Dynamics of communal vocalizations in a social songbird, the zebra finch (Taeniopygia guttata)

Julie E. Elie
Université de Saint-Etienne, Equipe de Neuro-Ethologie Sensorielle, Ecologie et Neuro-Ethologie Sensorielle/Centre de Neuroscience Paris-Sud Centre National de la Recherche Scientifique Unite Mistre de Recherche 8195, 23 rue Michelon, 42023 Saint-Etienne Cedex 2, France

Hédi A. Soula
Institut National de la Sante et de la Recherche Medicale U870, Institute National des Sciences Appliquées Lyon, Université de Lyon, Bât. Louis Pasteur, 20 Avenue Albert Einstein, 69621 Villeurbanne, France

Nicolas Mathevon and Clémentine Vignal
Université de Saint-Etienne, Equipe de Neuro-Ethologie Sensorielle, ENES/CNPS Centre National de la Recherche Scientifique, UMR 8195, 23 rue Michelon, 42023 Saint-Etienne Cedex 2, France

(Received 17 July 2010; revised 7 March 2011; accepted 7 March 2011)

Colonies or communities of animals such as fishes, frogs, seabirds, or marine mammals can be noisy. Although vocal communication between clearly identified sender(s) and receiver(s) has been well studied, the properties of the noisy sound that results from the acoustic network of a colony of gregarious animals have received less attention. The resulting sound could nonetheless convey some information about the emitting group. Using custom-written software for automatic detection of vocalizations occurring over many hours of recordings, this study reports acoustic features of communal vocal activities in a gregarious species, the zebra finch (Taeniopygia guttata). By biasing the sex ratio and using two different housing conditions (individual versus communal housing), six groups of zebra finches were generated, with six different social structures that varied both in terms of sex-composition and proportion of paired individuals. The results showed that the rate of emission and the acoustic dynamic both depended on the social structure. In particular, the vocal activity of a group of zebra finches depended mainly on the number of unpaired birds, i.e., individuals not part of a stably bonded pair. © 2011 Acoustical Society of America. [DOI: 10.1121/1.3570959]

PACS number(s): 43.80.Ka [MJO] Pages: 4037–4046

I. INTRODUCTION

Vocal communication has long been studied in the context of pairwise exchange between one sender and one receiver (McGregor, 2005). Identifying the information encoded in such signals has been a major objective for researchers, for instance, in demonstrating that acoustic parameters of vocalizations encode information about the identity and emotional state of the sender. However, communication is not only about a single sender and a single receiver. Audience effects are common, for example, in which the presence of additional conspecifics can influence a sender’s vocal behavior (Evans and Marler, 1994). In birds, it has also been shown that third-party “eavedroppers” can respond to vocal exchanges that they themselves are not participating in (Mennill et al., 2002). Multiple individuals may also be involved on both sides, such as when a group acts collectively as senders, directing acoustic signals to a group of receivers. When groups of animals are involved, any individual’s signals must be considered part of a larger communication network, one in which both vocalizations and responses to vocalizations may be influenced by activity of the group as a whole. To investigate this possibility, the current work examined the extent to which the vocal behavior of captive zebra finches (Taeniopygia guttata) is influenced by group composition.

So far, in strongly social species, most studies investigated the information encoded in acoustic signals exchanged between groups of animals, i.e., acoustic signals directed by a group of senders to a group of receivers. This type of conspicuous and coordinated communal vocalization has been described in numerous territorial species of mammals, especially primates (e.g., Bornean gibbons, Hylobates muelleri, Mitani, 1984) and carnivores (e.g., wolves, Canis lupus, Harrington and Mech, 1979; Frommolt, 1999 and female lions, Leo leo, McComb et al., 1994). But it has been reported also largely in birds that duet or chorus by more than two individuals (Farabaugh, 1982; Hall, 2000, 2004). Using playback experiments, several studies have shown that these coordinated territorial displays allow receivers to assess group size, coalition strength or sex ratio of the chorusing group (McComb, 1992; McComb et al., 1994; Seddon and Tobias, 2006; Hall and Magrath, 2007). In addition, acoustical analyses of these communal vocalizations have shown that the spectral parameters could also encode a group signature (Baker, 2004).

Besides these coordinated vocal displays, little is known about the information potentially conveyed by the apparently uncoordinated sound produced by an acoustic network as a
whole (i.e., multiple senders and receivers). Indeed in numerous animal species, the simultaneous vocal activity of several individuals results in apparently uncoordinated communal vocal production. This noisy sound may be produced by up to thousands of individuals, as it the case in choruses of frogs during the mating season (Marshall, 2003; Jones et al., 2009; Simmons et al., 2009; Bates et al., 2010), fish communities (Mann, 2003; D’spain and Berger, 2004; Locascio, 2004; Locascio and Mann, 2005), colonies of nesting birds (Aubin and Jouventin, 1998, 2002), and marine mammals (Schusterman, 1978; Southall et al., 2003).

In this present work, we try to relate some acoustic properties of such a background sound resulting from a network of vocalizing animals to the social relationships between individuals and/or the group composition. Up to now, these background sound properties have attracted little attention. The only experimental study that tackles this question showed that the communal sound of species living in coral reefs is used by fishes to locate and join the reef (Leis et al., 2002; Simpson et al., 2004). Apparently, no one has ever described any tool to analyze further such a complex acoustic signal. The purpose of the present paper is to introduce a method to analyze group vocal activity. This method was applied to groups of zebra finches (T. guttata) in order to test the hypothesis that the social structure of a group of birds can influence the acoustic parameters of the group vocal activity.

The zebra finch is a gregarious species native to subarid zones of Australia that relies heavily on acoustic communication for social interaction (Zann, 1996; Vignal et al., 2004, 2008; Elie et al., 2010). These songbirds are monogamous, i.e., male and female pair with an exclusive partner throughout their life.

Both males and females utter a wide repertoire of single-syllable calls while only males sing very stereotyped songs of several syllables (Zann, 1996). Among the calls categories, distance calls are the loudest calls given by zebra finches and can be heard from up to 100 m (Zann, 1996). These dimorphic calls convey the bird identity (Vignal et al., 2004, 2008). During most hopping movements, zebra finches also utter tet calls. These calls are softer and shorter at around 50 ms than distance calls (Zann, 1996). More rarely, one can hear calls emitted during aggressive encounters, by either the aggressor or the victim (Zann, 1996). Finally, paired birds may be heard emitting very soft calls (Whines, Arks, and Kakle) that they exclusively address to their partner around their nest during the breeding season (Zann, 1996; Elie et al., 2010).

In the wild, zebra finches form groups of dozens to hundreds of birds that gather in trees near watering points or feeding areas (Zann, 1996; J.E.E. and C.V., personal observation). These groups can emit an ensemble background sound that Zann (1996) described as “a soft background hum” (p. 197). Because of the amount of accumulated knowledge on both behavioral and neurobiological aspects of its acoustic communication (Clayton et al., 2009), the zebra finch is an interesting model to study communal vocalization and its relation to the social structure of the group.

In this study, six groups of zebra finches with varied social structures were set up. Five of them had various sex ratios and sizes and were housed in communal cages allowing social relationships. A control group was made of males housed in individual cages, allowing visual and acoustic contact but no physical interaction, thus preventing socialization within this group (Zann, 1996). Group vocal activity was recorded daily for several weeks and analyzed using custom-written computer programs that could detect vocalizations among hours of recordings.

First, the duration and the type of vocalizations used in each group were studied. Then, the rhythm of vocal emission was analyzed using vocalization rate and dynamics (i.e., pattern of vocalizations over time). Finally, the reliance of these acoustic parameters on three indices of the group social structure was examined: the number of birds, the sex ratio, and the percentage of paired birds within each social group.

## II. METHODS

### A. Subjects

Subjects were 54 zebra finches (T. guttata) including 34 males and 20 females. All birds were reared in groups containing young and adults birds of both sexes. Six groups were created for this experiment: five so-called social groups of varied sex ratio and number of birds 6M6F, 8M4F, 11M1F, 9F, and 9M (xM and yF being the number of males and females, respectively) and one group of ten isolated males referred to as 10IM (“I” standing for isolated condition). The sex ratio bias of each group was calculated as (M − F)/(M + F) that can be found in Table 1. All birds were used only once, except for the ten males that first composed the 10IM group and afterwards participated in the 11M1F group. The social group (6M6F, 8M4F, 11M1F, 9M, and 9F) was housed in a cage shaped like a 1 m$^3$ cube, whereas birds in the isolated condition (10IM) were individually housed in adjacent cages (40 × 25 × 35 cm) so that they could only...

### TABLE I. Recording procedures and social characteristics of experimental groups.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Number of recordings</th>
<th>Total duration of recordings (min)</th>
<th>Number of birds</th>
<th>Bias in sex ratio</th>
<th>Percentage of unpaired birds</th>
</tr>
</thead>
<tbody>
<tr>
<td>9M</td>
<td>28</td>
<td>2564</td>
<td>9</td>
<td>1</td>
<td>11.1</td>
</tr>
<tr>
<td>9F</td>
<td>30</td>
<td>2720</td>
<td>9</td>
<td>−1</td>
<td>33.3</td>
</tr>
<tr>
<td>6M6F</td>
<td>20</td>
<td>1933</td>
<td>12</td>
<td>0</td>
<td>41.6</td>
</tr>
<tr>
<td>8M4F</td>
<td>19</td>
<td>1836</td>
<td>12</td>
<td>0.33</td>
<td>25</td>
</tr>
<tr>
<td>11M1F</td>
<td>23</td>
<td>2208</td>
<td>12</td>
<td>0.83</td>
<td>50</td>
</tr>
<tr>
<td>10MI</td>
<td>6</td>
<td>870</td>
<td>10</td>
<td>1</td>
<td>100</td>
</tr>
</tbody>
</table>

interact visually and acoustically. Each group was housed and studied separately. Environmental conditions were the same for the six groups: the photoperiod was 14L/10D, food (seeds and fresh vegetables), water, and nest material were provided ad libitum, and the temperature was maintained between 25 °C and 28 °C. Experiments were performed under the authorization n°42-218-0901-38 SV 09 (ENES Lab, Direction Départementale des Services Vétérinaires de la Loire) and complied with the French legislation regarding experiments on animals.

B. Acoustic recordings

The acoustic activity of each group was recorded daily. While same-sex groups were allowed to stabilize for 2 weeks before recording began, mixed-sex groups were recorded as soon as the groups were created. The number of recording sessions and the total duration of recordings for each group are displayed in Table 1. Recording was conducted using two omni-directional microphones (Sennheiser MD42, Sennheiser electronic, GmbH, Wedemark, Germany) suspended from the ceiling. The microphones were positioned 20 cm from two different sides of the cage (social groups) or group of cages (isolated males) being recorded, and were connected to the two channels of a 16-bit, linear Pulse Code Modulation (PCM) stereo recorder (Marantz PMD670, 44.1 kHz sampling frequency, D&M Holdings Inc., Kanagawa, Japan).

C. Behavioral data

Behaviors of birds of the social groups were also monitored daily through focal observations performed by author J.E.E. Observation sessions (6M6F, n = 19; 8M4F, n = 19; 11M1F, n = 23; 9M, n = 32; 9F, n = 30) were conducted in the morning, from 30 min to 4 h after light onset. Each morning session consisted of focal observations of 10 min per individual with order of observation randomly assigned each day. Thus, a session lasted 90 min in the two same-sex groups and 120 min in the three mixed groups. The behavioral data gathered during these sessions aimed at identifying pair-bonded versus unpaired individuals. Several social behaviors were selected (Zann, 1996; Whitehead and Dufault, 1999): two birds cleaning each other’s feathers (alopreening), two birds perched side by side, often with erect feathers (clumping), two birds wiping each other’s beak (greeting beak fence), two birds performing a courtship dance by hopping to and fro between neighboring perches (Zann, 1996), two birds sharing their nest, one bird clearly addressing songs toward a particular individual (directed singing) and aggression (chase, threat, peck, aggressive beak fence; Zann, 1996). Birds that had established a pair-bond could be distinguished from unpaired individuals because they engaged in alopreening, often clumped side by side, performed greeting beak fence and courtship dance, shared the same nest but were never aggressive toward one another. Because birds in the 10M group could not physically interact, they were all considered as unpaired individuals. Thus, the social structure of each group of birds was described by the number of birds, the sex ratio bias, and the percentage of unpaired birds (Table 1). Note that in the same-sex groups 9M and 9F, pair-bonds were observed implying a percentage of unpaired birds different from 100%. The relation between these three indices of the social structure and the acoustic parameters of communal vocalizations were examined.

D. Automatic detection of vocalizations

Vocalizations were extracted from recordings using custom-written software. These programs were written in the PYTHON programming language (www.python.org) by author H.A.S. using open-source libraries. The scripts used in this paper are freely available from the authors upon request, and allowed processing of hours of recording in a batch process. Vocalization detection occurred through a series of three stages. The first identified all recorded events that exceeded a pre-determined energy threshold. Energy values were calculated from the power spectrum (1024-point Fourier transform) every 1000 samples (441 Hz). The energy contribution below 500 Hz was ignored. The second stage of detection was based on measuring the peak energy value of each analysis window, then expanding or contracting its length until mean energy over the window was greater than 10% of that peak value. This was a relatively conservative approach ensuring both that detection and the estimated duration of each sound was clearly above the pre-determined threshold value, and therefore robust. The third and final stage simply merged overlapping waveform segments. Together, the three stages produced start, end, and duration values for each sound event detected in the recording. In addition, the program could also extract each event as an individual sound file for further analysis.

The following checks were made to test the efficiency of the detection method. First, 12 files were artificially constructed using isolated calls between 5 and 20, which were extracted from a database of zebra finch sounds recorded in the laboratory. These calls were heterogeneous in spectral and temporal characteristics, with known durations and locations within each file. The automatic method successfully detected all calls in these artificial files. Second, the performance of the detection algorithm was examined empirically as a function of the threshold value used. To do so, threshold settings were systematically varied for each recording to determine the value at which a plateau was reached in the corresponding sound event count. Then, author J.E.E. manually counted the number of vocalizations in 12 randomly selected 3-min segments from the recordings, a total of 2294 sounds. As shown in Fig. 1, the threshold used in the experiment was chosen as that producing the best match between the manual count and automatic detection. Third, author J.E.E. manually determined the start and end points of 50 vocalizations in a 140-s group recording, finding these values to be in accordance with those provided by the detection routine. Finally, author J.E.E. manually verified whether 600 different sound events detected were in fact vocalizations. This total represented 100 samples from each group, with 25 samples selected randomly from each of 4 different recording days, also randomly selected. As shown in Fig. 2, 564 (94%) of the sounds events were in fact vocalizations, with only 36 (6%) representing background noise (e.g., wing-flapping, cage noise; see multimedia file noise.wav).
E. Durations of vocalizations

Durations of the sound events detected by the programs were first analyzed using a linear mixed effect model (LMM) with the group identity as a fixed factor (6 levels) and recording day as a random factor (126 levels representing the total number of daily recordings). Then, a second LMM was performed on the durations of vocalizations. Here, the three indices of social structure were fixed factors, with two levels for the number of birds (i.e., 9 versus 12), five levels of sex ratio bias (i.e., −1, 0, 0.33, 0.63, 1), and five levels for the percentage of unpaired birds (i.e., 11.1%, 25%, 33.3%, 41.6%, 50%). Random factors included the recording date (120 levels) nested within group identity (5 levels).

F. Repertoire use

Potential differences in the way birds from different groups used their vocal repertoire were studied. As mentioned in the Introduction, there are several kinds of vocal emissions. In addition to song syllables, the repertoire studied here was composed of the two most commonly used types of vocal emission, the distance and tet calls. Other vocal types were too infrequent to include in the work.

For both males and females, the distance call displays a harmonic structure and duration of around 100 ms [Fig. 2, male distance calls, this sound sample is available online at [url] DCM.wav, this is a file of type “wav” (16 Kb), and female distance calls, this sound sample is available online at [url]/DCF.wav, this is a file of type wav (24 Kb)]. As for tet calls, their spectrograms are displayed as series of chevron-shaped harmonics [Fig. 2, this sound sample is available online at [url/Tet.wav], this is a file of type wav (12 Kb)].

“Repertoire use” was defined as the relative rate of using the various vocalization types produced by zebra finches. This measure was determined for each group by examining the 564 automatically detected samples that had been manually identified as vocalizations among the 600 randomly selected events described above. Before group creation, distance calls of each bird and songs of males had been recorded. Based on this data bank, auditory identification, and spectrogram analysis, the same trained experimenter (J.E.E.) classified sound samples into four categories. Three of the categories were distance calls, tet calls, and song syllables [an example of this sample is available online at [url]/SS.wav, this is a file of type wav (12 Kb)]. A spectrogram for each category is displayed in Fig. 2. The fourth category included non-classifiable sound samples. They were either vocalizations emitted simultaneously with other birds or with a noise [this sound sample is available online at [url]/Noise.wav, this is a file of type wav (8 Kb)]. An example of a non-classifiable sound is also shown in Fig. 2 [an example of this sample is available online at [url/NC.wav], this is a file of type wav (12 Kb)]. Only 29 out of 564 sound samples were non-classifiable with no more than 8 sounds from any of the 6 groups. In the rest of the manuscript, the vocalization rate will be corrected by this error value of 6%. Apart from distance calls, tet calls, and song syllables, no other vocalization type was found in the 535 remaining sound events. Due to the low number of non-classifiable sound events, the repertoire was representative enough to estimate the percentage of each vocalization type per day for each group. Then, the repertoire use was compared between groups using an LMM on the percentage of vocalization type used per day, with the group and the vocalization type as fixed factors, and the date as a random factor. Finally, to identify which indices of the social structure could explain the differences between social groups, an LMM was conducted on the percentage of vocalization type with the three indices of social structure: the number of birds, the percentage of unpaired birds, and the sex ratio bias as fixed factors. The date nested in the social group identity was used as a random factor. First order level of interaction effects were considered in both LMMs.

G. Analysis of vocalization rate and emission dynamics

In addition to the mean vocalization rate, the autocorrelation function was also computed. It is another measure of
the temporal dynamics of the vocalizations. First, the instantaneous vocalization rate was calculated. The whole time course of a recording was quantized in bins of 250 ms. In each time bin, the value of the instantaneous vocalization rate was the number of vocalizations emitted during this bin. The autocorrelation of this signal was then computed over a total time of 10 s. Autocorrelation values significantly different from zero indicate a non-random pattern of vocalization occurrence. Whereas a positive value indicates that the same vocalization occurrence (either high or low vocalization rate) is regularly observed at a given temporal interval, a negative value indicates rate that constantly changes from high to low and the converse in the same interval. In the case of the vocalizations of zebra finches’ groups, the autocorrelation curves had very long tails. Thus, they could be fitted using non-linear, least-mean-square regression using the function \(1/(1+(4t)^a)\). The \(a\) value was a measure of the correlation of the signal. The lower the \(a\) value the higher the autocorrelation.

Finally, the instantaneous vocalization rate was highly variable within daily recordings. Periods of relative high vocal density were identified and named bursts of vocal activity \((n = 29,667)\). Several temporal parameters of these bursts were analyzed to investigate the vocal emission dynamics at a smaller scale than the vocalization scale. The start time of a burst was defined whenever the instantaneous vocalization rate first falls below the reference value after the burst started. The threshold used to define burst start and end times was chosen according to the sound pressure level based threshold used in a previous study on fish choruses (Locascio and Mann, 2005). Bursts were described using two measures: the burst length and the vocalization rate within bursts. Burst dynamics were also assessed by computing the burst rate within recordings and interval between the start times of successive bursts (“inter-burst interval”). For each recording, the coefficient of variation on inter-burst interval was calculated as the ratio between the standard deviation and the mean value of inter-burst interval.

To sum up, seven parameters measuring the rhythm of vocalization emission were computed: the vocalization rate within recordings, the \(x\) value given by the autocorrelation analysis, the burst length, the burst rate, the vocalization rate within bursts, the length of inter-burst interval, and the coefficient of variation on inter-burst interval. Two successive LMMs were used to compare these parameters between groups. The first LMM was defined using the group identity as fixed factor and the date as random factor. Then, to identify which indices of the social structure could explain differences between social groups, a second LMM was implemented using the three social indices as fixed factors and the date nested in the social group as a random factor.

### Statistical analysis

All statistical tests were conducted with R (R Development Core Team 2007). To perform LMM, lme() function of the nlme package was used. All post-hoc tests were done using paired \(t\)-test with Bonferroni correction. Then, the significance of correlations between indices of the social structure and acoustic measures was tested using Pearson’s test. Throughout the paper, values of acoustic parameters are given as mean ± standard deviation (SD).

### III. RESULTS

#### A. Durations of vocalizations

The durations of several thousands of vocalizations were analyzed for each group of birds. The mean duration of vocalizations was significantly different between all groups (Table II). In order to obtain a more precise picture, all the vocalizations from all recordings were classified according to their durations to obtain the duration histograms displayed in Fig. 3. All histograms were extremely comparable, but some differences could be observed. These differences can be explained by a higher fraction of vocalizations longer than 120 ms for the three groups displaying the highest mean durations. In addition, the three other groups that had lower mean duration had a higher fraction of vocalizations of duration less than 60 ms. These differences between histograms

---

**TABLE II.** Vocalization durations, measures of the rhythm of vocalizations, and statistical significance. The upper panel indicates mean and SD. The percentage of unpaired birds in each group is indicated in parenthesis. The lower panel indicates the results of analysis of variance (ANOVA) performed on the LMMs described in Sec. II (*p < 0.001; **p = 0.068). Superscript letters in the upper panel indicate significant differences between groups (post-hoc \(t\)-test, \(p < 0.05\)).

<table>
<thead>
<tr>
<th>Group type</th>
<th>Vocalization durations (ms)</th>
<th>Daily vocalization rate (per min)</th>
<th>(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9M (11.1%)</td>
<td>84 ± 34*</td>
<td>34.9 ± 10.7*</td>
<td>0.87 ± 0.1*</td>
</tr>
<tr>
<td>9F (33.3%)</td>
<td>87 ± 47*</td>
<td>63.1 ± 37.2*</td>
<td>1.31 ± 0.55*</td>
</tr>
<tr>
<td>6M6F (41.6%)</td>
<td>95 ± 44*</td>
<td>73 ± 31.9*</td>
<td>0.63 ± 0.13*</td>
</tr>
<tr>
<td>8M4F (25%)</td>
<td>86 ± 45*</td>
<td>39.2 ± 13*</td>
<td>0.73 ± 0.16*</td>
</tr>
<tr>
<td>11M1F (50%)</td>
<td>100 ± 50*</td>
<td>107.1 ± 21*</td>
<td>0.70 ± 0.08*</td>
</tr>
<tr>
<td>10M (100%)</td>
<td>98 ± 49*</td>
<td>103.8 ± 15.8*</td>
<td>0.48 ± 0.05*</td>
</tr>
</tbody>
</table>

Group type \(F_{1,124} = 19.74^{***} \) \(F_{1,124} = 40.25^{***} \) \(F_{1,124} = 18.53^{***} \)

| Group size | \(F_{1,1} = 18\) | \(F_{1,1} = 30.7\) | \(F_{1,1} = 7.53\) |
| Bias in sex ratio | \(F_{1,1} = 5.2\) | \(F_{1,1} = 4.49\) | \(F_{1,1} = 1.73\) |
| Percentage of unpaired birds | \(F_{1,1} = 29.9\) | \(F_{1,1} = 87.2^{**}\) | \(F_{1,1} = 1.32\) |
explained the slight but significant differences in mean durations. Nevertheless, the durations of vocalizations could not be related simply to any index of group social structure (Table II). Thus, social context had no visible impact on the durations of vocalizations.

B. Repertoire use

Since zebra finches can emit a wide range of calls and song syllables of various durations, the relative frequencies of these different vocalizations (the repertoire use) could be influenced by the group’s social structure. Indeed, females do not emit song syllables, so bias in sex ratio should influence the repertoire use. The results for all groups are displayed in Fig. 4.

The repertoire use was influenced by the group’s social structure ($F_{2,44} = 9.07, p < 0.001$). First, the bias in sex ratio was a major explanation for the differences between social groups in repertoire use ($F_{3,31} = 18.75, p < 0.001$). As anticipated, the all-female social group (9F) emitted no song syllables (0%) since all post-hoc t-tests on the proportion of song syllables between 9F and all other groups were statistically significant at $p < 0.001$. But surprisingly, the proportion of song syllables did not vary significantly according to the proportion of males (6M6F, 57.4% ± 14.5%; 8M4F, 57.4% ± 15.3%; 11M1F, 65% ± 12.6%; 9M, 55% ± 2%) with all post-hoc t-tests yielding p-value equal 1. As opposed to all groups with males—9.3% ± 6.8% for 6M6F, 8.2% ± 7.4% for 8M4F, 13.9% ± 13.3% for 11M1F, and 3% ± 6% for 9M—tet calls were predominantly used by 9F (50.7% ± 19.7%) with significance at $p < 0.01$ compared to all groups. For these same tet calls, all other post-hoc t-tests were not statistically significant at $p = 1$. The use of distance calls did not statistically differ between social groups (6M6F, 33.3% ± 10.9%; 8M4F, 34.4% ± 14.6%; 11M1F, 21.1% ± 10.2%; 9M, 42% ± 6.9%; 9F, 49.3% ± 19.7%).

Second, the number of birds had also an effect on the repertoire used by social groups ($F_{3,31} = 17.93, p < 0.001$). But this effect was mainly driven by the all-female group which emitted no song syllables. Indeed, the effect of the group size relied on the difference of proportions of song syllables between groups of 9 birds (i.e., 9M and 9F, 27.5% ± 29.4%) and groups of 12 birds (i.e., 6M6F, 8M4F, and 11M1F, 59.9% ± 13.4%). Finally, the proportion of unpaired birds also affected significantly the repertoire use of social groups ($F_{3,31} = 3.69, p = 0.022$). But, once again, this effect was only due to the 9F group emitting no song syllable and more tet calls than any other social group (all post-hoc t-tests, $p < 0.01$). Moreover, the percentage of each vocalization type did not differ significantly between the two all-male groups that mainly differed by the percentage of unpaired birds (post-hoc t-tests, $p > 0.3$). Thus, the repertoire use depended only on the sex ratio of groups and more precisely on the presence versus absence of males in the group.

C. Daily vocalization rate

The groups of zebra finches could be classified into three statistically different categories as shown in Table II. The 10IM group and the 11M1F social group formed the first category with high numbers of vocalizations (around 100 per min). A second category, with intermediate vocalization rates (between 60 and 70 per min), included the 6M6F and the 9F groups. Finally, the last two social groups (9M and 8M4F) constituted the last third category with much lower vocal activity (between 30 and 40 vocalizations per min). Finally, according to the previous Sec. III A, the two groups with the highest vocalization rates (11M1F and 10IM) also emitted longer vocalizations (Table II and Fig. 3). Thus, the total vocal activity of these two groups was particularly high compared to other groups. Moreover, as shown in Fig. 5, linear regression indicated that the vocalization rate was correlated with at least one index of the social structure: the percentage of unpaired birds ($R^2 = 0.84$; Table II). Thus, the more unpaired birds in a group, the more the group
vocalized. In accordance with this prediction, the all-male isolated group—10IM—vocalized more than the 9M group (Table II). Note that both groups only contained males. Furthermore, 10IM was considered to be composed solely of unpaired birds, as these individuals were physically separated, whereas 9M included only 11.1% unpaired birds.

Since the amount of communal vocal activity could at least discriminate between three categories of groups of zebra finches and was related to the social structure of these groups, Sec. III D further investigates the rhythm of emission, meaning the temporal structure of successive vocalizations.

D. Vocalization dynamics

The objective here was to separate two types of vocal activity. The first group activity could be described as a communal noise (uncorrelated chatter with a more or less constant average level of activity). The second group activity was the occurrence of intense vocal bouts (or bursts) separated by relatively long, quieter time periods. To give an example of such distinct vocal activities, Fig. 6 displays the vocal activity—instantaneous vocalization rate—of the two all-male groups (9M and 10IM) over a short time period. In order to make the comparison easier, the mean activity was subtracted from the signal. The 10IM group displayed more bouts of activity that were separated by quieter sequences compared to the 9M group. As Fig. 6 suggests, such dynamics were expected to give essential clues to the social structure. The autocorrelation (see Sec. II) of the vocal activity was computed to obtain quantitative figures. Mean autocorrelation values for all recordings are displayed in Fig. 7, while the values of $\alpha$ are displayed in Table II.

Autocorrelation curves of all groups were above the x-axis (i.e., zero), indicating correlated vocal dynamics. All the groups except the all-female one (9F) displayed positively correlated vocalizations over a relatively long period of time (up to around 5 s). Thus, the vocal activity of any of the groups was strongly influenced over a long time scale (several seconds) by past vocalizations or quieter time periods. As described above, vocal activity was following a burst-like pattern for all groups. In addition, the lower the $\alpha$, the more pronounced this pattern. Finally, autocorrelation curves allowed prediction of a mean duration of the bursts around 2 s for all groups except the all-female one (see next part for the detailed study of the burst-like activity).

The coefficient $\alpha$ was significantly different between groups (Table II). The 10IM group and the 9F group presented the most extreme values. The vocal activity of the 10IM group yielded the lowest value of $\alpha$, i.e., the more pronounced burst-like pattern in accordance with the result illustrated on Fig. 6. It was extremely and significantly

![FIG. 5. Vocalization rate (number of vocalizations per minute) versus percentage of unpaired individuals. The black line is the linear regression which yields an $R^2 = 0.84$.](image1)

![FIG. 6. Instantaneous rate of vocalization of the two all-male groups 9M (black curve) and 10IM (grey dashed curve). For each group, the average rate was subtracted from the instantaneous rate to allow comparisons.](image2)

![FIG. 7. (Color online) Autocorrelation of the instantaneous vocalization rate. Mean autocorrelation values ($AC(t)/AC(0)$) for all groups and for all recordings (mean ± SD). The time bin is 250 ms and total time is 5 s.](image3)
different from the 9M group (t-test $t_{14.9} = 13.4, p < 0.001$). Although when compared to all the other groups, post-hoc testing with Bonferroni correction only indicated a tendency (see Table II, $p = 0.057$). On the other hand, the vocal activity of the all-female group yielded the highest value of $x$ and was significantly different from all the other groups.

Although the indices of the social structure had no significant statistical effect (Table I), they could explain some variations of the $x$ values. First, the sex ratio influenced the vocal dynamic of birds, indeed, 9F and 6M6F groups displayed very different $x$ values, although they included a similar percentage of unpaired birds (33.3% in 9F, 41.6% in 6M6F). Moreover, $x$ values of the all-female group (9F) were extremely and significantly different from all the values of the other social groups containing males. Second, the possibility to establish relationships could also influence the vocal dynamics of birds. The $x$ values of the 10IM and 9M groups were significantly different and yet these two groups mainly differed by the fact that individually housed males could not pair with each other. So, vocalization dynamics could be related to the social structure of groups of zebra finches. The burst-like pattern of vocalizations emission is further analyzed in Sec. III E.

E. Burst-like activity

All the acoustic measures done on the burst-like activity are reported in Table III. By definition, a burst of activity was one that was high above the local background (see Sec. II). As the autocorrelation curves suggest in Fig. 7, bursts of vocal activity lasted around 1.4 s for all groups except 9F, which emitted the shortest bursts. Moreover, the rate of bursts in recordings differed significantly between groups. The 10IM and 9F groups, which presented the most extreme values of $x$ in the autocorrelation analysis, also showed the most extreme rates of bursts. Thus, the group that showed the lowest rate of autocorrelation (i.e., 9F) also showed the highest rate of bursts and the lowest duration of bursts. Thus, the 9F group presented the least pronounced burst-like pattern. However, the rate of bursts in social groups was not significantly influenced by any index of the social structure. Finally, the mean vocalization rate within bursts was highly correlated with the percentage of unpaired birds, in accordance with the results obtained for the daily vocalization rate ($R^2 = 0.90$).

In addition, hundreds of bursts were extracted and monitored by an experimenter. As expected, several birds participated in bursts of activity, and bursts were mainly composed of songs for all groups of birds, except for the all-female one. Their shorter bursts were only made up of distance calls.

With respect to the dynamics of bursts emission, the mean duration of the inter-burst interval was around a few tens of seconds, and two groups (11M1F and 10IM) emerged as different from all the others. The social structure of groups had no significant effect on the duration of inter-burst interval. Nevertheless, the extreme values of the 10IM and 9F groups were in accordance with the results of the autocorrelation analysis. Interestingly, the Coefficient of Variation (CV) on inter-burst intervals showed that bursts were not evenly displayed in the recordings. The CV was a useful measure to look for periodicity in temporal signal. A CV close to zero indicated a regular cyclic pattern in the burst occurrences. A CV close to 1 indicated a random, Poisson distribution, where the probability of a burst occurrence was constant and independent from previous occurrences of bursts. Most groups displayed the latter pattern (CVs close to 1) but some had even more variation, up to a CV of almost 3. This kind of variation indicated strongly correlated bursts and, as expected, happened for groups with high autocorrelation values (e.g., 10IM).

IV. DISCUSSION

This study used several acoustic tools to describe the communal noise of acoustic networks, meaning groups of multiple senders and receivers. The vocalization structure (duration and type) and various parameters of the rhythm of emission of vocalizations (rate of vocalizations, autocorrelation, bursts, and bursts dynamics) were extracted as relevant cues in communal vocalizations otherwise often considered

<table>
<thead>
<tr>
<th>Groups</th>
<th>Burst durations (s)</th>
<th>Burst rates (per min)</th>
<th>Vocalization rate within bursts (per second)</th>
<th>Inter-burst interval (s)</th>
<th>Inter-burst interval CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>9M (11.1%)</td>
<td>1.46 ± 0.96</td>
<td>2.55 ± 0.62&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>3.66 ± 0.9&lt;sup&gt;d&lt;/sup&gt;</td>
<td>23.38 ± 34.57&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.39 ± 0.17&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>9F (33.3%)</td>
<td>1.1 ± 0.36</td>
<td>2.79 ± 0.80&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>4.55 ± 2.28&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>21.1 ± 33.7&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.46 ± 0.47&lt;sup&gt;cd&lt;/sup&gt;</td>
</tr>
<tr>
<td>6M6F (41.6%)</td>
<td>1.33 ± 0.75</td>
<td>2.6 ± 0.55&lt;sup&gt;abc&lt;/sup&gt;</td>
<td>5.48 ± 1.57&lt;sup&gt;b&lt;/sup&gt;</td>
<td>22.62 ± 45.61&lt;sup&gt;bc&lt;/sup&gt;</td>
<td>1.88 ± 0.52&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>8M4F (25%)</td>
<td>1.59 ± 1.09</td>
<td>2.42 ± 0.44&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>4.27 ± 1.1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>24.41 ± 38.16&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.51 ± 0.2&lt;sup&gt;cd&lt;/sup&gt;</td>
</tr>
<tr>
<td>11M1F (50%)</td>
<td>1.27 ± 0.63</td>
<td>1.99 ± 0.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.49 ± 0.94&lt;sup&gt;c&lt;/sup&gt;</td>
<td>29.57 ± 56.06&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.73 ± 0.21&lt;sup&gt;bc&lt;/sup&gt;</td>
</tr>
<tr>
<td>10IM (100%)</td>
<td>1.27 ± 0.63</td>
<td>1.91 ± 0.39&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6.5 ± 0.68&lt;sup&gt;d&lt;/sup&gt;</td>
<td>29.67 ± 74.55&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.44 ± 0.39&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Group type

<table>
<thead>
<tr>
<th></th>
<th>$F_{1,124}$</th>
<th>16.28***</th>
<th>$F_{1,124}$</th>
<th>36.9***</th>
<th>$F_{1,124}$</th>
<th>13.67***</th>
<th>$F_{1,124}$</th>
<th>18.25***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group size</td>
<td>—</td>
<td>$F_{1,1} = 2.85$</td>
<td>$F_{1,1} = 30$</td>
<td>$F_{1,1} = 1.2$</td>
<td>$F_{1,1} = 2.34$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias in sex ratio</td>
<td>—</td>
<td>$F_{1,1} = 3.4$</td>
<td>$F_{1,1} = 4.6$</td>
<td>$F_{1,1} = 1.08$</td>
<td>$F_{1,1} = 0.14$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of unpaired birds</td>
<td>—</td>
<td>$F_{1,1} = 0.06$</td>
<td>$F_{1,1} = 35.7^{d}$</td>
<td>$F_{1,1} = 1.99$</td>
<td>$F_{1,1} = 0.65$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a very noisy signal. Applying the method on groups of zebra finches gave insights into how acoustic cues of the communal vocalizations could relate to the group’s social structure (size, sex ratio, percentage of paired birds).

In zebra finches, vocalizations can be sex-specific (as in the case of songs), vary between individuals (Vignal et al., 2004, 2008), and have different social functions (Zann, 1996). Thus, it was not surprising to find different mean durations of vocalizations and different repertoire use between groups that varied not only in the identity and the number of birds but also in sex ratio and in the percentage of paired birds. However, neither the durations of vocalizations nor the repertoire use were good discriminators of the social structure of the groups. Indeed, durations of vocalizations were not influenced by any index of the social structure. Moreover, the repertoire use could only discriminate the all-female social group from the other groups and did not differentiate between groups containing males.

On the other hand, acoustic parameters linked to the rhythm of emission of vocalizations represented good probes of the pairing status of birds within groups. First, these acoustic features well distinguished the most extreme perturbation of social links (physical isolation) from other social conditions. Compared to males housed in the same cage, the acoustic environment of the individually housed males was noisier (higher vocalization rate), extremely correlated, and burst-like. Second, the percentage of unpaired birds in social groups was positively correlated with the vocalization rate and influenced the acoustic dynamics. Thus, the unpaired birds drive the communal acoustic activity in a group of zebra finches.

Surprisingly, the rhythm of emission of vocalizations in the all-female group (9F) was completely different from all the other groups containing males, whatever the sex ratio. Although the 9F group presented mid-range values of daily vocalization rate, it showed the least pronounced burst-like pattern (the highest \( x \) value in the autocorrelation analysis, the highest rate of bursts, and the lowest duration of bursts). Moreover, mixed-sex groups did not differ from each other on this aspect, though they did not contain the same proportion of females. Thus, males are likely to drive the communal vocal rhythm in mixed-sex groups.

The vocalizations emitted by the zebra finches of this experiment were likely directed at within-group recipients. As such, each experimental group could be considered as a network of vocalizing animals where nodes represent individuals and links represent vocalization events. The programs of detection of vocalizations could not discriminate between senders. Thus, the six networks of vocalizing birds (six experimental groups) were studied at the scale of their communal activity, i.e., without knowing the identity of the sender of each detected vocalization. A next step in the study of communal vocalizations in this species will be to identify the individual identity or the sex of each sender. Having these more precise data would allow comparison of the network of social behaviors (e.g., aggressive or affiliative behaviors) with the network of vocal exchanges. Furthermore, it could be interesting to investigate the vocal network of various species of songbirds whose social structures that are similar to or different from that of zebra finches. Convergences in vocal dynamics may be highlighted.

V. CONCLUSIONS

Overall, the present study showed that the communal vocal output depended on the composition of the group in zebra finches. It remains to be tested whether this communal activity has a communicative function. Zebra finches are nomadic birds and very little is known on the stability and composition of wild groups of birds. Capture-recapture data are rare in this species but they suggest that zebra finches can travel very long distances. The mated pair is the only stable social unit and birds may join new groups between two breeding events (Zann, 1996). In this fission–fusion flocking system, individual birds could eavesdrop on the communal sound and use this “group signal” to decide whether or not they will join the group. Communal noise could thus participate in recruiting or repelling potential group members. Apparently no data are available and this point should be investigated in future studies.

The groups of zebra finches used in this study were very different from one another. The number of birds, the sex ratio of the group, and the proportion of paired birds were three important parameters of the social structure that varied at the same time. Although a simple relationship between each acoustic parameter and the indices of social structure could not be generated, acoustic cues linked to the rhythm of emission allowed discrimination between groups. Some parameters (e.g., vocalization rate) even allowed a definite grouping. The present results underline that acoustic parameters describing the communal sound of a whole vocalizing group could be a useful approach to the study of any social and highly vocal species.

ACKNOWLEDGMENTS

We warmly thank Colette Bouchut and Nicolas Boyer for all the technical support in the ENES laboratory. We are grateful to Brian Kelley for improving the English. This study was funded by the Agence Nationale de la Recherche (A.N.R., project “BIRDS’ VOICES” ANR-06-BLAN-0293-01) and Saint-Etienne Métropole. C.V. is supported by a Young Investigator Sabbatical of the Université de Saint-Etienne. N.M. is supported by the Institut Universitaire de France. J.E.E. is supported by the French Ministry of Research.
See supplemental material at http://dx.doi.org/10.1121/1.3570959 Document No. E-JASMAN-129-030105 for all the sound samples (of type “wav”). For more information see http://www.aip.org/pubservs/epaps.html