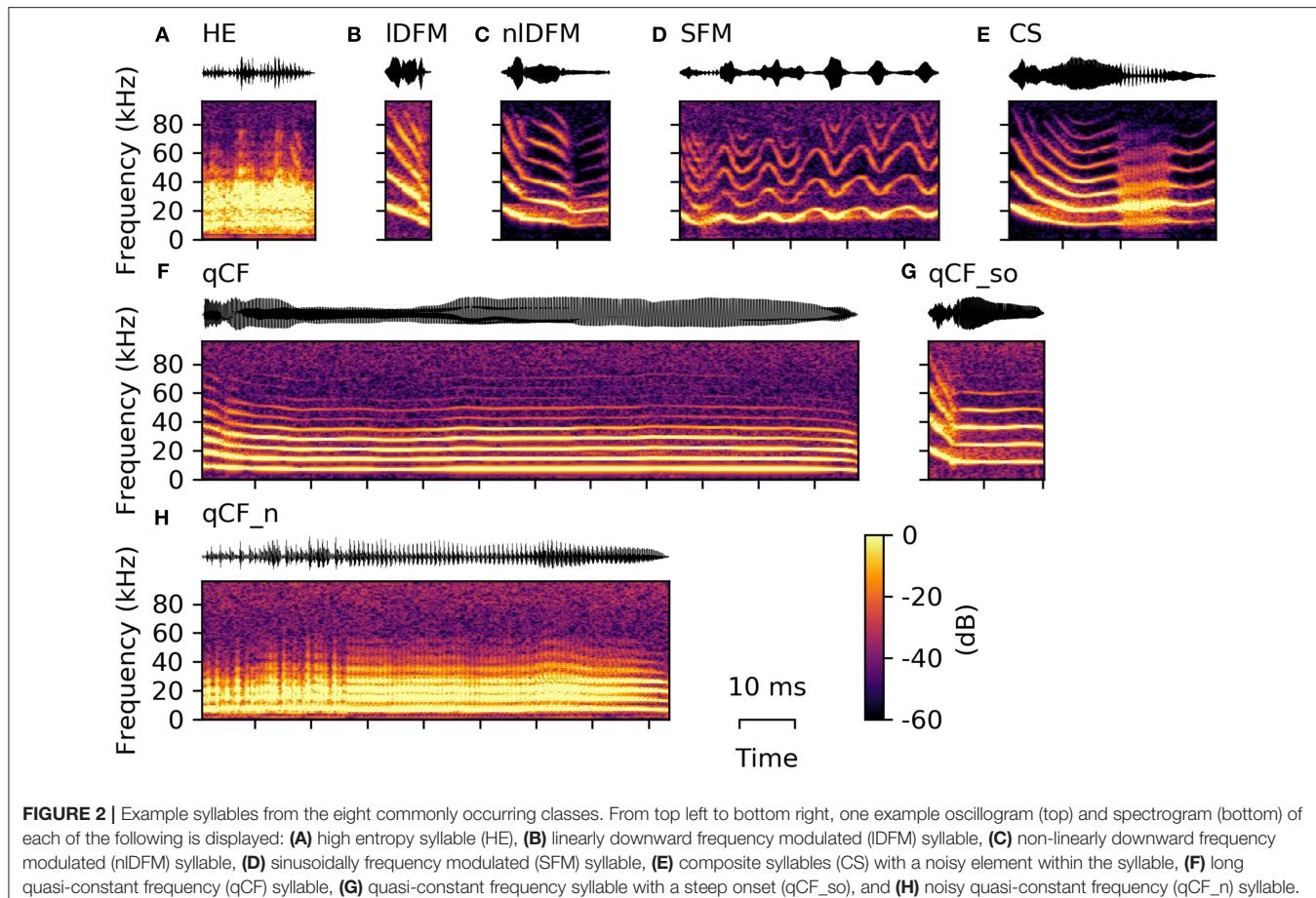
The majority of high quality, commonly emitted social syllables belong to the high entropy (HE) class ( $n = 3,860$ ; 63% of all syllables in the commonly occurring classes). HE syllables were termed according to their appearance in the spectrogram (i.e., smeared along the frequency axis), and can generally be described as noisy or screechy vocalizations (**Figure 2A**). They can still retain some degree of harmonicity, similar to synthesized tonal noises (iterated rippled noises) (Yost, 1996), and if the residual tonality was strong enough, modulations of the fundamental frequency (typically sinusoidal) could be observed (**Supplementary Figure S1**). As expected, HE syllables displayed a very high degree of aperiodicity (0.42 [0.34 0.48]; cf. Q50 [Q25 Q75], **Figure 3**). The short average duration of HE syllables (6.24 [4.74 10.27] ms) can be explained by our definition of syllable: The raters observed that long HE calls are often composed of several HE syllables (cf. **Figure 9A**), which were analyzed individually, if the call was strongly amplitude modulated and the modulation period longer than 5 ms (cf. 5 ms criterion for syllable separation).

### Linearly Downward Frequency Modulated (IDFM) Vocalizations

Seven hundred and twenty-seven syllables (12%) are composed of linear downward frequency modulations (IDFM) of the fundamental frequency (**Figure 2B**). Linearly DFM syllables are usually relatively short (6.74 [5.35 8.78] ms). They have a steep downward slope ( $-1.70 [-2.06 -1.41]$  kHz/ms) and the highest mean fundamental frequency (17.27 [15.83 18.65] kHz; **Figure 3**) of all commonly occurring syllables.

### Non-linearly Downward Frequency Modulated (nIDFM) Vocalizations

Non-linearly downward frequency modulated (nIDFM) syllables ( $n = 562$ ; 9%) also sweep downward, but they have a curved shape, or an irregular offset including small constant frequency or upward frequency modulated components (**Figure 2C**). These nIDFM syllables are generally longer than IDFM syllables (17.10 [13.72 20.13] ms; **Figure 3**) and have a lower mean  $f_0$  (14.72 [13.56 15.71] kHz; **Figure 3**). While IDFM and nIDFM syllables have a comparable bandwidth (IDFM: 28.50 [23.25 33.75] kHz; nIDFM: 28.50 [23.25 33.00] kHz), the slope of nIDFM syllables is less steep on average ( $-0.74 [-1.07 -0.53]$  kHz/ms).



**FIGURE 2 |** Example syllables from the eight commonly occurring classes. From top left to bottom right, one example oscilogram (top) and spectrogram (bottom) of each of the following is displayed: **(A)** high entropy syllable (HE), **(B)** linearly downward frequency modulated (IDFM) syllable, **(C)** non-linearly downward frequency modulated (nIDFM) syllable, **(D)** sinusoidally frequency modulated (SFM) syllable, **(E)** composite syllables (CS) with a noisy element within the syllable, **(F)** long quasi-constant frequency (qCF) syllable, **(G)** quasi-constant frequency syllable with a steep onset (qCF\_so), and **(H)** noisy quasi-constant frequency (qCF\_n) syllable.

### Sinusoidally Frequency Modulated (SFM) Vocalizations

Also frequently occurring were syllables with a sinusoidal  $f_0$  contour (SFM) ( $n = 445$ ; 7%). SFM syllables have a stable sinusoidal frequency modulation with small overall variation in modulation depth and modulation frequency, and they generally do not have an onset that notably exceeds the first frequency modulation (Figure 2D). However, SFM syllables can also have a steep linear downward sweep onset and a horizontal, ascending, or descending SFM tail (cf. Supplementary Figure S1). Irregular SFM syllables are also emitted and consist of inconsistent sinusoidal frequency modulations. SFM syllables can vary in both the rate and depth of oscillations. Similar to HE syllables, SFM vocalizations are often strongly amplitude modulated and our definition of syllables thus determines the rather short average durations of the SFM syllables (5.51 [4.66 7.90] ms; Figure 3).

### Composite (CS) Vocalizations

Composite syllables (CS;  $n = 286$ ; 5%) contain both tonal and noisy elements. Frequently, the syllable begins with a tonal, downward frequency-modulated sweep and then ends with a HE element. One or more HE elements can also occur within syllables (Figure 2E). In most cases, a CS is a SFM syllable that is interrupted by one or more HE elements. These syllables had

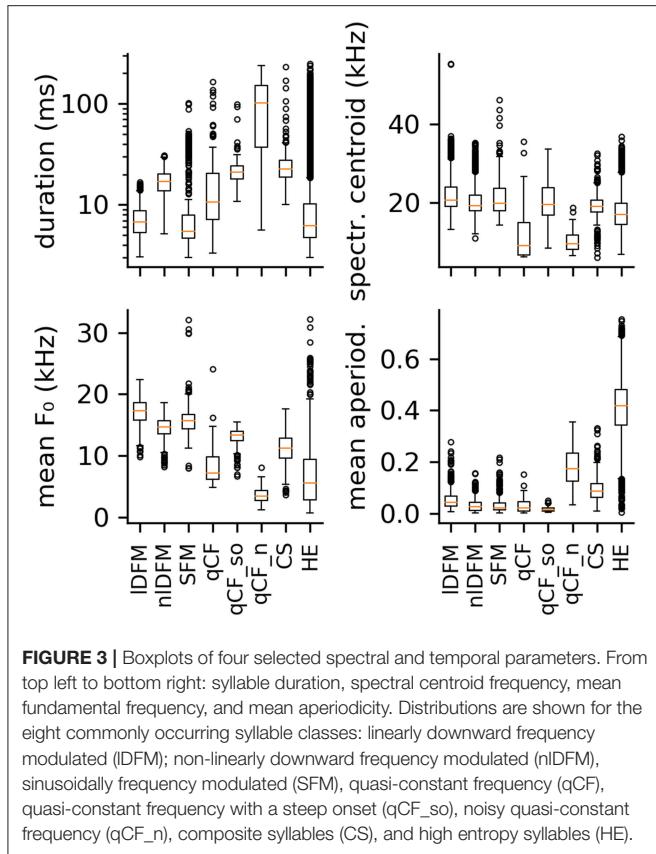
the third highest average aperiodicity (0.09 [0.06 0.12]; Figure 3) of the commonly emitted syllables.

### Quasi-Constant Frequency (qCF) Vocalizations

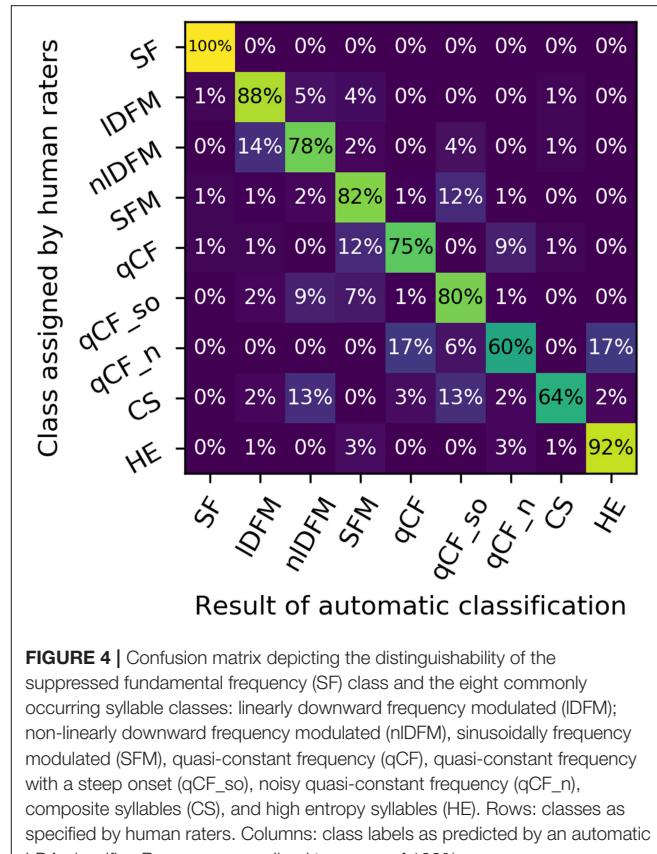
Quasi-constant frequency (qCF) syllables ( $n = 67$ ; 1%) have a near constant fundamental frequency for the duration of the entire syllable (Figure 2F). qCF syllables are tonal and have no specific onset, but rather start immediately with the constant frequency element. Overall, syllables in the qCF class tended to have low mean  $f_0$ s (7.19 [6.15 9.87] kHz; Figure 3).

### Quasi-Constant Frequency Vocalizations With a Steep Onset (qCF\_so)

Tonal qCF syllables can also have a steep downward frequency modulated onset (qCF\_so;  $n = 89$ ; 1%; Figure 2G). A separate class was created for those qCF\_so syllables as they necessarily differ in many parameters from pure qCF syllables, which lack such a clear onset. For example, qCF\_so syllables have stronger negative  $f_0$  slopes than the qCF syllables, because of the added onset (qCF\_so:  $-0.40 [-0.51 -0.28]$  kHz/ms; qCF:  $-0.05 [-0.20 0.01]$  kHz/ms). For the same reason, the qCF\_so syllables are generally longer (qCF\_so: 21.03 [17.98 24.33] ms; qCF: 10.64 [7.20 20.52] ms).



**FIGURE 3 |** Boxplots of four selected spectral and temporal parameters. From top left to bottom right: syllable duration, spectral centroid frequency, mean fundamental frequency, and mean aperiodicity. Distributions are shown for the eight commonly occurring syllable classes: linearly downward frequency modulated (IDFM); non-linearly downward frequency modulated (nlIDFM), sinusoidally frequency modulated (SFM), quasi-constant frequency (qCF), quasi-constant frequency with a steep onset (qCF\_so), noisy quasi-constant frequency (qCF\_n), composite syllables (CS), and high entropy syllables (HE).



**FIGURE 4 |** Confusion matrix depicting the distinguishability of the suppressed fundamental frequency (SF) class and the eight commonly occurring syllable classes: linearly downward frequency modulated (IDFM); non-linearly downward frequency modulated (nlIDFM), sinusoidally frequency modulated (SFM), quasi-constant frequency (qCF), quasi-constant frequency with a steep onset (qCF\_so), noisy quasi-constant frequency (qCF\_n), composite syllables (CS), and high entropy syllables (HE). Rows: classes as specified by human raters. Columns: class labels as predicted by an automatic LDA classifier. Rows are normalized to a sum of 100%.

## Noisy Quasi-Constant Frequency (qCF\_n) Vocalizations

Noisy quasi-constant syllables (qCF\_n) are essentially high entropy versions of the tonal qCF syllables ( $n = 126$ ; 2%; **Figure 2H**). They also did not start with a frequency modulated onset. Of all syllable classes, qCF\_n syllables had the longest average durations (102.20 [37.13 151.90] ms), lowest mean  $f_0$ s (3.45 [2.71 4.38] kHz), and lowest spectral centroids (9.51 [8.07 11.78] kHz). They had the second highest average aperiodicity (0.17 [0.12 0.23]; **Figure 3**).

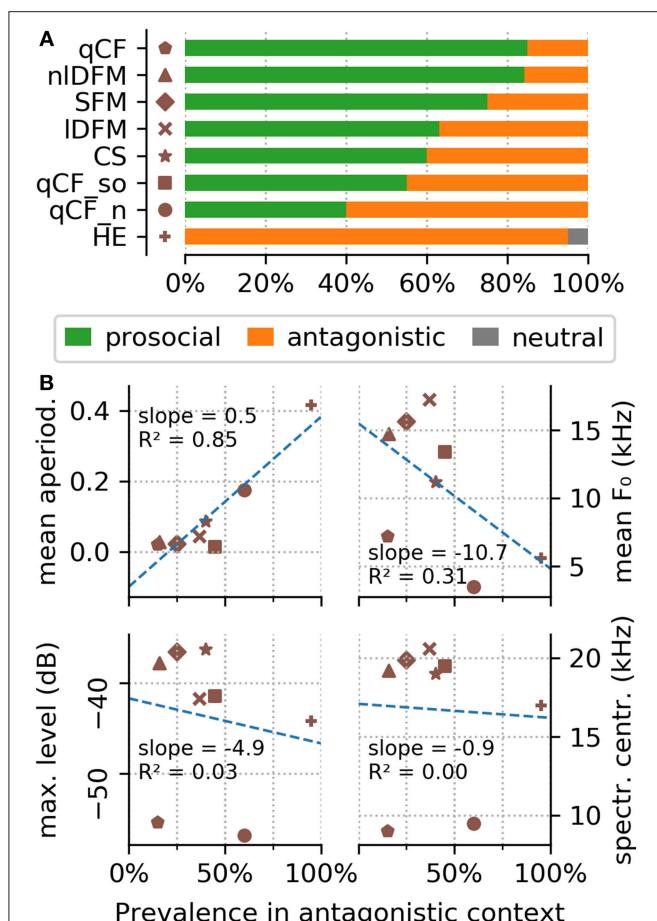
In the quantitative analysis, the LDA classifier performed with an overall accuracy of 87% over the eight classes described above (chance level: 12.5%) (**Figure 4**). The mean overall precision score was 89%, mean overall recall 87%, mean per-class precision 67%, and mean per-class recall 76%. **Figure 4** reproduces the row-normalized confusion matrix, i.e., each cell shows which percentage of calls of a specific human-rated class is assigned to a specific class label by the automatic classifier. The confusion matrix shows that particularly high recall scores are attained for IDFM and HE calls, which also separate comparatively well univariately (based on mean  $f_0$  and mean aperiodicity, respectively).

## Behavioral Context of the Common Syllable Classes

For each of the eight commonly occurring syllable classes, 20 videos were scored for the behaviors displayed by the bats

during syllable emission. For the IDFM and nlIDFM classes only 19 instances could successfully be scored as the behavior for one instance was performed outside the field of view of the camera. From the ethogram of 56 detailed behaviors, only 23 behaviors were observed during syllable emission (**Supplementary Table S3**). Only one single observation was ever made, where a vocalization was emitted in a neutral behavioral context (**Supplementary Table S3; Figure 5A**). More specifically, a single HE syllable was emitted in a context scored as “brachiating on walls or ceiling.” Other than that, syllables were always emitted either in a prosocial or an antagonistic behavioral context.

The behavioral analyses show that the HE syllables are emitted 95% of the time in antagonistic encounters (**Supplementary Table S3**). One exception is the above mentioned single observation of a HE syllable emitted in a neutral context. All other syllables were, with varying prevalence, emitted in both, prosocial and antagonistic contexts (**Supplementary Table S3; Figure 5A**). Syllables from the qCF, SFM, and nlIDFM classes were emitted in prosocial behavioral contexts in 75–85% of the scored videos (**Supplementary Table S3; Figure 5A**). CS, IDFM, and qCF\_so syllables were emitted slightly more often in prosocial than antagonistic contexts (in 55–63% of the videos, **Supplementary Table S3**). Noisy qCF syllables (qCF\_n) were emitted in antagonistic behavioral contexts in 40% of



the scored videos. Stable correlations were found between some acoustic parameters and the behavioral context in which a syllable was emitted: Specifically, the measured aperiodicity of the syllables is strongly positively correlated with their prevalence in antagonistic encounters (Figure 5B). Also syllable  $f_0$ s are lower during antagonistic behaviors (Figure 5B).

### Rare Syllable Classes

In addition to the commonly occurring syllable classes, several vocalizations were repeatedly, but extremely infrequently emitted. Specifically, out of the total of 18,658 high quality recordings fewer than 50 vocalizations per rare syllable class were recorded. Thus, not enough data are available to include these vocalizations in the statistical analysis. They are described in the following as purely observational and should be considered as rarely emitted, at least in a social roosting context.

### Puffs

During the recording sessions, the bats repeatedly emitted air puffs ( $n = 42$ ), which appeared to result from bats forcefully expelling air through their nostrils. These sounds are not necessarily to be considered sneezing, but are rather short nasal exhalation potentially used to clean the nostrils. The spectrograms of puffs appear to be noisy sound clouds with a sharp onset (Figure 6A). As the puffs did not contain a tonal component, the mean aperiodicity and bandwidth of these puffs were the highest of all recorded vocalizations (aperiodicity: 0.43 [0.40 0.47] and bandwidth: 45.75 [42.00 48.75] kHz).

### V-Shaped Vocalizations

Syllables from this class ( $n = 30$ ) consisted of a downward frequency modulated onset and a subsequent upward sweep, resulting in a characteristic “V”-shaped frequency contour (Figure 6B). Vocalizations in the V-shaped class are in shape comparable to the sinusoidal vocalizations, but always end within the first modulation.

### Noisy Quasi-Constant Frequency Vocalizations With Steep Onset (qCF\_nso)

The qCF\_nso syllables were recorded only five times and were a combination of the qCF\_n and the qCF\_so syllable classes (Figure 6C). They also consist of a steep downward frequency modulated onset followed by a quasi-constant syllable element. However, they were emitted with higher sound pressure levels than qCF\_n and higher aperiodicity than qCF\_so syllables (Supplementary Tables S1, S2), resulting in a noisy version of the qCF\_so syllable type.

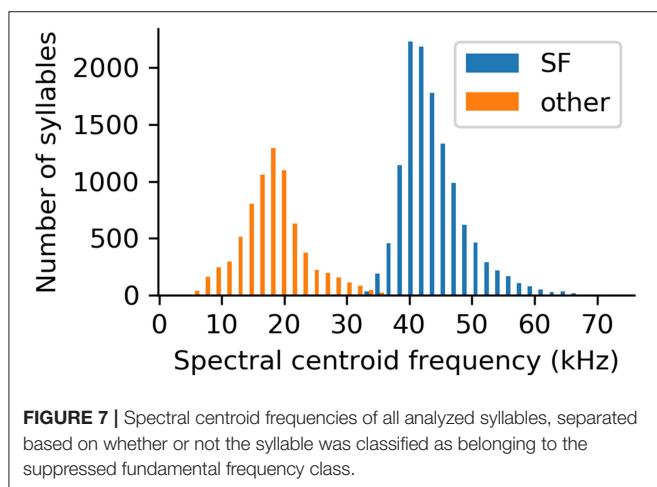
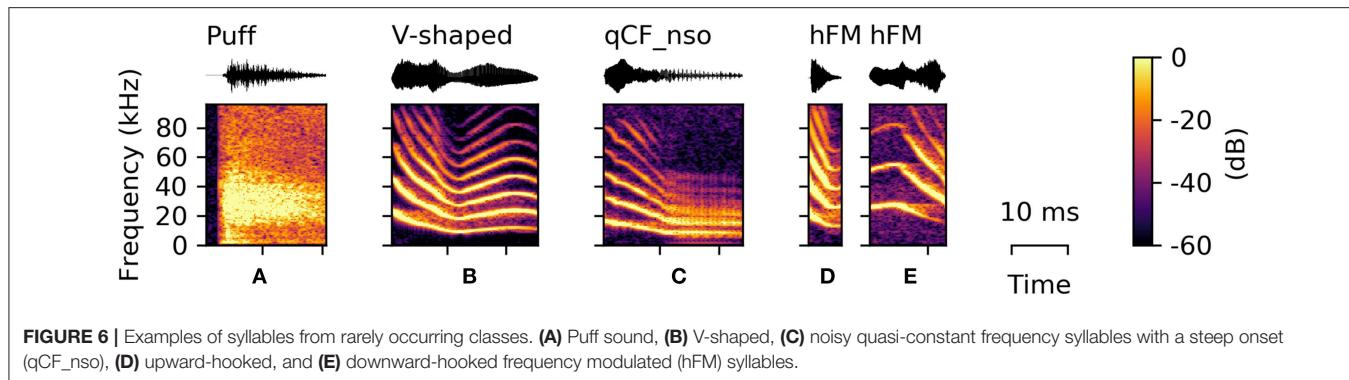
### Hooked Frequency Modulated (hFM) Vocalizations

Upward- or downward-hooked frequency modulated (hFM) syllables ( $n = 4$ ) are characterized by the similarity between the shape of the vocalization displayed in the spectrogram and a hook. These syllables are typically short and can appear in either an upward-hooked (Figure 6D) or a downward-hooked (Figure 6E) form. These two hFM syllable types were the least abundant (upward-hooked:  $n = 1$ ; downward-hooked:  $n = 3$ ). HFM syllables had the highest average spectral centroid aside from syllables with a suppressed fundamental frequency (27.08 [21.71 33.09] kHz). However, comparative results should be taken with care, as the quantitative characteristics of this class are not well-supported due to the small number of syllables detected.

### Suppressed Fundamental (SF) Class

The vast majority of recorded syllables belonged to the suppressed fundamental (SF) class ( $n = 12,416$ ; 66% of the high quality, uniformly rated syllables). This syllable class can easily be distinguished from all other recorded syllables by its high spectral centroid (Figure 7). In fact, the spectral centroid frequency is a parameter showing a clear bimodal distribution of the data, splitting SF syllables and syllables of all other classes (Figure 7).

Syllables in the SF class have either a fully or partially suppressed fundamental frequency, and the dominant harmonic is instead the second or even third harmonic (Figure 8). SF syllables typically had short durations (4.07 [3.46 5.04] ms,



**Supplementary Table S2**) and high spectral centroids (43.05 [40.65 46.51] kHz, **Supplementary Table S2**). Especially the very short durations indicate that this syllable class includes the species-specific echolocation calls, which typically range in duration between 0.3 and 2.5 ms (Rother and Schmidt, 1985; Kwiecinski, 2006; Luo et al., 2015). However, the SF class also included syllables, which structurally resembled syllables from other commonly occurring syllables classes with the only decisive difference that the fundamental frequency was fully or partially suppressed (**Figure 8**). Based on these strong characteristics and the varying shape of the SF syllables, this class can be easily separated from the other classes, but should rather be regarded as a meta-class, containing versions with suppressed fundamental frequency of most other syllable types. The function of these SF calls is currently uncertain and might or might not vary from the normal context of the syllable type with expressed fundamental frequency.

## Syllable Combinations: Trains and Phrases

Very few studies have investigated temporal emission patterns of syllables and the existence of consistently-occurring syllable combinations (e.g., Kanwal et al., 1994; Bohn et al., 2008; Knörnschild et al., 2014; Smotherman et al., 2016). Previous literature shows, however, that for certain bat species the

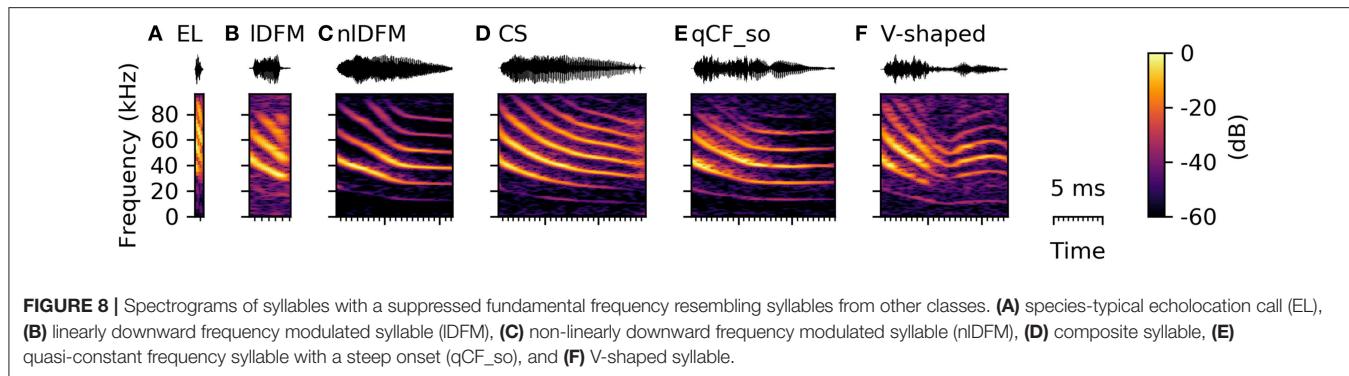
temporal emission pattern of social vocalizations can be highly complex. *Phyllostomus discolor* also emits combinations of syllables in a standardized order and with constant temporal emission patterns. Temporal relationships between syllables were not analyzed in the current work, thus we cannot draw qualitative conclusions about this aspect of the vocalizations. However, during syllable classification we observed several syllable combinations of varying length, complexity, and number of contained syllables (**Figure 9**).

Observed syllable trains consist of multiple syllables from the same class repeated with roughly the same temporal distance, whereby the silent interval can be longer than the preceding syllable (**Figures 9B,C**). Syllable trains can be of varying overall length, depending on the number of contained syllables. Phrases consist of syllables from two or more classes (**Figures 9D–F**), which can be repeated several times (usually in a fixed temporal distance). We found eight different types of syllable combination, which were repeatedly recorded over the duration of the experiment. The behavioral purpose of syllable trains and phrases is thus far purely speculative. A repetitive emission of phrases might serve to emphasize the transmitted information, but the number of phrase repetitions could also carry information by itself. Though the function and magnitude of syllable trains and phrases in these bats is currently unknown, we want to report our observation of them to encourage further research in this direction.

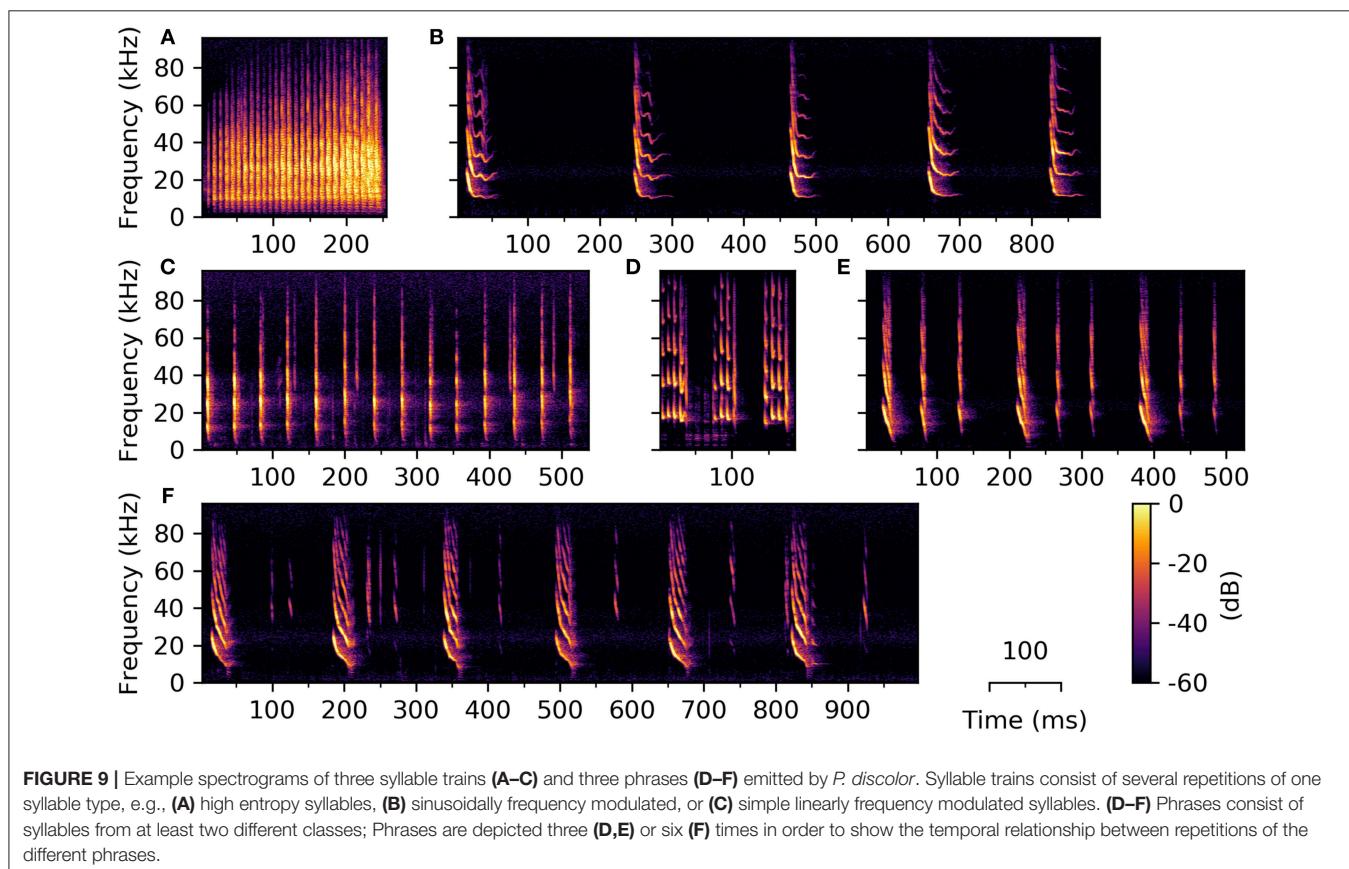
## DISCUSSION

### Vocalizations of *P. discolor*: Known and Novel

Here we present an extensive assessment of the vocal repertoire of the pale spear-nosed bat, *P. discolor*. As we recorded vocalizations in a social roosting context, which is the main pastime of *P. discolor* (Kwiecinski, 2006), we are confident that we identified the majority of social vocalizations emitted by this species. From 18,658 high-quality syllable recordings, we were able to define eight distinct classes, uniquely different from each other in their spectro-temporal parameters. We were also able to support the acoustic analysis with a detailed assessment of the behavioral contexts in which these eight syllable classes are generally emitted (**Supplementary Table S3; Figure 5**). Furthermore, we describe



**FIGURE 8 |** Spectrograms of syllables with a suppressed fundamental frequency resembling syllables from other classes. **(A)** species-typical echolocation call (EL), **(B)** linearly downward frequency modulated syllable (IDFM), **(C)** non-linearly downward frequency modulated syllable (nIDFM), **(D)** composite syllable, **(E)** quasi-constant frequency syllable with a steep onset (qCF\_so), and **(F)** V-shaped syllable.



**FIGURE 9 |** Example spectrograms of three syllable trains **(A–C)** and three phrases **(D–F)** emitted by *P. discolor*. Syllable trains consist of several repetitions of one syllable type, e.g., **(A)** high entropy syllables, **(B)** sinusoidally frequency modulated, or **(C)** simple linearly frequency modulated syllables. **(D–F)** Phrases consist of syllables from at least two different classes; Phrases are depicted three **(D,E)** or six **(F)** times in order to show the temporal relationship between repetitions of the different phrases.

four additional call classes, which were only infrequently emitted by the bats and are thus described here, but not analyzed on the basis of their spectro-temporal characteristics.

Most syllable classes described in the present study have never before been observed for this species. Especially the quasi-constant frequency modulated (qCF) class and classes containing qCF elements (i.e., qCF\_so and qCF\_n) have hitherto not been reported for *P. discolor*. From our behavioral observations (**Supplementary Table S3**) it becomes apparent that all three classes containing syllables with a qCF element are used in very versatile behavioral contexts. This could indicate a loose behavioral association with the syllable structure and

one could speculate about a behaviorally more meaningful variation of these syllables in their specific context (e.g., duration of qCF element could indicate special emphasis on a particular meaning). However, such speculations await experimental confirmation.

Sinusoidally frequency modulated (SFM) syllables have received considerable attention in previous literature. In *P. discolor*, SFM syllables were found to be used in mother-infant communication (as e.g., maternal directive calls and late forms of infant isolation calls) and can encode individual signatures, and even vocal dialects (Gould, 1975; Esser and Schmidt, 1989; Esser and Lud, 1997; Esser and Schubert, 1998). We can

confirm that the majority of the analyzed SFM syllables were emitted in the behavioral contexts “attention seeking” or “vocal contact,” which are both in line with previous observations (**Supplementary Table S3**). In addition to the usage of SFM syllables in these contexts, we also demonstrated their emission in antagonistic encounters (**Supplementary Table S3; Figure 5A**). Emission of one syllable type in a variety of different behavioral contexts (cf. **Supplementary Table S3; Figure 5A**) suggests complex communicative function or purpose. Thus, our results support previous findings, which advocate syllable subgroups, in which vocalizations with very similar acoustic parameters can be further split up based on associated behaviors (Bohn et al., 2008; Kanwal, 2009). As described above, the syllable classification here presented is based purely on spectrogram shape and the extracted syllable parameters. This allows us to present mathematically distinct syllable classes and validates our first, subjective classification scheme. Nevertheless, the established classes may be further differentiated according to their behavioral contexts. Our behavioral assessments show that syllables from a single class with very similar acoustic characteristics can be used in up to 10 different behavioral contexts (**Supplementary Table S3**). The establishment of syllable subgroups (i.e., splitting of the presented syllable classes) based on their contextual usage would require extensive, detailed behavioral observations and ideally confirmation via playback experiments. We also want to highlight the possibility that additional syllable classes might be contained in the *P. discolor* repertoire, which were not emitted in the here reported social roosting context.

## Comparison to the Closely Related Species (*P. hastatus*): Emerging Vocal Complexity

The number of distinct syllable classes assessed in this study (eight) is comparable to vocal repertoire descriptions of other bat species, which also found between 2 and 10 syllable types (e.g., Nelson, 1964; Gould, 1975; Barclay et al., 1979; Kanwal et al., 1994; Pfalzer and Kusch, 2003; Bohn et al., 2004; Wright et al., 2013; Knörnschild et al., 2014). When comparing the vocal repertoire of *P. discolor* to a closely related species (*P. hastatus*, which lives under essentially identical social and ecological conditions), it is noticeable, that the vocal repertoire of *P. hastatus* is less expansive. In addition to their echolocation calls, only two types of social calls are reported for *P. hastatus*, namely group-specific foraging calls, so-called screech calls, and infant isolation calls (Bohn et al., 2004). The screech calls of *P. hastatus* were shown to be used for the recognition of social group members during foraging, while infant isolation calls help mothers to recognize offspring (Boughman, 1997; Boughman and Wilkinson, 1998; Wilkinson and Boughman, 1998). Vocalizations reported as infant isolation calls are distinctly different between *P. discolor* and *P. hastatus*, with the former using single, clearly sinusoidally frequency modulated calls (Esser and Schmidt, 1989) and the latter typically using a pair of linear or bent frequency modulated calls (Bohn et al., 2007). The broadband, noisy screech calls of *P. hastatus* are similar in their spectral characteristics to the here defined high entropy (HE) syllables (Boughman, 1997), they are,

however, used for the coordination of foraging activities and are not emitted in antagonistic behaviors contexts as observed in this study (**Supplementary Table S3**). The surprising difference in the size of the vocal repertoires of these closely related species, which are so similar in their ecology and lifestyle, only highlights the value of *P. discolor* as a model species for vocal communication and vocal learning. The vocal repertoires of the other members of the genus (*P. elongatus* and *P. latifolius*) are still unknown. Uncovering the evolutionary background of the emergence of such differences in vocal complexity in closely related species might help us to shed light on the evolution of communicative systems and the capacity for vocal learning in bats.

## Similarities to Distantly Related Species: Acoustic Universals

A number of distantly related bat species were reported to emit high entropy calls during aggressive encounters (e.g., Russ et al., 2004; Hechavarria et al., 2016; Prat et al., 2016). It has been hypothesized that aggressive vocalizations tend to always be long, rough, and lower in frequency (Briefer, 2012). We confirmed a strongly positive correlation between the mean syllable class aperiodicity and its prevalence in antagonistic confrontations (**Figure 5B**). We also detected a negative correlation between the mean fundamental frequency of a syllable class and its occurrence during aggressive encounters. Overall, these findings support the idea of shared characteristics of mammalian vocalizations in strongly emotional behavioral contexts and provide further evidence for acoustic universals and potential for interspecies communication (Filippi, 2016; Filippi et al., 2017).

## Temporal Emission Patterns: Evidence for Higher Order Vocal Constructs

Previous studies suggest that syllable sequences such as trains or phrases can encode combinational meaning or emphasis, thus increasing the available vocal complexity for a given bat species (e.g., Behr and Von Helversen, 2004; Bohn et al., 2008; Smotherman et al., 2016; Knörnschild et al., 2017). Sequences of syllables, which present higher order vocal constructs, have been described for a few bat species (for review see Smotherman et al., 2016). However, for the family Phyllostomidae, which is a very ecologically diverse and speciose bat family [i.e., >140 described species within 56 genera (Wetterer et al., 2000)], to date there have been only two published observations of the use of such hetero-syllabic constructs. Specifically, only for Seba’s short-tailed bat (*Carollia perspicillata*) and the buffy flower bat (*Erophylla sezekorni*) descriptions of syllable combinations (i.e., simple trains and phrases) are available (Murray and Fleming, 2008; Knörnschild et al., 2014). Here we provide further evidence for syntax usage in a phyllostomid bat, which opens this family up for future in-depth research on this topic.

## CONCLUSIONS

In the framework of this study, 18,658 high-quality social vocalizations of the pale spear-nosed bat, *P. discolor*, were

recorded under laboratory conditions. From 6,162 of these, it was possible to define eight robust syllable classes, including some vocalizations not previously known to be produced by these bats. Furthermore, we were also able to assess the behavioral contexts in which these syllable classes are generally emitted, and could show that e.g., high entropy syllables are exclusively emitted in aggressive encounters. We also describe four additional, rarely occurring syllable classes (i.e., 81 recordings in total). The majority of recorded syllables ( $n = 12,416$ ) present evidence for a meta-class of vocalizations, i.e., syllables from different classes with the joint characteristic of having a suppressed fundamental frequency. Finally, we present tentative evidence for emission of syllable trains and phrases in this Neo-tropical bat species, highlighting the described complexity of *P. discolor* vocalizations. Together, these results present an extensive assessment of the vocal repertoire of *P. discolor* in a social roosting context and the associated behavioral contexts.

## AUTHOR CONTRIBUTIONS

LW, ML, SV, and EL conceived and supervised the study. SS recorded the data. SS, EL, and ML developed the classification key. LW wrote the syllable detection and analysis program. EL and SS performed the syllable classification. MS conducted the statistical analyses and data presentation. JR rated the behavioral context. EL wrote the first draft of the manuscript. All authors contributed to the writing, editing, and revising of the final paper.

## FUNDING

The research was funded by a Human Frontiers Science Program Research Grant (RGP0058/2016) awarded to LW and SV. SS was funded by the Fulbright U.S. Student Program.

## ACKNOWLEDGMENTS

The authors want to thank Mirjam Knörnschild for helpful discussions concerning this manuscript and the LMU workshop for their help in constructing the

recording setup. We would also like to thank the two reviewers for their helpful and constructive feedback on earlier versions of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2019.00116/full#supplementary-material>

**Supplementary Figure S1** | Syllable diversity in the commonly occurring syllable classes. The different commonly occurring classes contain syllables with some structural variety. Here we want to give an impression about the different shapes syllables from any class can take. From top left to bottom right, example oscillograms (top) and spectrograms (bottom) of the following are displayed: (A) noisy, (B) long, and (C) short high entropy syllables (HE), (D) linearly downward frequency modulated (IDFM) syllable, (E,F) non-linearly downward frequency modulated (nIDFM) syllables, (G) regular sinusoidally frequency modulated (SFM) syllable, (H) SFM syllables with a downward-frequency modulated onset, (I) ascending and (J) short SFM syllables, composite syllables (CS) with a noisy element (K) at the end or (L) within the syllable, (M) short, and (N) long quasi-constant frequency (qCF) syllables, (O) quasi-constant frequency syllable with a steep onset (qCF\_so), and (P) noisy quasi-constant frequency (qCF\_n) syllable.

**Supplementary Table S1** | Measured and calculated acoustic parameters of the common syllable classes. For the 19 extracted spectro-temporal parameters, the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles (i.e., first, second, and third quartiles) are reported to represent data distribution.

**Supplementary Table S2** | Measured and calculated acoustic parameters of the rare syllable classes and the suppressed fundamental frequency class (SF). For the 19 extracted spectro-temporal parameters, the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles (i.e., first, second, and third quartiles) are reported to represent data distribution.

**Supplementary Table S3** | Behavioral contexts scored for 20 syllables per class. For each of the eight commonly occurring syllable classes, 20 videos were scored for the behaviors displayed by the bats during syllable emission. For the IDFM and nIDFM syllable classes only 19 instances could successfully be scored as the behavior during syllable emission was performed outside the field of view of the camera in the remaining two cases. A single vocalization from the HE class was emitted in a neutral behavioral context, which was scored as “brachiating on walls or ceiling”. Other than that, all's syllables were emitted either in a prosocial or an antagonistic behavioral context. From the ethogram of 56 detailed behaviors, which was used for the behavioral scoring, only 23 behaviors were observed during syllable emission.

## REFERENCES

- Barclay, R. M. R., Fenton, M. B., and Thomas, D. W. (1979). Social behavior of the little brown bat, *Myotis lucifugus*. *Behav. Ecol. Sociobiol.* 6, 137–146
- Bastian, A., and Schmidt, S. (2008). Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am.* 124, 598–608. doi: 10.1121/1.2924123
- Behr, O. (2006). The vocal repertoire of the sac-winged bat, *Saccopteryx bilineata*. Doctoral thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.
- Behr, O., and Von Helversen, O. (2004). Bat serenades - Complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behav. Ecol. Sociobiol.* 56, 106–115. doi: 10.1007/s00265-004-0768-7
- Bohn, K. M., Boughman, J. W., Wilkinson, G. S., and Moss, C. F. (2004). Auditory sensitivity and frequency selectivity in greater spear-nosed bats suggest specializations for acoustic communication. *J. Comp. Physiol. A Sens. Neural Behav. Physiol.* 190, 185–192. doi: 10.1007/s00359-003-0485-0
- Bohn, K. M., Schmidt-French, B., Ma, S. T., and Pollak, G. D. (2008). Syllable acoustics, temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats. *J. Acoust. Soc. Am.* 124, 1838–1848. doi: 10.1121/1.2953314
- Bohn, K. M., Wilkinson, G. S., and Moss, C. F. (2007). Discrimination of infant isolation calls by female greater spear-nosed bats, *Phyllostomus hastatus*. *Anim. Behav.* 73, 423–432. doi: 10.1016/j.anbehav.2006.09.003
- Boughman, J. W. (1997). Greater spear-nosed bats give group-distinctive calls. *Behav. Ecol. Sociobiol.* 40, 61–70. doi: 10.1007/s002650050316
- Boughman, J. W., and Wilkinson, G. S. (1998). Greater spear-nosed bats discriminate group mates by vocalizations. *Anim. Behav.* 55, 1717–1732. doi: 10.1006/anbe.1997.0721
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *J. Zool.* 288, 1–20. doi: 10.1111/j.1469-7998.2012.00920.x
- Davidson, S. M., and Wilkinson, G. S. (2004). Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Anim. Behav.* 67, 883–891. doi: 10.1016/j.anbehav.2003.06.016

- de Cheveigné, A., and Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *J. Acoust. Soc. Am.* 111, 1917–1930. doi: 10.1121/1.1458024
- Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631.
- Esser, K. H. (1994). Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *Neuroreport* 5, 1718–1720.
- Esser, K. H., and Lud, B. (1997). Discrimination of sinusoidally frequency modulated sound signals mimicking species specific communication calls in the FM bat *Phyllostomus discolor*. *J. Comp. Physiol. A* 180, 513–522.
- Esser, K. H., and Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae) - evidence for acoustic learning. *Ethology* 82, 156–168.
- Esser, K. H., and Schubert, J. (1998). Vocal dialects in the lesser spear-nosed bat *Phyllostomus discolor*. *Naturwissenschaften* 85, 347–349. doi: 10.1007/s001140050513
- Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: a comparative approach to the emergence of language. *Front. Psychol.* 7:1393. doi: 10.3389/fpsyg.2016.01393
- Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Paškunis, A., et al. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: evidence for acoustic universals. *Proc. R. Soc. Biol. Sci.* 284, 1–9. doi: 10.1098/rspb.2017.0990
- Firzlaff, U., Schörnich, S., Hoffmann, S., Schuller, G., and Wiegreb, L. (2006). A neural correlate of stochastic echo imaging. *J. Neurosci.* 26, 785–791. doi: 10.1523/JNEUROSCI.3478-05.2006
- Forman, G., and Scholz, M. (2010). Apples-to-apples in cross-validation studies. *ACM SIGKDD Explor. Newsl.* 12:49. doi: 10.1145/1882471.182479
- Friard, O., and Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol.* 7, 1235–1230. doi: 10.1111/2041-210X.12584
- Gadziola, M. A., Grimsley, J. M. S. S., Faure, P. A., and Wenstrup, J. J. (2012). Social vocalizations of big brown bats vary with behavioral context. *PLoS ONE* 7:e44550. doi: 10.1371/journal.pone.0044550
- Gould, E. (1975). Neonatal vocalizations in bats of eight genera. *J. Mammal.* 56, 15–29. doi: 10.2307/1379603
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*, 2nd Edn. New York, NY: Springer New York.
- Hechavarria, J. C., Beetz, M. J., Macias, S., and Kössl, M. (2016). Distress vocalization sequences broadcasted by bats carry redundant information. *J. Comp. Physiol. A* 202, 503–515. doi: 10.1007/s00359-016-1099-7
- Heinrich, M., and Wiegreb, L. (2013). Size constancy in bat biosonar? Perceptual interaction of object aperture and distance. *PLoS ONE* 8:e61577. doi: 10.1371/journal.pone.0061577
- Hoffmann, S., Baier, L., Borina, F., Schuller, G., Wiegreb, L., and Firzlaff, U. (2008). Psychophysical and neurophysiological hearing thresholds in the bat *Phyllostomus discolor*. *J. Comp. Physiol. A* 194, 39–47. doi: 10.1007/s00359-007-0288-9
- Kanwal, J. S. (2009). “Audiovocal communication in bats,” in *Encyclopedia of Neurosciences*, ed L. R. Squire (Oxford: Academic Press), 681–690.
- Kanwal, J. S., Matsumura, S., Ohlemiller, K., and Suga, N. (1994). Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. *J. Acoust. Soc. Am.* 96, 1229–1254. doi: 10.1121/1.410273
- Knörnschild, M. (2014). Vocal production learning in bats. *Curr. Opin. Neurobiol.* 28, 80–85. doi: 10.1016/j.conb.2014.06.014
- Knörnschild, M., Behr, O., and von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften* 93, 451–454. doi: 10.1007/s00114-006-0127-9
- Knörnschild, M., Blüml, S., Steidl, P., Eckenweber, M., and Nagy, M. (2017). Bat songs as acoustic beacons - Male territorial songs attract dispersing females. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-14434-5
- Knörnschild, M., Feifel, M., and Kalko, E. K. V. (2014). Male courtship displays and vocal communication in the polygynous bat *Carollia perspicillata*. *Behaviour* 151, 781–798. doi: 10.1163/1568539X-00003171
- Knörnschild, M., Glöckner, V., and Von Helversen, O. (2010b). The vocal repertoire of two sympatric species of nectar-feeding bats (*Glossophaga soricina* and *G. commissarisi*). *Acta Chiropterol.* 12, 205–215. doi: 10.3161/150811010X504707
- Knörnschild, M., Nagy, M., Metz, M., Mayer, F., and Von Helversen, O. (2010a). Complex vocal imitation during ontogeny in a bat. *Biol. Lett.* 6, 156–159. doi: 10.1098/rsbl.2009.0685
- Kwiecinski, G. G. (2006). *Phyllostomus discolor*. *Mamm. Species* 1–11. doi: 10.1644/801.1
- La Val, R. K. (1970). Banding patterns and activity periods of some costa Rican Bats. *Southwest. Nat.* 15, 1–10.
- Lattenkamp, E. Z., and Vernes, S. C. (2018). Vocal learning: a language-relevant trait in need of a broad cross-species approach. *Curr. Opin. Behav. Sci.* 21, 209–215. doi: 10.1016/j.cobeha.2018.04.007
- Lattenkamp, E. Z., Vernes, S. C., and Wiegreb, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *J. Exp. Biol.* 221:jeb180729. doi: 10.1242/jeb.180729
- Luo, J., Goerlitz, H. R., Brumm, H., and Wiegreb, L. (2015). Linking the sender to the receiver: vocal adjustments by bats to maintain signal detection in noise. *Sci. Rep.* 5, 1–11. doi: 10.1038/srep18556
- Luo, J., Lingner, A., Firzlaff, U., and Wiegreb, L. (2017). The Lombard effect emerges early in young bats: implications for the development of audio-vocal integration. *J. Exp. Biol.* 220, 1032–1037. doi: 10.1242/jeb.151050
- Ma, J., Kobayasi, K., Zhang, S., and Metzner, W. (2006). Vocal communication in adult greater horseshoe bats, *Rhinolophus ferrumequinum*. *J. Comp. Physiol. A* 192, 535–550. doi: 10.1007/s00359-006-0094-9
- Murray, K. L., and Fleming, T. H. (2008). Social structure and mating system of the buffy flower bat, *Erophylla sezekorni* (Chiroptera, Phyllostomidae). *J. Mammal.* 89, 1391–1400. doi: 10.1644/08-MAMM-S-006.1
- Nelson, J. E. (1964). Vocal communication in Australian flying foxes (Pteropodidae; Megachiroptera). *Z. Tierpsychol.* 21, 857–870. doi: 10.1111/j.1439-0310.1964.tb01224.x
- Pfalzer, G., and Kusch, J. J. (2003). Structure and variability of bat social calls: implications for specificity and individual recognition. *J. Zool.* 261, 21–33. doi: 10.1017/S0952836903003935
- Prat, Y., Taub, M., and Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Sci. Rep.* 6:39419. doi: 10.1038/srep39419
- Rodenas-Cuadrado, P., Chen, X. S., Wiegreb, L., Firzlaff, U., and Vernes, S. C. (2015). A novel approach identifies the first transcriptome networks in bats: a new genetic model for vocal communication. *BMC Genomics* 16, 1–18. doi: 10.1186/s12864-015-2068-1
- Rodenas-Cuadrado, P. M., Mengede, J., Baas, L., Devanna, P., Schmid, T. A., Yartsev, M., et al. (2018). Mapping the distribution of language related genes FoxP1, FoxP2, and CntnaP2 in the brains of vocal learning bat species. *J. Comp. Neurol.* 526, 1235–1266. doi: 10.1002/cne.24385
- Rother, G., and Schmidt, U. (1985). Die ontogenetische Entwicklung der Vokalisation bei *Phyllostomus discolor* (Chiroptera). *Z. Säugetierkd.* 50, 17–26. doi: 10.1017/CBO9781107415324.004
- Russ, J. M., Jones, G., Mackie, I. J., and Racey, P. A. (2004). Interspecific responses to distress calls in bats (Chiroptera: Vespertilionidae): a function for convergence in call design? *Anim. Behav.* 67, 1005–1014. doi: 10.1016/j.anbehav.2003.09.003
- Smotherman, M., Knörnschild, M., Smarsh, G., and Bohn, K. (2016). The origins and diversity of bat songs. *J. Comp. Physiol. A* 202, 535–554. doi: 10.1007/s00359-016-1105-0
- Wetterer, A., Rockman, M. V., and Simmons, N. B. (2000). Phylogeny of phyllostomid bats (Mammalia: Chiroptera): data from diverse morphological systems, sex chromosomes, and restriction sites. *Bull. Am. Museum Nat. Hist.* 248, 1–200. doi: 10.1206/0003-0090(2000)248<0001:POPBMC>2.0.CO;2
- Wilkinson, G. S. (1995). Information transfer in bats. in “Ecology, Evolution and Behaviour Bats” eds P. A. Racey and S. M. Swift Symp. Zool. Soc. London 67, 345–360.
- Wilkinson, G. S. (2003). “Social and vocal complexity in bats,” in *Animal Social Complexity: Intelligence, Culture and Individualize Societies*, Chapter 12, eds

- F. B. M. de Waal, and P. L. Tyack (Cambridge, MA: Harvard University Press), 322–341.
- Wilkinson, G. S., and Boughman, J. W. (1998). Social calls coordinate foraging in greater spear-nosed bats. *Anim. Behav.* 55, 337–350. doi: 10.1006/anbe.1997.0557
- Wright, G. S., Chiu, C., Xian, W., Wilkinson, G. S., Moss, C. F., Gadziola, M., et al. (2013). Social calls of flying big brown bats (*Eptesicus fuscus*). *Front. Physiol.* 4:214. doi: 10.3389/fphys.2013.00214
- Yost, W. A. (1996). Pitch strength of iterated rippled noise. *J. Acoust. Soc. Am.* 100, 3329–3335.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Lattenkamp, Shields, Schutte, Richter, Linnenschmidt, Vernes and Wiegreb. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.