

# Calling dynamics and call synchronization in a local group of unison bout callers

Douglas L. Jones · Russell L. Jones · Rama Ratnam

Received: 30 August 2013 / Accepted: 24 October 2013 / Published online: 19 November 2013  
© Springer-Verlag Berlin Heidelberg 2013

**Abstract** In many species of chorusing frogs, callers can rapidly adjust their call timing with reference to neighboring callers so as to maintain call rate while minimizing acoustic interference. The rules governing the interactions, in particular, who is listening to whom are largely unknown, presumably influenced by distance between callers, caller density, and intensities of interfering calls. We report vocal interactions in a unison bout caller, the green tree frog (*Hyla cinerea*). Using a microphone array, we monitored bouts from a local group of six callers embedded in a larger chorus. Data were analyzed in a 21-min segment at the peak of the chorus. Callers within this group were localized and their voices were separated for analysis of spatio-temporal interactions. We show that callers in this group: (1) synchronize with one another, (2) prefer to time their calls antiphonally, almost exactly at one-third and two-thirds of the call intervals of their neighbors, (3) tolerate call collision when antiphonal calling is not possible, and (4) perform discrete phase-hopping between three preferred phases when tracking other callers. Further, call collision increases and phase-locking decreases, with increasing inter-caller spacing. We conclude that the precise phase-positioning, phase-tracking, and phase-

hopping minimizes acoustic jamming while maintaining chorus synchrony.

**Keywords** Frog breeding chorus · Male–male interactions · *Hyla cinerea* · Weakly coupled oscillators · Array-based sound source localization and separation

## Abbreviations

Hc *Hyla cinerea*  
SD Standard deviation  
SNR Signal-to-noise ratio

## Introduction

Vocal communication in large social groups involves multiple receivers and senders. These social situations, often called cocktail parties with human speech (Cherry 1953; Bregman 1990; Bronkhorst 2000) and choruses in animals (Gerhardt and Huber 2002; Greenfield 2002; Catchpole and Slater 1995) are constrained by the use of a common (acoustic) channel for communication. Consequently, the received sound is a mix of competing messages from which a desired message must be extracted. In many species of frogs, chorusing by aggregates of males serves to attract females and plays a critical role in sexual selection (Blair 1958; Ryan 1985; Gerhardt and Huber 2002). Females actively listen to callers and select a mate. This can be a difficult task when the call density is high because of significant spectral and temporal overlap of calls. Thus, in this highly competitive milieu, females would find it difficult to locate and separate individual calls

---

D. L. Jones · R. L. Jones  
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

D. L. Jones · R. L. Jones · R. Ratnam (✉)  
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801, USA  
e-mail: ratnam@illinois.edu

D. L. Jones · R. Ratnam  
Beckman Institute for Advanced Science and Technology,  
University of Illinois at Urbana-Champaign, 405 N. Mathews Ave., Urbana, IL 61801, USA

without some mechanisms for minimizing acoustic jamming. Sexual selection therefore posits that female requirements have driven the chorusing behavior of males so that individual callers can be readily detected, located, and identified.

Current theories propose that males use selective auditory attention and focus on a subset of callers, and choose the best opportunity to place calls while minimizing acoustic jamming (Ryan 1985; Greenfield and Rand 2000 in the Túngara frog (*Physalaemus pustulosus*); Brush and Narins 1989 in the Puerto Rican tree frog (*Eleutherodactylus coqui*); Schwartz 1993 in a neotropical treefrog (*Hyla microcephala*); Bates et al. 2010 in the American bullfrog (*Rana catesbeiana*); for reviews in anuran and orthopteran species see Greenfield 1994; Greenfield et al. 1997; Gerhardt and Huber 2002; Bee and Micheyl 2008). It should be noted that these theories are based on data from a small number of frog species and on a limited analysis of specific choruses. Not all choruses have the same structure (Greenfield 1994; Gerhardt and Huber 2002) and determining the particular structure of vocal interactions will require more extensive surveys over multiple species.

Vocal interaction in a social situation is a dynamical multi-way exchange, with sending and receiving influencing one other. The sender, who is also a listener, must find a strategy to broadcast a signal with the best chance of being received. In the absence of other sensory cues, an effective strategy is to parse the acoustic scene and demix the received signals so that competing callers and their timing are identified. This can be done using intensity, directionality, spectral, and timing cues, allowing an individual to focus attention on one or a few senders (Cherry 1953; Bregman 1990; Klump and Gerhardt 1992; Schwartz 1993; Greenfield 1994; Bronkhorst 2000; Gerhardt and Huber 2002). However, parsing an acoustic scene is not sufficient. It still leaves open the question of determining an effective strategy for broadcasting a signal. This work describes a possible strategy by which a small group of frogs can time their broadcast based on a rapid adjustment of the relative phase between different callers.

The strategies for broadcasting a message, upon receiving multiple messages, have received limited attention in natural frog choruses where the density of callers can be high (see Brush and Narins 1989). It is likely to be species-dependent because calling behavior is influenced by a variety of ecological and physiological constraints (see Ryan 1986; Greenfield and Rand 2000). Irrespective of the species, a determination of these strategies (Greenfield 1994; Greenfield et al. 1997) requires, at the least, detailed information on the spatial distribution of the callers and the vocal output of each of them. This is often difficult to obtain in the wild using conventional hand-held recorders. Therefore, most of the available data on call timing is

restricted to playback experiments in controlled settings, or passive recordings of frogs in known locations paired with microphones (Ryan 1985; Brush and Narins 1989; Schwartz 1991, 1993, 2001; Greenfield 1994; Greenfield and Rand 2000).

The problem of determining location and call timing for each frog in a large natural chorus has therefore remained largely intractable. As a result, many of the rules and strategies that have evolved for vocal interactions remain unknown (see Schwartz 1991; Klump and Gerhardt 1992; Greenfield 1994; Greenfield et al. 1997; Greenfield and Rand 2000; Boatright-Horowitz et al. 2000). For instance, it is known that frogs pay attention to a fixed subset of neighbors (Brush and Narins 1989; Greenfield et al. 1997; Greenfield and Rand 2000; Bates et al. 2010), but are they able to switch attention between subsets of callers (see Greenfield and Rand 2000, for a discussion)? And while it has been shown that the call oscillator of a frog is capable of rapid and precise changes in timing (Zelick and Narins 1983, 1985), only one study in a natural chorus of *Eleutherodactylus coqui* (Puerto Rican tree frog) has shown that individuals can rapidly modify their auditory attention by actively hunting for a suitable temporal gap within and between subsets of callers (Brush and Narins 1989, but see Schwartz 1993).

In natural settings, spatial separation between callers and the directionality of the individual call patterns can significantly influence vocal interactions (as suggested by numerous controlled playback studies, for example Greenfield and Rand 2000). While spatial distribution can be sometimes determined using multi-microphone and video techniques, extracting the individual vocal output, i.e., unmixing or blind source separation (see Comon and Jutten 2010), has remained a difficult problem. This has limited studies of vocal signaling in natural assemblies. Until recently, only one study had teased apart vocal interactions in a natural chorus using multiple microphones positioned close to several callers and shown that frogs pay attention to a subset of their neighbors (Brush and Narins 1989). Other studies have used microphone array techniques to localize individual callers from the differences in the arrival times of sounds at the microphones (i.e., using time-delay estimation). While these methods cannot be used to separate, i.e., unmix, the calls of the individuals they do provide a breakthrough over conventional single-microphone recordings. Notable among these studies on chorusing frogs are those by Grafe (1997) in the African reed frog (*Hyperolius marmoratus*), and by Simmons et al. (2008) and Bates et al. (2010) in the American bullfrog (*Rana catesbeiana*). Grafe (1997) used a 4-microphone array to record chorusing in male reed frogs so that he could monitor female choice. He did not study chorus dynamics although the data were available. The earliest use

of microphone arrays, explicitly for the purpose of locating frogs so that chorus dynamics can be analyzed, is from the Simmons group (the technique and analysis were first presented in Simmons et al. 2008, with further studies reported in Bates et al. 2010). Their most intriguing result shows that bullfrogs prefer to alternate calls (sequence of notes) with individual callers from a distant cluster, but synchronize or alternate notes with neighbors within their cluster.

It should be noted that localization does not imply unmixing of sources. By itself a location map cannot be used to unmix or separate sources without an additional source extraction filter. But the location map can be used in conjunction with acoustic data to disambiguate and correctly identify callers (without unmixing or source separation). Simmons et al. (2008) and Bates et al. (2010) visually examined time-delay maps from two pairs of microphones along with the spectrogram of the sounds recorded at one of the microphones. The use of such combined information made it possible to disambiguate and correctly identify callers and determine patterns of call interactions. This is a very useful method when the technical difficulty of trying to unmix or separate sources is not worth the cost. It does suffer from poor temporal resolution because location estimates require longer averaging times and therefore time-delay maps have reduced temporal precision. As with all methods, the best technique may depend on the problem at hand. Bullfrogs have relatively long call durations (about 500–600 ms) and slower calling rates in comparison with green tree frogs (reported here).

To unmix the calls, an additional filtering step is necessary where the waveform corresponding to each caller is extracted from the mixture, as if the other callers did not exist. Adding all the individual waveforms will recreate the original mixture. This is the purpose of the adaptive beamformer used in this report and originally reported in Jones and Ratnam (2009).

Details of the array processing methods used in these and other studies have been extensively reviewed by Jones and Ratnam (2009) and Blumstein et al. (2011). Recently, the group of Aihara developed a novel “sound imaging” technique called the “firefly” (Mizumoto et al. 2011). This method uses spatially dispersed microphones combined with light-emitting diodes and an image processing algorithm to locate callers. It serves as a useful alternative to microphone array-based approaches. Based on this approach, Mizumoto et al. (2011) analyzed the timing of Japanese tree frogs (*Hyla japonica*) in the laboratory and in the field. Although they analyzed only two calling individuals in both laboratory and field experiments, they showed that two individuals in close proximity are capable of call alternation (anti-phase synchronization) and synchronous calling, with the former being greatly preferred

over the latter. Earlier, the same group developed a non-linear coupled-oscillator model (Aihara et al. 2011) and demonstrated the same results for three captive Japanese tree frogs. When only two of the three frogs are calling, the tendency is to call in anti-phase synchrony and less often in synchrony (as also shown by Mizumoto et al. 2011). When three frogs are calling, the system demonstrates bifurcations with more complex patterns of call timing including anti-phase, triphase, 1:2 anti-phase, and in-phase synchronization. These results on Japanese tree frogs will be compared with our results.

Aihara et al. carried out unmixing of sound sources (Aihara et al. 2011; Mizumoto et al. 2011) using Independent Component Analysis or ICA (see Hyvarinen et al. 2001 for a review of the method). In the laboratory, each frog was paired with a nearby microphone (3 mics for 3 frogs in the case of Aihara et al. 2011 and 2 mics for 2 frogs in the case of Mizumoto et al. 2011). Matching the number of microphones with the number of sound sources is necessary for source separation using ICA (unmixing  $N$  sources requires  $N$  sensors). The array-based adaptive beamforming technique does not suffer from this constraint (see Jones and Ratnam 2009), although it has other limitations that will be discussed in this report.

Thus, many pressing questions in the evolution of male–male vocal interactions can be attacked if location and call-timing information can be *simultaneously* obtained from a natural chorus by passively recording sounds at will. The recently developed microphone array technique by these authors (Jones and Ratnam 2009) closes this gap (but see also Mizumoto et al. 2011 for another technique). Microphone array technology has the capability to deconstruct a frog chorus into its spatial and temporal components (as demonstrated in Jones and Ratnam 2009). In principle, the technology makes it possible to determine locations and unmix source waveforms (vocal output) of individuals within a natural chorus so that we can tease apart the strategies underlying vocal communication. It should however be noted that the number of calling individuals that can be analyzed depends on the size of the array. If the array does not spatially cover the entire chorus, then only a local group of frogs within the larger chorus can be analyzed, as is shown here.

Most choruses of frogs involve a stereotypical and repetitive calling pattern that serves to attract females (Blair 1958; Gerhardt and Huber 2002). In the species reported here, the American green tree frog (*Hyla cinerea*), individuals resort to unison bout calling (Schwartz 1991; Gerhardt and Huber 2002) where a local group calls in unison for some time before falling silent. The call is a single note and is repeated periodically. We report a 21-minute segment during peak chorus activity and closely examine the call timings within a local group of six frogs

that were part of a larger chorus distributed around the breeding site. In this time, there were 20 bouts, about one every minute with a bout duration of about 40 s. The individuals that participated in a bout were largely fixed and did not move once they entered the chorus. We report the detailed spatio-temporal deconstruction of the local bout and show that individuals within this bout are able to rapidly adjust their call timing in relation to members of the bout. More specifically, we show that in a local group of six frogs, there is preference for only three discrete phase slots separated by phase intervals of zero, one-third, and two-thirds of the inter-call interval. This finding has some similarities with the findings of Aihara et al. (2011) in the Japanese tree frog (*Hyla japonica*). Further, we show that *H. cinerea* is capable of rapidly switching between these phase points suggesting that the call oscillator may be capable of rapidly modifying call dynamics from one call interval to the next.

## Materials and methods

### Field site

We recorded calls from a local group of frogs embedded in a large assemblage of green tree frogs (*Hyla cinerea*) in a breeding pond (Creekfield Lake) at Brazos Bend State Park, TX (USA). The recording site was at the intersection of a causeway that bisected the lake and the eastern shore, approximately 29°22'30.30"N and 95°35'40.50"W, at an elevation of 49 ft (14.9 m) (Google Earth, Google Inc.). *H. cinerea* were distributed along the shoreline of the lake (approximately 1.6 km in extent). Although the breeding site was large and many tree frogs were heard at the peak of the chorus, the site itself was mostly inaccessible for much of its perimeter. Consequently, we could not determine the total size of the chorus. Equipment were setup and tested during daylight hours, and the site was cleared of personnel. Recordings were carried out continuously for about 5 h from 9 pm to 2 am, for the entire duration of the chorus on consecutive days. None of the frogs were handled or disturbed during the chorus, all lighting equipment including headlamps were extinguished, and silence was maintained. Permission to record chorusing activity was granted by Texas Parks and Wildlife Department.

### Microphone array

We deployed a microphone array system consisting of 15 microphones (Sennheiser MKE-2, 0.02–20 kHz, omnidirectional) deployed in three spatially dispersed modules of five microphones each. All modules were mounted on tripods and were identical. In each module, four microphones

were mounted on the ends of a 1.4-m cross-arm positioned 2.65–2.9 m above the ground, and a fifth microphone was mounted 1 meter below the cross at the end of a single cross-arm of 0.7 m. Microphone outputs in each module were amplified by five battery-powered pre-amplifiers (Sound Devices MP-1) housed in a box mounted below the tripod. The amplifier outputs from each module were fed to the data acquisition system via balanced XLR cables. Five microphones provide unambiguous source location in three dimensions (Jones and Ratnam 2009) and provide the basis for the modular design adopted here. The spatial coverage can be extended by adding more modules as needed. In this study, the array modules were deployed to cover an analysis area of about 150 m<sup>2</sup>. The coordinates of the microphones were measured with respect to an arbitrary origin. Speed of sound  $c$  (in m/s) at ambient temperature  $T$  (in °C) was estimated from the formula  $c = 20.0457(T + 273.15)^{1/2}$ .

### Data acquisition

Microphone data were acquired synchronously at a sampling rate of 20 kHz (National Instruments PXI 4498, 16-channel, 24-bit) by a data acquisition computer (National Instruments PXI-8186 controller running Windows XP, mounted in a PXI-1042Q chassis). Data acquisition programs were developed in LabVIEW (National Instruments Inc). All equipment were powered with DC (battery) sources. Data were analyzed offline using Matlab (The MathWorks Inc.).

### Source localization and source extraction

The theory, algorithms, and analysis procedures have been described in detail in an earlier report (Jones and Ratnam 2009). Signal processing and data analysis are performed offline. Briefly, in the first step, a localizer determines the location of each caller in three dimensions. In the next step, an adaptive beamformer “steers” the array beam in the direction of each source and selectively extracts that source while suppressing all other sources. Localization of callers is performed in the time domain, whereas beamformer extraction of the waveform for each caller is performed in the frequency domain. See Jones and Ratnam (2009) for details of the following steps in the procedure:

1. The microphone data were bandpass filtered to include the spectral band of the *H. cinerea* advertisement call (700–5,500 Hz).
2. Filtered data were run through a cross-correlator on a frame-by-frame basis (0.12 s) that analyzed pairs of microphones. We set a threshold for the maximum of the cross-correlation function so that any acceptable

frame was dominated by a single caller. For each caller, this provided an estimate of the arrival time delay between microphones.

3. The location of each caller was estimated from the time-delay estimate using a gradient descent procedure to minimize the total mean-squared time-delay error.
4. Location estimates were clustered and individual callers were identified.
5. Steering vectors (impulse response functions) were estimated for all sources from the cross-correlation matrix in those frames where only one caller was present.
6. The steering vectors and the cross-correlation matrix were used to calculate the optimal beamformer weights to be applied to the Fourier-transformed microphone data. The weights were applied to each source and the calls were extracted. An inverse Fourier transform recovered the beamformer output in the time domain (at the sampling rate of 20 kHz or 50  $\mu$ s resolution). This is the unique waveform for each caller.

This work focuses on the timing interactions between callers rather than the acoustical characteristics of the calls themselves. Therefore, the extracted waveforms were further processed to correct for timing delays and to extract the onset and offset times of the individual calls (indicator functions). These analyses are presented in the next section.

## Post-processing

### *Absolute and relative time frames*

There are two reference time frames in this analysis: (1) an absolute time frame, where all callers are timed with respect to a single master clock (universal time) and (2) a relative time frame, where all callers are timed with respect to any given caller (the local time).<sup>1</sup>

1. *Absolute time frame* The beamformer combines the microphone signals as described above and arbitrarily references an extracted source to one of the microphones. Thus, the source waveforms do not have the same time of origin due to propagation times between the spatially dispersed sources and microphones. Let  $M_i$  be the coordinates of the microphone extracting the waveform of frog  $i$  with coordinates  $F_i$ . Then the absolute time of origin of the waveform of  $i$  (with time-base  $t_i$ ) must be corrected by  $\Delta t_i = d(M_i, F_i)/c$ , where  $d(x, y)$  is the distance between the points  $x$  and

$y$ , and  $c$  is the speed of sound. The corrected time  $t_i - \Delta t_i$  over all frogs  $i$  will align the time of origin for the population to a universal or absolute time frame.

2. *Relative time frame* From the point of view of a “receiver frog  $j$ ,” the calls received from other frogs are with respect to a local or relative time frame rather than an absolute time frame. The local or relative time at the coordinates of frog  $j$  depends on the spatial separation between receiver  $j$  and sender  $i$ . If the distance between these frogs is  $s_{i,j}$ , then the time taken for the call of focal frog  $i$  to reach frog  $j$  is  $\Delta s_{i,j} = s_{i,j}/c$ . This correction must be applied to obtain the local arrival times of senders calls at the coordinates of frog  $j$ .

Absolute and relative time frames are used in combination or, on occasion, used separately. In this study, we frequently wish to determine the time difference between the call of frog  $j$  (occurring at mic  $M_j$  at time  $t_j$ ) in response to the call of frog  $i$  (occurring at mic  $M_i$  at time  $t_i$ ) given that their spatial separation is  $s_{i,j}$ . From the perspective of the receiver (frog  $j$ ), a correction  $\psi_{i,j}$  must be applied to the time difference  $t_j - t_i$  to determine the local time at frog  $j$ . From the preceding formulae, the correction is  $\psi_{i,j} = \Delta t_i - \Delta t_j - \Delta s_{i,j}$ , and the true time difference is  $t_j - t_i + \psi_{i,j}$ . This shift in time frames is applied extensively in much of the analysis presented here.

### *Indicator functions for callers*

This study analyzes the timing interactions between calling frogs. Thus, we are interested in determining the onset and offset times for each call in the extracted waveforms. A threshold was applied to the extracted waveform, and the time point at which the waveform exceeded the threshold was marked as the call onset time. Similarly, the return of the waveform below threshold marked the call offset time. Onset/offset time estimation was implemented using an algorithm with a threshold setting that was adjusted to eliminate false alarms and misses. The beamformer output selectively extracts a target sound (call) at sufficiently high target-to-background ratios so that SNR does not become an issue (see Jones and Ratnam 2009 for a detailed assessment of beamformer performance). Thus, it is possible to maintain a low threshold while assessing the call onset and offset times.

We constructed an indicator function for each frog from the sequence of estimated call onset and offset times. For every frog, the indicator function is a waveform sampled at the same rate as the extracted source. It takes the value 1 during a call, and 0 elsewhere.

<sup>1</sup> We thank the members of the Sound Communication and Behavior Group at the Institute of Biology, University of Southern Denmark, Odense, for alerting us to the time delay corrections.

### Analysis of call phase

Consider the sequence of calls of a frog  $m$ , received at the coordinates of frog  $r$ , with local call onset times  $t_{m,k}, t_{m,k+1}, t_{m,k+2}, \dots$ , where  $k$  is the index of calls in the sequence. Our goal was to analyze the temporal positioning of the calls of the receiver frog  $r$ ,  $t_{r,j}, t_{r,j+1}, t_{r,j+2}, \dots$  with respect to the calls of frog  $m$ . We first determine the calls of  $r$  such that  $t_{m,k} \leq t_{r,j} \leq t_{m,k+1}$  for some  $k$  and  $j$  (see Fig. 2a). Then we define the normalized phase of the caller  $r$  measured with respect to  $m$  as:

$$\phi_{m,r}(k,j) = \frac{t_{r,j} - t_{m,k}}{t_{m,k+1} - t_{m,k}}. \quad (1)$$

The normalization is carried out with respect to the instantaneous call interval of  $m$ , i.e.,  $t_{m,k+1} - t_{m,k}$ . Thus, we designate frog  $m$  as the focal frog and determine the timing of all other frogs with respect to  $m$  according to Eq. (1). This method for measuring phase with respect to a focal frog is identical to the method of Aihara et al. (2011), except that they report phase angle from 0 to  $2\pi$ . The analysis proceeds by equating each frog in the population to the focal frog with respect to which the timing of all other frogs is measured in a pair-wise fashion. For a population containing  $N$  frogs, there are  $N(N-1)$  such pairs. The average of  $\phi_{m,r}(k,j)$  over all  $k, j$  for the given pair  $(m, r)$  is denoted simply as  $\phi_{m,r}$ . For all pairs  $(m, r)$  in the population, the average is denoted as  $\phi$ . When we are interested in examining the sequential evolution of the phase between any given pair, we will use a simple notation  $\phi(i), \phi(i+1), \dots$ , etc. The usage of  $\phi$  will be clear from the context in which it appears.

### Calculation of vector strength

A measure of phase-locking to a periodic signal is provided by the vector strength,  $V$  (Goldberg and Brown 1969; Mardia 1972). It can be used to measure the extent to which any given caller is time-locked to any other caller.  $V$  varies from 0 for a uniform distribution (no phase preference) to 1 if all calls are perfectly phase-locked to one another (there is a unique preferred phase). We use the call onset time for each frog to determine  $V$  from the following equation:

$$V = \frac{1}{N} \left\{ \left( \sum_i^n R_i \sin\left(\frac{2\pi i}{n}\right) \right)^2 + \sum_i^n \left( R_i \cos\left(\frac{2\pi i}{n}\right) \right)^2 \right\}^{1/2}, \quad (2)$$

where  $n$  is the number of phase bins in the histogram,  $R_i$  the number of calls with phase in bin  $i$ , and  $N$  is the total number of calls for any given pair.

If there are multiple peaks in the phase histogram, the phases are wrapped to line up the peaks. The vector strength is then calculated from the wrapped phases.

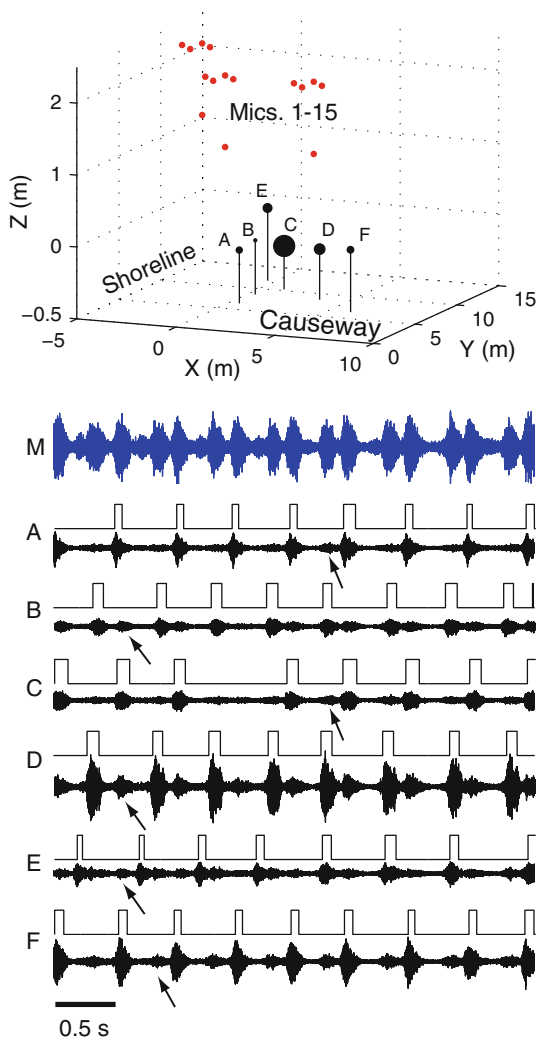
The probability of phase-locking in a set of random periodic calls emitted by  $N$  callers was also calculated. Random inter-call times for each of  $N$  callers were generated from a normal distribution with mean and standard deviation taken from the observed inter-call intervals. The starting time for each caller in the bout was randomized, and the pair-wise synchronization ( $V$ ) was determined over 1,000 repeated trials. This provides a bound on the strength of synchronization between randomly arranged call oscillators.

## Results

### Caller locations and call waveforms

Recordings were carried out on June 18 and 19, 2009, continuously for about 5 h from 9 pm to 2 am, encompassing the entire duration of the chorus. *H. cinerea* (*Hc*) was present all along the shore of the lake and the microphone array was positioned in the midst of a local group of six frogs. The ambient weather conditions were: surface temperature 79 °F (26.1 °C); dew point 71.1 °F (21.7 °C); relative humidity 77 %; barometric pressure 29.94 in (101.488 kPa). At this temperature, the speed of sound was 346.8 m/s.

An automated procedure for source localization and source extraction is available. However, the data presented in this study were additionally verified manually to ensure the accuracy of location estimates and call timings. This is a time-consuming process, but is necessary because the algorithms are still in a developmental stage and they need to be verified before being fully automated. Due to the large data throughput, we analyzed a 21-min segment of the data in great detail. We selected this segment based on the peak root-mean-square (RMS) level of the signals recorded at the microphones. Briefly, the RMS levels of the microphone signals were estimated for three microphones (one from each module, mic-1, mic-6, and mic-11) in 1-s blocks over the entire duration of recording (16,100 s or 268 min and 20 s). The RMS data were normalized, averaged over the three microphones, and smoothed with a low-pass filter (100 s time-constant) to provide the RMS level of calling activity averaged over three spatial locations. We observed that the RMS level ramped up over time and peaked at about 140 min from the start of the chorus (not shown). Thereafter, it was roughly constant for a further 40 min, and then it declined rapidly for the remainder of the duration. The data presented here are from a 21-min segment taken at the peak of the recorded energy,



**Fig. 1** Localizing callers and extracting individual calls using a 15-microphone array. *Top* Mean locations in three dimensions of six calling individuals (A–F, center of black sphere; diameter proportional to std. dev.), and locations of the 15 microphones (red). *Bottom* Example traces of 4 s duration taken from 20 to 24 s into bout# 20. Mixed sounds received at one microphone (M), and calls extracted by the beamformer (A–F, lower traces). Cross-talk from other sources (arrows) is usually 10–40 dB below the extracted source. The indicator function (A–F, upper traces) shows the on–off periods in the call sequence. The indicator functions pick only the desired source and do not include cross-talk

from 150 to 171 min (approximately 11:30–11:51 pm on June 19).

*Hc* is a unison bout caller (Schwartz 1991; Gerhardt and Huber 2002). That is, individuals in a local group call together in a bout before falling silent. This pattern is repeated over the duration of the chorus. At the peak of the chorus, the time between bouts was  $57.4 \pm 19.8$  s (mean  $\pm$  SD) with bout duration of  $44.6 \pm 10.5$  s (mean  $\pm$  SD). As a rule of thumb, the bout lasts for about 40 s and repeats every minute after a pause of 20 s.

**Table 1** Estimated coordinates of the frogs

Frog ID	Coordinates (m)			SD (m) $(x^2 + y^2 + z^2)^{1/2}$	Total calls (21 min)
	x	y	z		
A	0.51	6.38	0.85	0.16	1,138
B	0.37	8.62	0.89	0.09	949
C	1.15	10.24	0.38	0.54	1,031
D	3.62	8.66	0.71	0.27	1,114
E	−0.38	11.88	1.79	0.23	1,016
F	6.10	6.49	1.28	0.17	668

The localization algorithm estimated the 3-dimensional coordinates of the frog positions from the microphone array data. Also shown are the total number of calls output by the frogs in a 21-min period over a total of 20 bouts

**Table 2** Distance matrix

	A	B	C	D	E	F
A	0	2.24	3.94	3.85	5.64	5.61
B		0	1.87	3.25	3.46	6.12
C			0	2.95	2.64	6.28
D				0	5.24	3.35
E					0	8.44
F						0

The pair-wise distance between frogs (meters)

**Table 3** Nearest neighbor distances

A	B	C	D	E	F
B (2.24)	C (1.87)	B (1.87)	C (2.95)	C (2.64)	D (3.35)
D (3.85)	A (2.24)	E (2.64)	B (3.25)	B (3.46)	A (5.61)
C (3.94)	D (3.25)	D (2.95)	F (3.35)	D (5.24)	B (6.12)
F (5.61)	E (3.46)	A (3.94)	A (3.85)	A (5.64)	C (6.28)
E (5.64)	F (6.12)	F (6.28)	E (5.24)	F (8.44)	E (8.44)

The individuals nearest to a focal frog (column) are shown with distance (meters) in ascending order. B and C were closest neighbors, and E and F were furthest apart

A segment of one bout recorded at one of the microphones is depicted in Fig. 1 (waveform M, blue). This waveform is a mix of calls from six individual *Hc* (denoted by A, B, C, D, E, and F) forming a local group, and background noise from biotic and abiotic sources. Using data from a spatially deployed array of 15 microphones (Fig. 1, top panel, red circles), each caller within the group was located in 3-dimensional space (Fig. 1, top panel; location data are reported in Table 1). Spatial separation (Tables 2, 3) was smallest between frogs B and C (1.87 m) and greatest between E and F (8.44 m). The individual locations were fairly constant with mean positional error of 0.24 m. Error estimates (Table 1) include estimation error

from the localization procedure and possibly small movements made by the individuals.

Once the locations of callers had been determined, an adaptive beamformer (Jones and Ratnam 2009) selectively extracted the sound of each caller (Fig. 1A–F, lower trace shows extracted waveform, cross-talk is indicated by arrows). The time resolution was the same as the sampling period of digitization (50  $\mu$ s) and the cross-talk from other callers (target-to-interference ratio) was generally between  $-10$  and  $-40$  dB (Jones and Ratnam 2009). To investigate timing behavior, we created an indicator function for each caller rather than working with the extracted calls. These functions specify the on/off times for each caller (Fig. 1A–F, upper traces) and were determined from the extracted source waveform. The threshold for determining the on/off times and hence the indicator function was fixed at 70 % of the source amplitude. This removed the cross-talk while retaining the source. The onset time has a small delay (with respect to the actual onset time of the call) due to this procedure, while the offset time has a small lead. The error in the on/off time estimate is difficult to measure because of the ambient noise level. However, we believe that it is no more than a few milliseconds because the ramping times for the call are rapid. Further, the threshold is proportional to the source amplitude (see above), and this makes the estimates of the on/off times more consistent than if we were to use an absolute threshold. The calls are regularly spaced, with a median inter-call interval of 0.5 s [interquartile range (IQR): 59 ms and coefficient of variation (CV): 0.2]. The call duration was more variable with a median duration of 79 ms (IQR 23 ms and CV 0.24). The mean ratio of the call duration to the call interval was 0.15 or 15 %. The periodicity of calling is evident in Fig. 1A–F, with occasional change in timing. Individuals are more likely to delay calling, and less likely to call at a faster rate than normal. A total of 5,916 calls were extracted from the individuals over 20 bouts totaling 21 min (Table 1 provides the breakdown on the number of calls per frog).

### Timing and phase relationships

A glance at the indicator functions of A–F (Fig. 1) demonstrates that the periodic calling behavior of individuals in a local bout makes it difficult to determine who is leading and who is following. We measured the timing or phase relationship between pairs of individuals and calculated the normalized phase  $\phi_{m,r}$  of each frog ( $r$ ) with respect to a focal frog ( $m$ ) (Fig. 2a). For six frogs A–F this gives 30 pairs (excluding self-pairing). The normalized phase distribution for all pairs is shown in Fig. 2b. The horizontal bar below the abscissa indicates the mean normalized call duration over all calls (0.15). Following Fig. 2a, this is

obtained by averaging  $\phi_{m,r}(k, j)$  over all calls  $k$  for frog  $m$ , over all calls  $j$  for frog  $r$ , and over all 30 pairs of  $m$  and  $r$ .

To determine whether frogs preferred to maintain their phase over successive cycles, we examined the phase return map (Fig. 2c). The return map plots the current phase  $\phi(i)$  between frog  $m$  (focal) and frog  $r$  versus their next phase  $\phi(i + 1)$  over all  $i$  for all 30 pairs ( $m, r$ ). The return map can be interpreted as follows: (1) points along the diagonal indicate that phase is maintained between successive call periods, (2) points along the edges of the map indicate call collisions (synchrony), with (0, 0) and (1, 1) indicating successive collisions, and (3) points elsewhere indicate phase skipping. Note that the phase points 0 and 1 are the same due to wrap around. It should be noted that Fig. 2c is a recurrence map because it plots the recurrence of the phase of a given pair of frogs in successive call intervals. Aihara et al. (2011) also measure phase with respect to a focal frog but they present their data in the phase plane of a given pair (say AB) versus another pair (say AC). Further comparisons with the data of Aihara et al. (2011) are presented in the “Discussion”.

We determined the vector strength of the phase-locking as a function of pair-wise distance (Fig. 2d; inset shows detail for the most strongly phase-locked pairs). Each point in the graph indicates the vector strength of the phase between the indicated pairs (the first letter is the focal frog). Note that although the distances are symmetric,  $V(m, r)$  is not in general equal to  $V(r, m)$ . However, the values are not widely divergent.

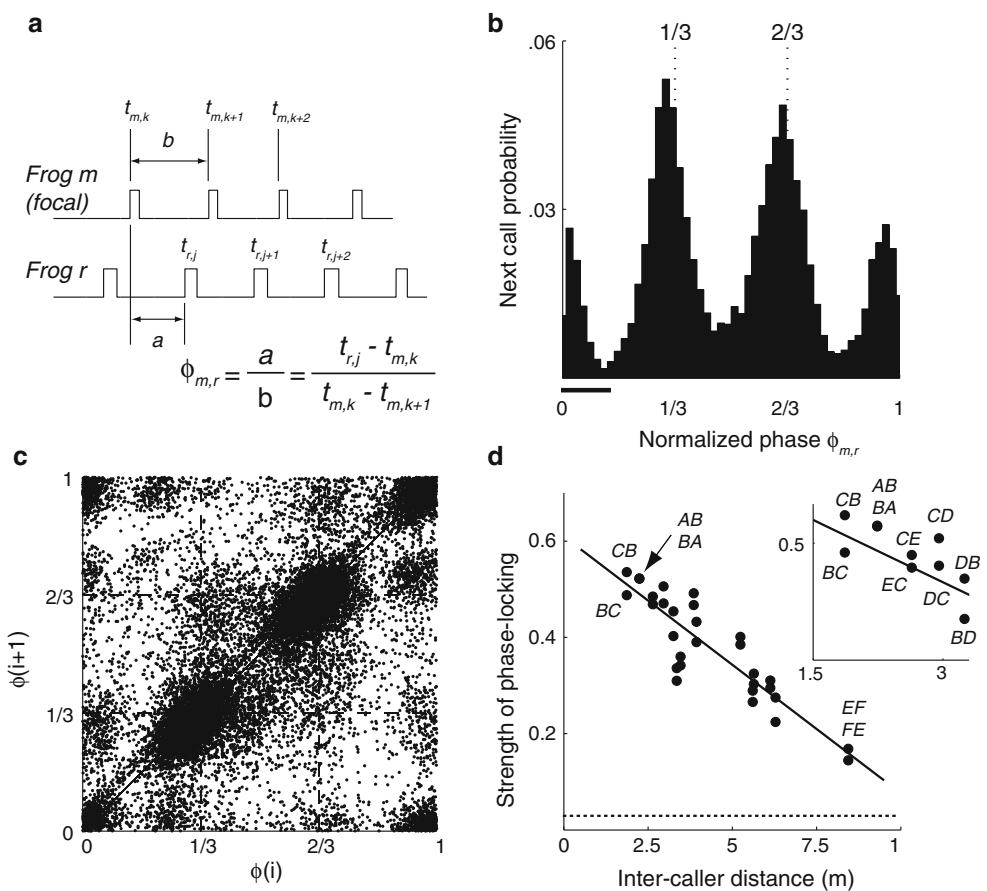
Vector strength was calculated for random call oscillators (see “Materials and methods”). We randomly generated 1,000 calls for each of six frogs using the observed mean inter-call interval of 0.5 s and standard deviation (SD) of 0.1 s. Pair-wise synchronization was estimated over all 30 pairs. This simulation was repeated 1,000 times and yielded  $V = 0.028 \pm 0.003$ . Most pairs, even the most weakly synchronized (F and E,  $V = 0.14$ ) exceeded the simulated control trials.

### The evolution of phase over time

The phase histogram (Fig. 2b) and phase return map (Fig. 2c) do not provide information on the temporal evolution of phase relationships. Even though certain phase points are preferred, it is by no means certain that individuals maintain their phase throughout the bout. Further, it is not known how inter-frog spacing, and the entry and dropping out of individuals in a bout affects phase preference. We illustrate the data with phase trajectories from three representative bouts. These examples provide insight into the dynamical interaction between individuals in a chorus (Fig. 3). The traces in each of the panels (Fig. 3a–c) are constructed by plotting the time evolution of  $\phi_{m,r}(k, j)$ , where  $m = C$  (Fig. 3a) and  $m = B$  (Fig. 3b, c). The



**Fig. 2** Phase histograms and phase-locking in a bout. **a** Phase relationship between a focal frog  $m$  and any other individual  $r$ . The normalization procedure converts phase to values between 0 and 1. **b** The phase distribution between all possible pairs (excluding self-pairings). Almost all frogs time their calls at  $\phi = 0$  (collision),  $1/3$ , or  $2/3$  with respect to another frog (horizontal bar, normalized duration 0.15 or 80 ms). **c** The phase return map shows the clustering of preferred phases between successive calls. Call collision lies along the boundaries. **d** Degree of phase-locking measured by the vector strength of the preferred phase, shown for each pair as a function of their spatial separation. Pairs are treated distinctly. The vector strength from a random interleaving of periodic calls is  $V = 0.028$  (dashed line)



remaining individuals ( $r$ ) participating in the bout are shown in the three panels. Each individual’s call is depicted by a point on the phase trajectory. To avoid excessive details only the initial portions of the bouts are shown in Fig. 3a–c (first 22, 24, and 16 s, respectively).

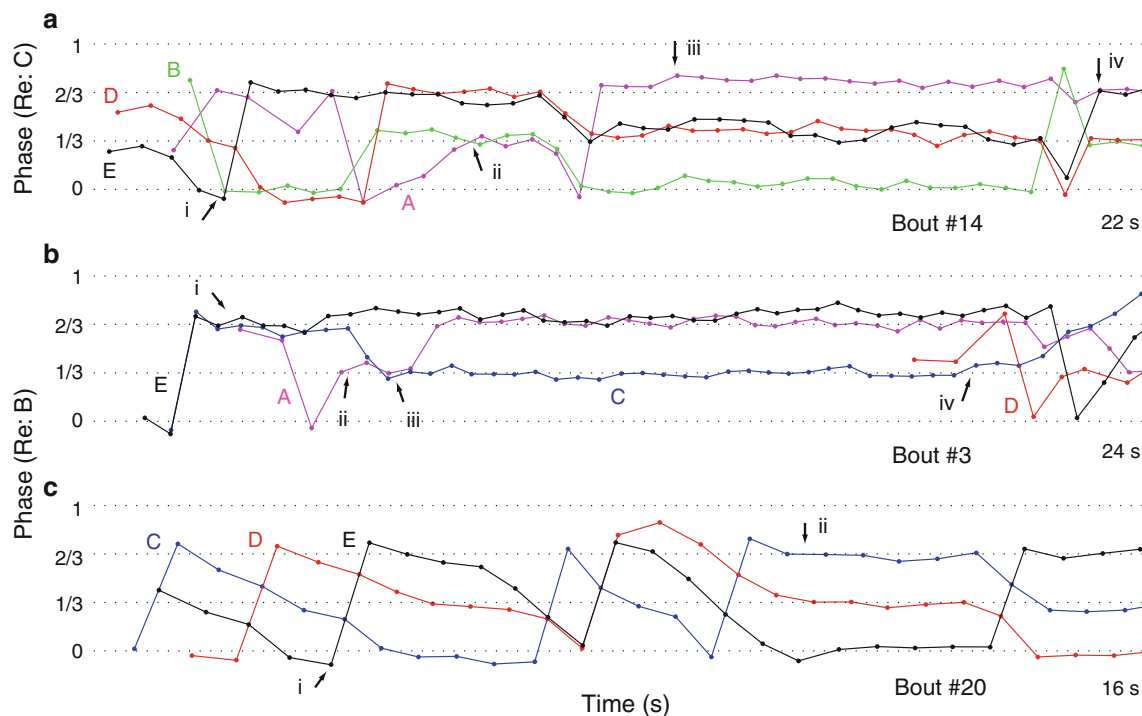
The panels illustrate that bouts may not include all individuals in the local group. During bout initiation, individuals hunt for a phase before settling into preferred phase slots (Fig. 3a, b). Some rearrangement may occur if any of the individuals switch phase, and sometimes individuals may drop out early or join later in the bout (D in Fig. 3b). The entry of an individual into a bout or its exit greatly affects the dynamics, as seen in the rearrangement of phases (Fig. 3b, after entry of D). Intriguingly, phase trajectories can provide information on the extent of coupling between the call oscillators of the frogs. For example, in Fig. 3c three frogs (C, D, and E) exhibit a linear drift in phase (with respect to B) suggesting that they are decoupled from B. These findings are discussed in detail further below.

**Discussion**

The calling patterns within a local group six frogs were analyzed to test for vocal interactions. A major drawback

of this study is that we could not analyze a larger group of callers or spatially cover a larger area. The spatial coverage is dependent on the number of microphones and their spacing. The array geometry used in this study restricted our focus to a local group of six frogs (denoted as A through F). We were able to localize and unmix the calls from this group and analyze their call dynamics and synchronization in great detail. We were also able to carry out a limited analysis on the effects of spatial separation on calling dynamics within this group. The spatial separation (Table 2) ranged from 1.87 m (BC) to 8.44 m (EF).

The calling patterns and interactions were analyzed using a semi-automated procedure that involved manual verification of the beamformer output. While this procedure is capable of being fully automated, it is still computationally demanding and further, it needed to be verified for accuracy. As a result of these constraints, we focused our analysis efforts on 21 min of calling at the peak of the overall chorus activity (from 150 to 171 min, see “Results”). We use the term “overall chorus” to refer to activity in the entire breeding pond as opposed to the local group analyzed here. Even though we cannot locate or unmix all the callers in the chorus, a measure of overall chorus activity can be obtained by calculating a running average of the RMS levels at the microphones over the



**Fig. 3** Switching of a follower's preferred phase as a function of time. Phase is measured with respect to a focal frog (C in **a**, and B in **b**, **c**). Frog F has been removed from these traces. Each point represents one call. All bouts exhibit a phase hunt in the initial period of the bout, followed by phase separation. Sometimes phase rearrangement and a hunt for a new phase can occur when another frog joins the bout or when collision cannot be tolerated. Collisions with focal frog occur at  $\phi = 0$  or 1; collisions between other pairs occur when their phase with respect to the focal frog is at  $\phi = 1/3$  or  $2/3$ . **a** (i) Phase hunting at the start of the bout, (ii) phase separation,

and (iii, iv) phase rearrangement. B–C and A–B distances are smallest and collisions between these pairs are usually not tolerated. D–E distance is much greater and collisions are permissible. **b** (i) Phase hunt with A, C, and E colliding at  $\phi = 2/3$ . (ii, iii) Phase separation and phase rearrangement. (iv) Entry of D causing disruption of phase and a new phase hunt. **c** (i) Linear phase drift of C, D, E with respect to B. While C, D, and E are phase-locked to each other at  $\phi = 0, 1/3,$  and  $2/3$ , they are not following B. (ii) Phase rearrangement with D and C locking to B at  $\phi = 1/3$ , and  $2/3$ , respectively, but E collides with B

duration of the chorus (see “Results” for details). The chorus peaked at 140 min into the recordings and lasted for about 40 min before declining gradually until activity had almost ceased after 2 am or about 268 min from the start of the recordings. To determine whether the results shown in Figs. 2 and 3 are applicable at other points in time (outside the 21-min analysis interval), we analyzed calling patterns at randomly selected intervals of time. The phase synchronization and alternation behavior that was observed in the 21-min interval at the peak of the chorus was largely maintained at other times as well, although the calling rates and numbers of callers participating in the local group was more variable in the early hours of the evening (before 11 pm) and the later hours (after 12 am). The data are not shown because the analysis was not exhaustive. An analysis of call dynamics within the local group is necessary during the initial transient in overall chorus activity (at  $t < 140$  min) and during the offset transient (at  $t > 180$  m), and is a topic for future study. We acknowledge the limitations of this study, and recognize that more data covering a bigger group of frogs, with a wider range of

inter-frog distance, and over the entire duration of the chorus will be necessary to draw firm conclusions.

#### Timing and phase relationships

Throughout this report, we use the term “call collision” and “synchrony” interchangeably without necessarily implying that they are either a form of cooperative interaction or a form of competitive interaction. Passive recordings, like those performed here, cannot determine motivation. Determination of motivation will require more intrusive experiments, such as passive recordings combined with controlled playback experiments. Figure 2b shows that calls are strongly phase-locked and the lagging frog in the pair is most likely to place its calls antiphonally at either one-thirds or two-thirds of the interval of the focal frog ( $\phi = 1/3$  or  $2/3$ ). The probability of call collision (synchrony) is shown by the peaks around  $\phi = 0$  and 1 and is lower than the probability of antiphonal calling. This indicates that collisions are not preferred over the  $1/3$  and  $2/3$  phase slots, but they will occur more frequently than

any other phase. This is surprising given that the call duration is 15 % of the inter-call interval (see horizontal time bar in Fig. 2b) and about six calls can be accommodated without collision. This is further discussed below. Note that for any pair of frogs ( $\phi_{m,r}$ ) that are calling in alternation, if  $\phi_{m,r} = 1/3$ , then it is not necessary that  $\phi_{r,m} = 1 - \phi_{m,r} = 2/3$ . This is because  $\phi_{m,r}$  is normalized with respect to the instantaneous period of  $m$ , whereas  $\phi_{r,m}$  is normalized with respect to the instantaneous period of  $r$  (see Eq. 1; Fig. 2a). It would be true only if the inter-call intervals of both frogs were identical and both callers maintained their preferred phase over successive call periods. The minima of the histogram are located at the normalized duration (Fig. 2b, horizontal bar), suggesting that calls are least likely immediately after the focal frog terminates its call.

Aihara et al. (2011) developed a nonlinear oscillator model (supported by experimental data) where they showed that a system of three coupled oscillators (frog callers) can find a stable triphasic synchronization pattern that corresponds to the phase separation of 0, 1/3, and 2/3 reported here (see “Materials and methods” for a brief description of the nomenclature followed by Aihara et al. 2011). In addition, their model predicts bifurcations leading to occupancy of states equivalent to (0, 1/2). When there are three frogs and the only stable states are (0, 1/2), the system becomes “frustrated” and only two pairs can maintain perfect anti-phase synchrony, namely (0, 1/2) and (1/2, 0), whereas the third pair must occupy the same phase slots (that is, they must synchronize). This is a stable state which they refer to as 1:2 anti-phase synchronization. Phase slots that corresponded to 1/2 are uncommon in our data ( $\phi = 1/2$  would occur at the trough lying between the 1/3 and 2/3 peaks in Fig. 2b). Thus, in our data, and as supported by Fig. 2b, the overwhelming number of phase slots are (0, 1/3, and 2/3) with the system becoming frustrated when there are four or more frogs. The Aihara et al. (2011) model cannot exhibit frustration for three frogs calling in triphasic synchrony because there are three slots, and each has a single occupant. We will go into the details of the frustrated occupancy of these phase slots when we discuss Fig. 3. Mizumoto et al. (2011) from the Aihara group examined the calls of two frogs in laboratory and field conditions. They report robust and stable anti-phase synchronization (0, 1/2) and occasional synchronization (0, 0). Aihara et al. (2011) reported that when there are two callers, synchronization (0, 0) is uncommon, and when there are three callers, synchronization (0, 0, 0) is rare.

While the (0, 1/3, 2/3) states observed here are supported by the nonlinear coupled-oscillator model of Aihara et al. (2011), it should be noted that: (1) the species are different (*Hyla japonica* in their case and *Hyla cinerea* in

ours), and so it is possible that their model may need a different set of parameters to produce the almost exclusively triphasic data that we observe. (2) They studied only three frogs (A, B, and C) kept in the laboratory in three cages, separated by 50 cm, and laid out in a line. Our recordings are carried out in a free-behaving group with much larger separation distances (minimum separation was 1.87 m, Table 2). Thus, the coupling parameters ( $K$  and  $\gamma$ ) in their oscillator model may need to be adjusted based on a more general spatial distribution. (3) We had more than three frogs in our group, and so the number of coupled oscillators is larger. Their coupled nonlinear oscillator system will have to be enhanced by adding more equations for the extra oscillators, with a bigger ( $6 \times 6$ ) coupling ( $K$ ) matrix. It should be noted that we have 30 distinct pairs of callers in this study, whereas only 3 distinct pairs are possible with 3 callers (as with Aihara et al. 2011). This makes the system greatly frustrated if only three stable phase slots (0, 1/2, 1/3) are available. It would be necessary to expand the Aihara et al. (2011) model for systems with more than three callers, and this is a topic for future research.

The return map (Fig. 2c) provides further evidence that frogs are strongly phase-locked and prefer to maintain their phase relationship with one another. Further, it shows that they do so over successive call periods. In particular, pairs of frogs show marked preference for maintaining synchrony (collisions,  $\phi = 0, 1$ ) or antiphonal calling between successive phases ( $\phi = 1/3, 2/3$ ). This is shown by the density of phase points along the diagonal. There are occasional transitions from antiphonal calling ( $\phi = 1/3$  or  $2/3$ ) to synchronized calling ( $\phi = 0$ ) and vice versa, and from  $\phi = 1/3$  ( $2/3$ ) to  $\phi = 2/3$  ( $1/3$ ), but surprisingly these are relatively rare events. Thus, callers appear to prefer discrete phase positions, overwhelmingly favoring the 1/3 and 2/3 phase slots. In fact, these two antiphonal phase points are so attractive that frogs are willing to suffer collisions even though a continuum of positions are available where calls may be positioned without collision. Aihara et al. (2011) do not report the evolution of phase over time (recurrence map) for a single pair of callers (as we do in Fig. 2c) but instead they report the phase plane behavior of two pairs of frogs (AB and AC) in the triphasic, 1:2 anti-phasic, and in-phase synchronous conditions (see Fig. 4b–e in Aihara et al. 2011). The recurrence map and the phase plane map are not the same and do not report the same information. An interesting contrast with the phase-occupancy observed by Aihara et al. (2011) is the occupancy observed for the states  $\phi = 0$  (in-phase synchronization) and  $\phi = 1/2$  (perfect anti-phase synchronization). They report that the former is relatively uncommon whereas the latter is common and stable. The opposite is

true for our data. The in-phase synchronous condition (0, 0) is fairly common in our data, and the anti-phase synchronization (0, 1/2) is uncommon. This can be seen in Fig. 2c as a clustering of phases at the corners of the map (in-phase synchronization) and the sparsity of points at the mid-points along the four edges of the map (anti-phase synchronization). These differences may be due to the different species examined in the two studies, or they may require a more complete bifurcation analysis for three and more callers.

The phase return map does not provide information on dependencies beyond two successive phases nor does it indicate the temporal evolution of the phase relationship between any pairs of frogs. This is taken up later.

The results shown in Fig. 2b and c conclusively demonstrate that *H. cinerea* prefer three-way interleaving and collision avoidance with their neighbors. That is, three individuals A, B, and C would prefer to position their calls in sequence so that B (and C) will occupy  $\phi = 1/3$  (and  $2/3$ ), with respect to A. This minimizes collision avoidance. However, if another frog were to join the bout, then two of them would rather suffer controlled collisions, as opposed to asynchronous calling which would result in uncontrolled collisions. We conjecture that *Hc* require a longer interval between successive phases than four-way interleaving would allow. This may be (1) to monitor other callers and to maintain synchrony, or (2) to allow for variability in call duration and timing, or (3) to allow for an inhibitory period following a call, or (4) to reduce the deleterious effects of forward masking. An inhibitory period has been proposed (see Brush and Narins 1989; Greenfield 1994; Greenfield et al. 1997; Greenfield and Rand 2000) and our data qualitatively support this model.

In this respect, if only three callers can interleave with one another, then the phase points  $1/3$  and  $2/3$  are theoretically the best phase locations because they achieve maximum temporal separation from each other. Other frogs attempting to follow will then prefer collision with any of the three. The data in Fig. 2b suggest, for example, that if a fourth frog were to join the bout at an intermediate phase, then it will result in a reordering of phases so that any two of the four will suffer collisions or one of them will retire from the bout. We examine this aspect in detail later.

Data from earlier studies have suggested that a caller is more tightly synchronized with neighbors who are closer than with those who are farther. However, barring a few reports (Brush and Narins 1989; Narins 1992; Schwartz 1993; Greenfield and Rand 2000; Bates et al. 2010) quantitative field data are not available. To test this hypothesis and measure the synchronization with neighbors, we measured the degree of phase-locking between neighbors as a function of distance. We hypothesized that

the phase-locking of calls between B and C (closest neighbors) would be greater than phase-locking between E and F (furthest apart). The sharpness of the peaks shown in Fig. 2b is a measure of the strength of phase-locking ( $0 \leq V \leq 1$ ) between pairs over all pairs. When the data were broken down into histograms for individual pairs (not shown), we observed that the height and width of the peaks demonstrated substantial variability between pairs. For some pairs, the peaks were narrow and sharp indicating strong phase-locking (large  $V$ ), whereas for others it was broad and more-or-less uniform indicating weak phase-locking (small  $V$ ). The strength of phase-locking between individual pairs is shown in Fig. 2d as a function of distance. As we hypothesized earlier, there is a strong distance effect in phase-locking ability. The results quantitatively confirm that nearest neighbors such as B and C (or A and B) are more tightly synchronized with one another than more distant pairs, such as E and F. Nearest neighbors maintain tighter synchrony with one another.

#### Phase-tracking and phase-hopping

The ability of callers to track the phase of their neighbors is most strikingly apparent in the phase-time plots (Fig. 3). They depict the time evolution of the phase of a frog with respect to a focal frog. In Fig. 3a, the phase relationship of frogs A, B, D, and E are shown with respect to focal frog C, and in Fig. 3b and c the phase relationship of frogs A, C, D, and E are shown with respect to focal frog B.

When a bout is initiated, there is a period of “phase hunting” where the callers attempt to find a preferred phase [(i) in Fig. 3a, b]. When a preferred phase is found, it is usually a synchronous phase (collision) or  $\phi = 1/3$  or  $2/3$  [(ii) and (iii) in Fig. 3a]. A long period of stable phase relationship can also be seen in Fig. 3b. There can be an abrupt phase rearrangement that takes place within two successive calls [(i) in Fig. 3a and (iii) in Fig. 3b]. Phase rearrangement can lead to a new preferred phase or to complete incoherence of the phases and the beginning of a new phase hunt [(iv) in Fig. 3a, b]. This is likely when the arrival of another individual leads to collisions that cannot be tolerated (entry of D into bout in Fig. 3b).

However, not all collisions are unfavorable as seen in Fig. 3a where E and D prefer to synchronize, and even switch phases at the same time (from  $\phi = 2/3$  to  $\phi = 1/3$ , about 9 s into the bout). E and D maintain synchrony for almost the entire duration until the terminal calls of the bout. This suggests that call collision may be preferred when other phase arrangements are less tolerable. But under what conditions are collisions favorable? Distance may be a factor because collisions are less likely between nearest neighbors, but is it the sole determining factor? In open air, sound attenuates by 6 dB when distance from the

source is doubled, and so a caller may tolerate collisions with more distant callers but not with callers who are nearby (for example, see Greenfield and Rand 2000). There is evidence for a distance effect on call collisions in the data. In Fig. 3b, A and C do not tolerate collisions [(iii) in Fig. 3b]. Frog A abruptly switches from  $\phi = 1/3$  to  $\phi = 2/3$  when C transitions from  $\phi = 2/3$  to  $\phi = 1/3$ . This results in collisions between A and E which seem to be tolerated by both callers. From Tables 2 and 3, it can be seen that the separation between AC (3.94 m) is less than the separation between AE (5.64 m) lending support to the idea that distance is a factor. Similarly, collisions between A and B are also unlikely (separation of 2.24 m) as seen in Fig. 3a, b. Indeed, A and B are among the most phase-locked pairs in the group ( $V = 0.5$ ; Fig. 2d). The pair most similar to AB is BC. The separation between B and C (1.87 m) is not very different from the separation between A and B, with similar phase-locking ( $V = 0.52$ ). These data provide strong evidence for a distance effect in collision tolerance as was suggested for the Puerto Rican tree frog (Brush and Narins 1989).

Inter-caller distance may not be the only factor in call collisions. During a bout, the situation is highly dynamic and other factors may determine whether collisions are tolerable. The pair BC again serves as an example. Although their collision probabilities are lower than probabilities for most other pairs (see above), from Fig. 3a it can be seen that B and C collide for a significant duration of the bout ( $\phi = 0$ ) even though the separation of BC is less than the separation of AC or AE. Thus, smaller separation may not always make collisions unfavorable. It is likely that factors, in addition to distance, such as directionality of the vocal beam pattern or presence of sound absorbing and sound deflecting objects may influence collision probability. The drawback of the microphone array method is that such conditions cannot be readily teased apart.

Figure 3c shows a linear phase drift of frogs C, D, E with respect to B (note the resetting of phase due to wrap around). A linear drift in phase occurs when two oscillators with different periods of oscillation become decoupled and run freely. The drift indicates that none of these frogs is following B, but they are following one another because of the nearly parallel phase drifts. However, C, D, and E settle into a preferred phase with respect to B [(ii) in Fig. 3c], although this is more likely due to B trying to track C, D, or E. In general, phase drift can be seen in a phase return map (such as Fig. 2c) as a line running parallel to the main diagonal. The distance from the diagonal is proportional to the difference in the time periods of the oscillators (normalized to the call period of the frog). In the case of the data reported here, the drift can be seen as a faint band running just above and parallel to the main diagonal. Since

the call periods of the frogs are nearly similar and do not demonstrate large variability, only a single band is present. The width of the band is proportional to the variability in the call periods.

The time evolution of the phases provides detailed information on the behavioral link between the auditory and vocal-motor systems. The consistent preference for collision or antiphonal calling (three phase slots) is remarkable given that there are other phase slots available but are unutilized. The availability of just three phase slots may be a physiological limitation imposed by vocal refractoriness, metabolic constraints, or auditory masking. More behavioral and neurophysiological work is needed to determine whether the limitations are purely auditory, or vocal-motor, or a combination of both. Two of the phase slots ( $\phi = 1/3$  or  $2/3$ ) provide a temporal window for clear broadcast whereas one ( $\phi = 0$ ) allows for controlled collisions with more distant callers. Any other combination will lead to uncontrolled collisions and reduce the availability of a time window that is free of interference. And so it is simply better to skip between these phases. Thus, *H. cinerea* is capable of discrete phase-hopping.

It would be interesting to extend Aihara et al. (2011) model and determine the evolution of phase as is shown in Fig. 3. The frustrated nature of the multi-caller system is readily observed in Fig. 3a, b. The desynchronized or free-running oscillators seen in Fig. 3c (with reference to a focal frog B) should also be predicted by their model when the coupling constant for one of the frogs (with respect to all other frogs) is set to 0.

The data presented here should also be contrasted with the data from Simmons et al. (2008) and Bates et al. (2010). *Rana catesbeiana* (American bullfrog) is not a unison bout caller. Their calls tend to have many notes of about 500–600 ms duration, and the time between calls is variable. Their longer note-duration and slower calling-rate (in contrast with tree frogs) may not require rapid adjustments in timing of the call oscillator. While we cannot compare timing data with the bullfrog, we can compare qualitative features of note and call placement. Both studies by the Simmons group have shown that bullfrogs appear to call in local clusters. Callers within a cluster generally prefer to call in alternation with callers from a more distant cluster (i.e., anti-phase synchronization). However, at the level of the individual notes of the call, callers may overlap or alternate notes with callers from their own cluster. Bates et al. (2010) suggest that note-overlapping may increase the strength of the amplitude modulations (AM, the envelope of the waveform) in the local cluster thus avoiding masking with clusters farther away. This could serve to attract female listeners to a local cluster. Tree frogs, on the other hand, may need to broadcast short-duration calls at relative high call rates with precise temporal control of the calling

rhythm while still maintaining chorus synchrony. These findings suggest that in-phase and anti-phase calling (or phase adjustment in general) may be adapted to suit the unique needs of a species.

Broadly, our data suggest that when a bout is initiated, individuals merely pay attention to their neighbors and try to minimize call collision primarily by three-way phase-locking and secondarily by controlled collisions with more distant neighbors if it is unavoidable. The most intriguing results of this study are that three-way phase-locking does not lead to random selection of phase. Instead the preferred phases are discrete with  $\phi = 0, 1/3$  and  $2/3$  (Fig. 2b). In time division multiplexing, these are the optimum temporal positions for broadcasting discrete packets from three senders at a constant rate. When there are four or more callers, these phase preferences are still maintained, but now some pairs of callers will suffer collisions with neighbors rather than hunt for vacant phase slots. These collisions typically take place with more distant neighbors so that masking effects are reduced. Thus, “discrete phase-hopping” is a useful way of maintaining controlled collisions (over random asynchronous collisions) so as to minimize acoustic jamming, without sacrificing chorus synchrony or a high rate of vocal output. Is the system really discrete? We cannot be certain because we need to study a larger group of individuals. However, one line of evidence supports this idea. The nonlinear oscillator model developed by Aihara et al. (2011) also suggests that bifurcations can lead to discrete phase-occupancy in stable states. The idea of a frog-call being a “discrete packet” is not new. It was originally proposed by Brush and Narins (1989) who studied chorus dynamics in Puerto Rican tree frogs.

The study reported here provides evidence that *H. cinerea* may have evolved a system for optimizing vocal communication using phase-positioning, phase-tracking, and phase-hopping. However, *H. cinerea* is a unison bout caller, and conclusions drawn from this paper may apply only to this species, and not to other anurans that are not unison bout callers. The notion that communication is optimized in vocally communicating anurans was first suggested by Brush and Narins (1989) in another tree frog species, *Eleutherodactylus coqui*. Our observations on *H. cinerea* support these earlier ideas but are based on limited data. More extensive field work is needed over multiple days in the breeding season, over multiple seasons, and across multiple breeding sites. Further, it is necessary to record from a larger group of frogs by increasing the array aperture. These are topics for future work, some of which are ongoing.

The microphone array technique for localizing and separating sources provides a powerful analytical tool to quantitatively examine vocal timing behavior (Jones and Ratnam 2009). This has so far not been possible except in a few studies where microphones were placed next to calling

frogs (Brush and Narins 1989; Schwartz 1993, 2001; Mizumoto et al. 2011; Aihara et al. 2011). In *H. cinerea*, a unison bout singer, individuals time their calls with respect to their nearest neighbors, preferring collisions when an antiphonal phase slot is not available. We show that these frogs are capable of rapid modification of their calling phase and closely time their calls with respect to their neighbors. More work is necessary in a larger bout to determine the effect of the entry of multiple participants. The most intriguing results pertain to the precise ability to shift the call timing to an alternate phase slot between two successive calls (discrete phase-hopping, Fig. 3). Precise adjustments in vocal timing have been noticed in the frog *Eleutherodactylus coqui* (Zelick and Narins 1983, 1985; Brush and Narins 1989), and along with the work reported here, open the door to new questions on the physiological connection between the auditory and vocal-motor systems in frogs. Finally, it should be noted that these results and conclusions apply only to *H. cinerea* which is a unison bout caller. Recordings from other species of anurans are necessary to determine the range of communication strategies that have evolved across species.

**Acknowledgments** This research was partly supported by the Norman Hackerman Advanced Research Program (NHARP) Grant No. 010115-0101-2007 from the Texas Higher Education Coordinating Board (to R. R.), and by funds from the College of Engineering, UIUC (to R. R.). We thank David Riskind and Steve Killian of Texas Parks and Wildlife Department for permission to collect chorus data in Brazos Bend State Park (Needville, TX). Special thanks to the members of the Sound Communication and Behavior Group at the Institute of Biology, University of Southern Denmark, Odense, for their helpful comments and critique of this research. We also thank the two anonymous reviewers for their constructive criticisms on this manuscript. They have helped to significantly strengthen the presentation of this work.

## References

- Aihara I, Takeda R, Mizumoto T, Otsuka T, Takahashi T, Okuno HG, Aihara K (2011) Complex and transitive synchronization in a frustrated system of calling frogs. *Phys Rev E* 83:031913
- Bates ME, Cropp BF, Gonchar M, Knowles J, Simmons JA, Simmons AM (2010) Spatial location influences vocal interactions in bullfrog choruses. *J Acoust Soc Am* 127:2664–2677
- Bee M, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *J Comp Psychol* 122:235–251
- Blair WF (1958) Mating call in the speciation of anuran amphibians. *Am Nat* 92:27–51
- Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark C, Cortopassi KA, Hanser SF, McCowan B, Ali AM, Kirsche ANG (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *J Appl Ecol* 48:758–767
- Boatright-Horowitz SL, Horowitz SS, Simmons AM (2000) Patterns of vocal response in a bullfrog (*Rana catesbeiana*) chorus: preferential responding to far neighbors. *Ethology* 106:701–712

- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge
- Bronkhorst AW (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica* 86:117–128
- Brush JS, Narins PM (1989) Chorus dynamics of a neotropical amphibian assemblage: comparison of computer simulation and natural behavior. *Anim Behav* 37:33–44
- Catchpole C, Slater P (1995) Bird song: biological themes and variations. Cambridge University Press, Cambridge
- Cherry EC (1953) Some experiments on the recognition of speech, with one ear and with two ears. *J Acoust Soc Am* 25:975–979
- Comon P, Jutten J (2010) Handbook of blind source separation. Academic Press, Burlington
- Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans. University of Chicago Press, Chicago
- Goldberg JM, Brown PR (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32:613–636
- Grafe TU (1997) Costs and benefits of mate choice in the lek-breeding reed frog, *Hyperolius marmoratus*. *Anim Behav* 53:1103–1117
- Greenfield MD (1994) Cooperation and conflict in the evolution of signal interactions. *Ann Rev Ecol Syst* 25:97–126
- Greenfield MD (2002) Signalers and receivers: mechanisms and evolution of arthropod communication. Oxford University Press, Oxford
- Greenfield MD, Rand AS (2000) Frogs have rules: selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). *Ethology* 106:331–347
- Greenfield MD, Tourtellot MK, Snedden WA (1997) Precedence effects and the evolution of chorusing. *Proc R Soc B* 264:1355–1361
- Hyvarinen A, Karhunen J, Oja E (2001) Independent component analysis. Wiley Interscience, New York
- Jones DL, Ratnam R (2009) Blind location and separation of callers in a natural chorus using a microphone array. *J Acoust Soc Am* 126:895–910
- Klump GM, Gerhardt HC (1992) Mechanisms and function of call-timing in male–male interaction in frogs. In: McGregor PK (ed) Playback and studies of animal communication. Plenum Press, New York, pp 153–174
- Mardia KV (1972) Statistics of directional data. Academic Press, New York
- Mizumoto T, Aihara I, Otsuka T, Takeda R, Aihara K, Okuno HG (2011) Sound imaging of nocturnal animal calls in their natural habitat. *J Comp Physiol A* 197:915–921
- Narins PM (1992) Biological constraints in anuran acoustic communication: auditory capabilities of naturally behaving animals. In: Webster DB, Fay RR, Popper AN (eds) The evolutionary biology of hearing. Springer, New York, pp 439–454
- Ryan MJ (1985) The Túngara frog: a study in sexual selection and communication. University of Chicago Press, Chicago
- Ryan MJ (1986) Synchronized calling in a treefrog (*Smilisca sila*). *Brain Behav Evol* 29:196–206
- Schwartz JJ (1991) Why stop calling? A study of unison bout singing in a neotropical treefrog. *Anim Behav* 42:565–577
- Schwartz JJ (1993) Male calling behavior, female discrimination and acoustic interference in the Neotropical treefrog *Hyla microcephala* under realistic acoustic conditions. *Behav Ecol Sociobiol* 32:401–414
- Schwartz JJ (2001) Call monitoring and interactive playback systems in the study of acoustic interactions among male anurans. In: Ryan MJ (ed) Anuran communication. Smithsonian, Washington, pp 183–204
- Simmons AM, Simmons JA, Bates ME (2008) Analyzing acoustic interactions in natural bullfrog (*Rana catesbeiana*) choruses. *J Comp Psychol* 122:274–282
- Zelick RD, Narins PM (1983) Intensity discrimination and the precision of call timing in two species of neotropical frogs. *J Comp Physiol* 153:403–412
- Zelick RD, Narins PM (1985) Characterization of the advertisement call oscillator in the frog *Eleutherodactylus coqui*. *J Comp Physiol A* 156:223–229