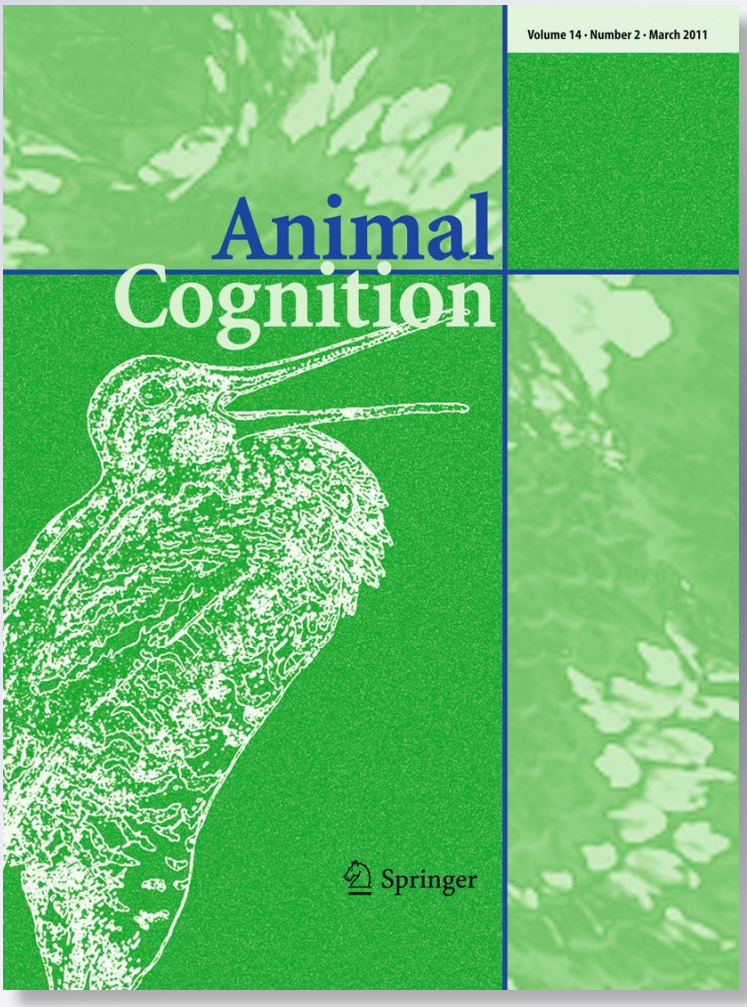
Closed-loop bird–computer interactions: a new method to study the role of bird calls

Animal Cognition

ISSN 1435-9448
Volume 14
Number 2

Anim Cogn (2010) 14:203-211
DOI 10.1007/
s10071-010-0353-6



Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Springer

Author's personal copy

Anim Cogn (2011) 14:203–211
DOI 10.1007/s10071-010-0353-6

ORIGINAL PAPER

Closed-loop bird–computer interactions: a new method to study the role of bird calls

Alexandre Lerch · Pierre Roy · François Pachet · Laurent Nagle

Received: 3 June 2010 / Revised: 1 September 2010 / Accepted: 8 October 2010 / Published online: 30 October 2010
© Springer-Verlag 2010

Abstract In the field of songbird research, many studies have shown the role of male songs in territorial defense and courtship. Calling, another important acoustic communication signal, has received much less attention, however, because calls are assumed to contain less information about the emitter than songs do. Birdcall repertoire is diverse, and the role of calls has been found to be significant in the area of social interaction, for example, in pair, family, and group cohesion. However, standard methods for studying calls do not allow precise and systematic study of their role in communication. We propose herein a new method to study bird vocal interaction. A closed-loop computer system interacts with canaries, *Serinus canaria*, by (1) automatically classifying two basic types of canary vocalization, single versus repeated calls, as they are produced by the subject, and (2) responding with a preprogrammed call type recorded from another bird. This computerized animal–machine interaction requires no human interference. We show first that the birds do engage in sustained interactions with the system, by studying the rate of single and repeated calls for various programmed protocols. We then show that female canaries differentially use single and repeated calls. First, they produce significantly more single than repeated calls, and second, the rate of single calls is associated with the context in which they interact, whereas repeated calls are context independent. This experiment is the first illustration of how closed-loop bird–computer interaction can be used productively to study social relationships.

Keywords Single call · Repeated call · Social interaction · Computer · Methodology · Canary · *Serinus canaria*

Introduction

The acoustic environment is a major source of information in animals and often influences their behavior (Maynards Smith and Harper 2003). Sound production is typically used for self-defense (e.g., stridulation in insects, Mitchell Masters 1979), localizing prey (e.g., echolocation in insectivorous bats, Aldridge and Rautenbach 1987), courtship (e.g., songs in frogs, Gerhardt and Huber 2002; and birds, Catchpole and Slater 1995), or social group cohesion (e.g., songs or calls in birds, Kroodsma 2004). So far, scientists have focused on information contained in conspecific vocalizations in a large number of taxa (Smith 1977).

Two main methods are traditionally used to understand the meaning of vocalizations. The first approach is to analyze the correlation between vocalization and behavior. Indeed, in many species, vocalization acts as an honest signal of emitter quality (Clutton-Brock and Albon 1979; Johnstone 1995). The other approach is observation of behavioral changes in the receiver (Castellano 2009; Grodzinski and Lotem 2007; Struhsaker 1967). To that end, researchers commonly use playback techniques, which allow the design of experiments to test unrelated

Electronic supplementary material The online version of this article (doi: 10.1007/s10071-010-0353-6) contains supplementary material, which is available to authorized users.

A. Lerch (✉) · L. Nagle
Laboratory of Compared Ethology and Cognition,
University of Paris Ouest Nanterre—La Défense,
200 avenue de la République, 92000 Nanterre Cedex, France
e-mail: alexandre.lerch@u-paris10.fr

P. Roy · F. Pachet
Sony Computer Science Laboratory of Paris,
6 rue Amyot, 75005 Paris, France

Author's personal copy

individuals (Catchpole and Slater 1995) independently of social relationships within the group.

However, even if playback techniques are efficient for focusing on the behavioral responses of receivers (Amrhein and Erne 2006; Brumm and Todt 2004), their use for interactions with the emitter is still unsatisfactory. Indeed, responses can be preprogrammed by researchers, but they still have to interfere to set off the mechanism (Mennill and Ratcliffe 2004; Sprau et al. 2010). This is unfortunate, as in nature communication is usually not unidirectional but includes mutual interaction: the behavioral responses of the receiver influence those of the emitter (Marler et al. 1986). Moreover, interactions between two individuals can be meaningful to other conspecifics (McGregor and Dabelsteen 1996); for example, in male songbirds, singing interactions influence the social ranks of competitors and the choices of eavesdropping females (Amy et al. 2008; Mennill et al. 2003; Naguib 2005).

In this paper, we propose a new experimental framework that overcomes the limitations induced by human intervention by introducing a closed-loop automatic interaction, based on fully automated sound analysis and playback techniques. We illustrate this framework with a study of the vocal production of female canaries *Serinus canaria* in the context of a simulated acoustic interaction.

In birds, two main types of vocalizations have been described: (1) songs, which are long, complex, and emitted mostly by males in the reproductive period, and (2) calls, which are shorter, simple, and produced by both sexes all year long (Catchpole and Slater 1995).

Although songs have been studied extensively for decades, few studies have been conducted on songbird calls. Nevertheless, calls are thought to be of great importance for social relationships. Marler (2004) reviewed the different contexts in which birds use calls. Firstly, many birds use alarm calls in the context of predator disturbance (Klump and Shalter 1984) and to warn relatives of potential danger. Secondly, different types of calls can inform receivers of the presence of spatial resources (e.g., roosting location, Wyndham 1980) or feeding resources (Hauser 1997). Lastly, it is also assumed that calls are used to maintain cohesion within the pair (Vignal et al. 2004), family (Beecher et al. 1981; Draganoiu et al. 2006), and flock (Bradbury 2003).

At the same time, the terminology "call" itself also hides a subtler categorization (Marler 2004). In domestic canaries, for example, a subdivision between "single calls" and "repeated calls" (Fig. 1) has been described by Mulligan and Olsen (1969). The use of several types of calls suggests transmission of different information on the emitter or on the environment and raises the question of referential communication. In vervet monkeys, *Cercopithecus aetiops*, three different alarm calls are produced when confronted by a leopard, a snake, or an eagle (Cheney and Seyfarth 1990). In the same way, Evans et al. (1993) have shown that alarm calls of domestic fowl (*Gallus gallus domesticus*) are adapted to the type of predator approaching. This article suggests that, in canaries, "single calls" are contact calls, whereas "repeated calls" are mostly used in a context of stress. A previous experiment corroborated this hypothesis, showing that female domestic canaries in reproductive state responded more frequently with single calls to the playback of a good-quality song (with a "sexy" phrase, Vallet and Kreutzer 1995; Vallet et al. 1998) than to other types of playback, whereas there was no significant difference in the production of repeated calls (Nagle et al. 2002). In contrast, the use of single and repeated calls during social interaction remains unclear. Indeed, to date, no satisfactory method has been able to mimic a real interaction between two vocalizing birds.

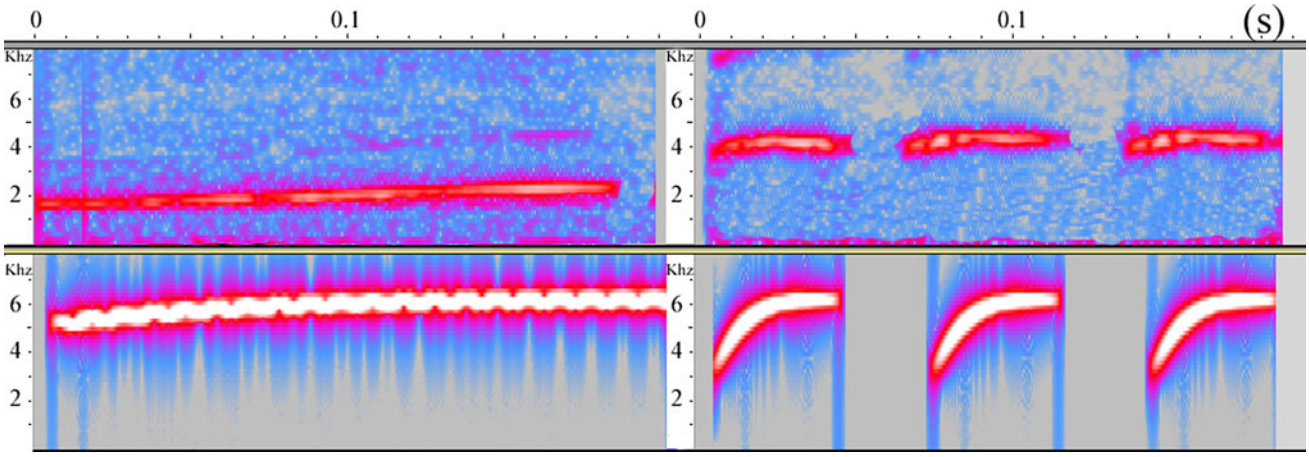


Fig. 1 Spectrograms of the played calls: *top left* real single call, *top right* real repeated call, *bottom left* synthesized single call, and *bottom right* synthesized repeated call

Author's personal copy

To study the role of calls in social interaction in a systematic fashion, we designed an experiment in which single birds were involved in a closed-loop acoustic interaction with a computer program that mimics the vocal behaviors of a conspecific without any intervention from the experimenter. In this closed-loop interaction, the computer iteratively listens to the bird, classifies its vocalization, and responds with a prerecorded sound. The details of the experimental protocol (including start and stop conditions) are described below in the "Experiment 1."

The experiments described below aim precisely at showing that: (1) birds do enter into an actual acoustic interaction with the system, and (2) this interaction confirms the role of single calls in social bird communication.

Methods

Software tools

The software used for the experiments consisted of four interconnected modules (Fig. 2). The recording module M1 gets an input signal from a microphone. The recorded signal is sent to the segmentation module M2 that detects salient acoustic events in real time. Each segment is in turn sent to a classification module M3 that labels segments as "single call," "repeated call," or "noise." After a segment has been classified, an appropriate response is computed and played by the response module M4 according to a predefined experimental protocol. We describe each module in detail below.

The recording module M1 is a MaxMSP patch (http://www.cycling74.com) that records a 16-bit, 44.1-kHz monophonic signal from the microphone. The recorded signal is sent to the segmentation module in real time (in frames of 1.5 ms).

The segmentation module M2 is a Java program that detects salient acoustic events based on amplitude variations in the signal and on the structure of canary vocalization, namely that canary single calls have typical duration of 200–350 ms, separated by several seconds of silence (Fig. 1, left), whereas repeated calls consist of a series of 3–5 very brief acoustic events (∼25 ms) separated by short silent frames (∼45 ms) (Fig. 1, right). The segmentation algorithm discards segments that do not fit these two categories, such as nonvocal noises produced by the bird in the cage (scratching, wing flapping, etc.).

The classification module M3 is a Java program that uses feature extraction and supervised machine-learning techniques to classify each segment output by M2 into one of the three categories "single call," "repeated call," and "noise."

In this work we use a particular technique for feature extraction called feature generation (Pachet and Roy 2009). The common approach to audio classification is to use standard acoustic features. However, we have shown in previous research that features generated for specific problems significantly outperform standard features for

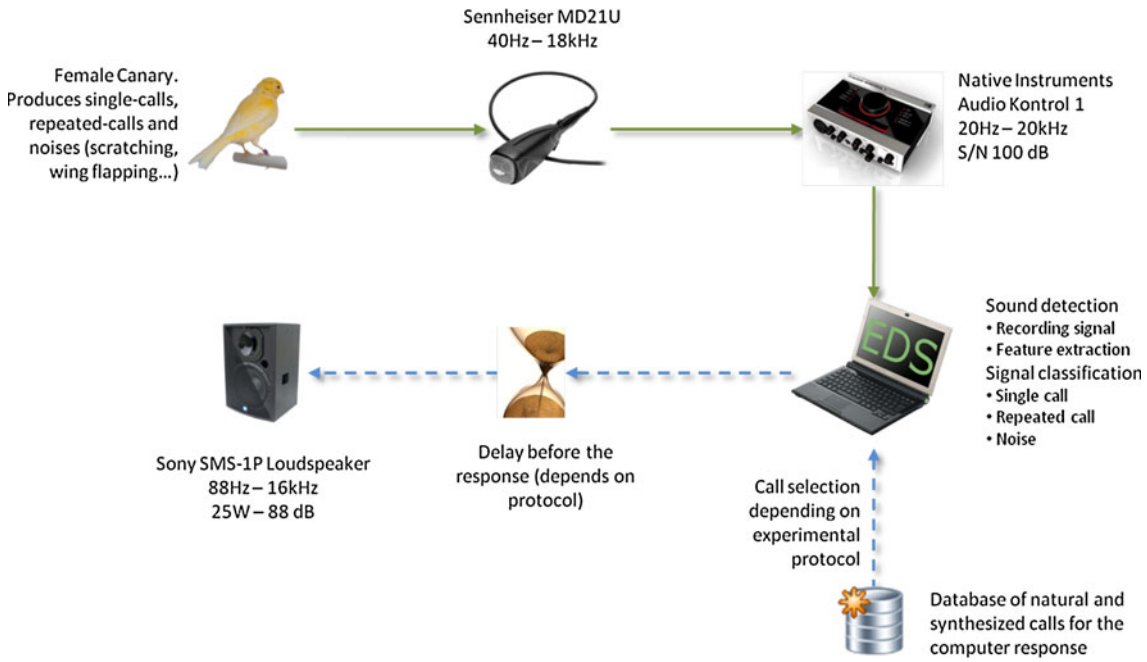


Fig. 2 Scheme of the closed-loop interaction between a canary and the computer system. *Solid green arrows* indicate the flow of acoustic signals produced by the bird. *Dashed blue arrows* indicate the signal produced by the system as a response to the bird's vocalization

Author's personal copy

animal vocalization classification problems (Molnár et al. 2008; Giret et al. in press). Consequently, we used a feature generation system to generate features adapted to the canary vocalization classification task. We trained the feature generation system using 1,800 labeled segments recorded in the experimental conditions described in this article. These 1,800 training segments were divided into 600 single calls, 600 repeated calls, and 600 noises. We used proprietary feature generation and feature selection algorithms (Pachet and Roy 2009) to create a feature set with the six features indicated below.

The feature generation algorithm (Pachet and Roy 2009) was used to generate a feature space of about 60,000 signal-processing features. A feature selection algorithm was then used to select a high-quality feature set that contained the six features below. The feature selection algorithm is based on feature analytical syntax and individual feature discrimination performance. It is described in detail in (Pachet and Roy 2009).

1. Bandwidth (Autocorrelation (Mean (SplitOverlap (x, 512, 0.5))), 50)

The signal is split into 50%-overlapping frames of 512 samples (i.e., 11.6 ms, as we consider 44,100-Hz signals). The frames are averaged into a single 512-sample frame. The autocorrelation of this 512-sample signal is computed, and its bandwidth is returned, i.e., the width of the highest peak at a threshold of 50% of the peak's value.

2. Median (SpectralRolloff (SplitOverlap (x, 512, 0.6)))

The signal is split into 60%-overlapping frames of 512 samples. The spectral roll-off is the frequency under which most of the energy is found in the spectrum (a rough approximation of the cut frequency between harmonic signal and noise), for each frame. The feature returns the median value over all the frames.

3. Norm (Low Short-Time Energy Ratio (LSTER) (SplitOverlap (Derivative (x), 512, 0.8), 0.15, 256, 0.5))

The signal is differentiated and split into 80%-overlapping 512-sample frames. For each frame, the LSTER is computed (Lu et al. 2002) as the ratio of 50%-overlapping 256-sample frames whose short-time energy is less than half the average. The feature returns the norm of the vector of LSTER values across the 512-sample frames.

4. Norm (Integration (Hamming (HarmonicSpectralSpread (SplitOverlap (x, 512, 0.7)))))

The signal is split into 70%-overlapping 512-sample frames. The harmonic spectral spread is computed for each frame, i.e., the amplitude (linear scale)-weighted standard deviation of the harmonic peaks of the spectrum, normalized by the instantaneous harmonic spectral centroid. A Hamming window operator is applied to the vector of harmonic spectral spread values. The resulting vector is integrated, and the feature returns its norm.

5. Norm (Integration (Pitch (SplitOverlap (x, 441, 0.8))))

The signal is split into 80%-overlapping frames of 441 samples each. The pitch of each frame is computed. The vector of pitches is integrated, and the feature returns the norm of the resulting vector.

6. Bandwidth (FFT (Triangle (FFT (Norm (SplitOverlap (x, 3307, 0.8)))))), 50)

This feature is particularly hard to interpret. FFT denotes the fast Fourier transform, and Triangle denotes a triangle filter. Other operators have been described above.

More details on the syntax and semantics of the features can be found in Pachet and Roy (2009).

Note that the six features above were not devised by human experts, but rather generated automatically by a feature generation algorithm. Their expression is therefore not straightforward and some of them may be difficult to interpret (e.g., feature 6).

A decision tree classifier was trained using the values of the six aforementioned features on the 1,800 samples. The algorithm is J48 in its Weka implementation (Witten and Eibe 2005). This classifier is used to classify the segments coming from M2 in real time.

The performance of the classifier was assessed on 1,800 labeled segments (different from the 1,800 training segments). The accuracy, averaged for the three classes "single call," "repeated call," and "noise," was 94.7%. A better performance assessment is provided by the F-measure, which avoids the inherent problems of precision or recall. For each class, the recall R is the proportion of positive examples that are correctly predicted, and the precision P is the proportion of the predicted positive examples that are correct. Where, for a given class, the proportion of positive examples is high compared with that of negative examples, the precision is usually artificially very high, and the recall very low, regardless of the actual quality of the classifier. The F-measure addresses this issue and is defined, for each class, by

$$F = {}^{2RP}/_{(R+P)}.$$

The average F-measure on the three classes was 0.947, with a minimum F-measure of 0.924 for the class "single call" (see Table 1 for more details on classifier performance).

The response module M4 is a program whose input is the last classified segment detected by the classification module. The module computes an acoustic response to the bird that depends on: (1) the last classified segment detected by the classification module, and (2) the predefined experimental protocol. The response can be any

Author's personal copy

Table 1 The performance of the classification module

TP rate	FP rate	Precision	Recall	*F*-measure	ROC area	Class
0.915	0.033	0.934	0.915	0.924	0.974	Single call
0.977	0.017	0.967	0.977	0.972	0.988	Noise
0.95	0.03	0.941	0.95	0.945	0.986	Repeated call
0.947	0.026	0.947	0.947	0.947	0.983	Weighted Avg.

Various statistics on top, and the confusion matrix below

A	B	C	←Classified as
549	15	36	A = single call
14	586	0	B = noise
25	5	570	C = repeated call

ROC receiver operating characteristic, *TP* true positive, *FP* falsepositive

acoustic signal. For the experiments described in this article (see "Experiment 1"), the responses are either a prerecorded canary vocalization from another canary, a synthesized canary vocalization or a silent period.

Subjects and housings

We used 15 unmated females of domestic canaries in this experiment, taken from our laboratory (LECC—University of Paris Ouest Nanterre-La Défense). They had been reared in aviaries ($118 \times 50 \times 50$ cm^3, six to seven canaries per aviary) in a room with many other birds, so these artificial conditions probably had little impact on their social behavior. Before the test period, birds had been kept in a short-day photoperiod (8 h light: 16 h dark). In this photoperiod, domestic canaries do not reproduce.

The day before the test, we changed the birds' conditions. Firstly, we put them in individual metal cages ($38 \times 33 \times 26$ cm^3), provided with two perches, a nest, and material to arrange it. The cages were arranged in soundproof chambers (two-floored boxes; $60 \times 55 \times 90$ cm^3 inside; $90 \times 80 \times 110$ cm^3 outside). We also changed the day length into a long-day photoperiod (16 h light:8 h dark), to simulate the start of spring and make them enter a reproductive state, in which they produce more calls (Nagle et al. 2002). Despite this change, no eggs were laid during the 2 weeks of the test. Moreover, no effect of the date was observed in the vocalization rate of the birds. No males were present in the room, and females were not in competition with others to mate. Birds were fed with seeds, fresh food (apple), and a special mixture for canaries. Fresh water was provided every day.

Experiment 1

We tested the birds individually in a dedicated sound-attenuation room. Each bird was placed in a plastic cage (less noisy than metal cages). Two perches only were inside the cage in order to minimize noise produced when birds made contact with the perches. A microphone (Sennheiser MD21U) linked the cage to an external sound card (Native Instruments Audio Kontrol 1) that conducted the signal to a VAIO Sony (PCG-6G1M) computer. In these conditions, only two types of calls were produced by the canaries: single calls and repeated calls. The computer recorded the whole test and then classified the vocalizations into one of these classes (Fig. 2).

The microphone used had a frequency range of 40–18,000 Hz and sensitivity of 1.8 mV/Pa ± 2.5 dB. The sound card had a frequency range of 20–20,000 Hz and a signal-to-noise (S/N) ratio of 100 dB. The loudspeaker was a single Sony SMS-1P with nominal power of 25 W, sensitivity of 88 dB, and frequency range of 88–16,000 Hz.

We simulated an acoustic contact with an unknown female in a new environment. Each test was composed of three periods. The total duration of the whole experiment was 20 min. The first period (habituation) was a 10-min silence, for habituation of the bird to the conditions. In the second period (stimulation), the computer emitted single calls of an unrelated conspecific female at constant time intervals until the first bird response (single call or repeated call). We played single calls, because two birds initiating an interaction use this type of call (Mulligan and Olsen 1969; L.N. and A.L. pers. obs.). In the third period (interaction), five types of interactions were designed: (1) positive interaction (PI): the computer responded with a single call when the bird produced a single call and with a repeated call when the bird produced a repeated call; (2) negative interaction (NI): the computer responded with a single call when the bird produced a repeated call, and with a repeated call when the bird produced a single call; (3) silence (S): the computer did not respond after the bird's first vocalization; (4) positive synthetic interaction (PSI): the computer responded with a single-call-like sound artificially synthesized with the computer program AVISOFT-SAS Lab Pro (R. Specht version 4.38, 2005) when the bird produced a single call, and with an artificially synthesized repeated-call-like sound when the bird produced a repeated call; (5) negative synthetic interaction (NSI): as for PSI, but with the interaction types reversed.

Author's personal copy

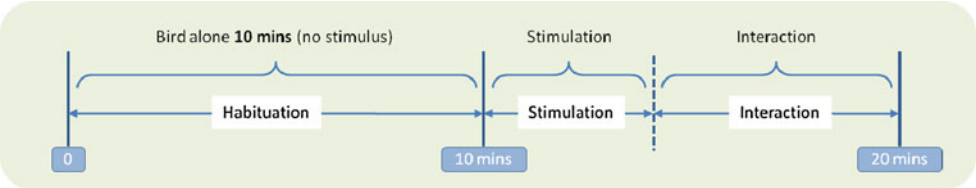


Fig. 3 The three-phase experimental protocol: (1) habituation: the bird is in the cage for 10 min and the computer does not produce or record any sound; (2) stimulation: the computer plays single calls every 4 s to stimulate the canary; the stimulation ends when the canary produces a call; (3) interaction: the computer produces a response to every call; the response depends on the type of interaction (positive, negative, etc.)

The stimulation period and the interaction period lasted 10 min together. As the length of the stimulation period varied, the length of the interaction period was also not constant (Fig. 3). The computer responses were played by a loudspeaker (Elipson Minihoms Sat) placed near the cage via an external sound card.

To avoid any problem with pseudoreplication, we used the repertoire of a different bird for each experiment, chosen at random. Moreover, each repertoire consisted of ten different single calls and ten different repeated calls of a same bird that were randomly played in the second and in the third period. All played calls (single and repeated) came from females recently moved into a long-day photoperiod.

Each bird was tested once a day for a period of 10 days. The order of playback was randomly assigned to each bird. Their responses were thus recorded twice for each interaction type.

Experiment 2

We designed a second experiment to compare our interaction method with traditional playback. The frequency of the vocal production of ten female domestic canaries was tested in both an interaction simulation and in a playback experiment.

The birds were put in the same conditions (cages, photoperiod, and soundproof chambers) as in the first experiment. The birds were divided into two groups of equal size, A and B. On day 1, the birds in group A experienced the interactions PI (see experiment 1) and those in group B experienced a traditional playback context.

In the playback test, canary single calls (the same as those used for the PI experiment) were played at a regular frequency of 3.4 calls per minutes (the average number of single calls produced by canaries in experiment 1). The use of a "natural pattern" seems difficult because of the high variability of bird vocal production, both in isolation and in PI.

On day 2, the birds in group A underwent the playback experiment, and the birds in group B the positive interaction test. Thus each bird was its own control.

Analysis of the results

For the analysis, we counted the number of responses (single calls and repeated calls) for each interaction and in the playback test. The classification module could have done the counting automatically, but only with some imprecision due to the error rate of the classifier (see "Software tools"). To avoid this imprecision, we undertook a manual count. In experiment 1, we first compared the total number of single calls versus repeated calls for each bird, using a Mann–Whitney sum-rank test. We then tested the single call rate and the repeated call rate for significance with a Friedman test, using Sigmastat V3.1 (2004). We also compared the total number of vocalizations between the different days of the test period, using a Friedman test, followed by Student-Newman-Keuls posthoc analysis.

In experiment 2, we tested the number of vocalizations (single and repeated calls) in both the interaction PI and the playback contexts. Nevertheless, three of the ten birds did not engage in the interaction (i.e., they did not produce any vocalization) and were excluded from the analysis. The number of responses in each condition was compared using a Wilcoxon signed-rank test.

Results

Experiment 1

During the entirety of the interactions, females produced 25.94 ± 5.67 [mean ± standard error (SE)] single calls per minute and 13.39 ± 4.58 (mean ± SE) repeated calls. The number of single calls was significantly higher than the number of repeated calls in each type of interaction (Mann–Whitney rank-sum test, $T = 287$, $n = 15$, $p = 0.025$; Fig. 4).

The distribution of single calls between the different types of interaction ($\chi^2 = 13.88$, $df = 4$, $p = 0.008$, $n = 15$; Fig. 5a) was significantly heterogeneous. Post hoc analysis using the Student–Newman–Keuls (SNK) method revealed that the number of single calls was significantly

Author's personal copy

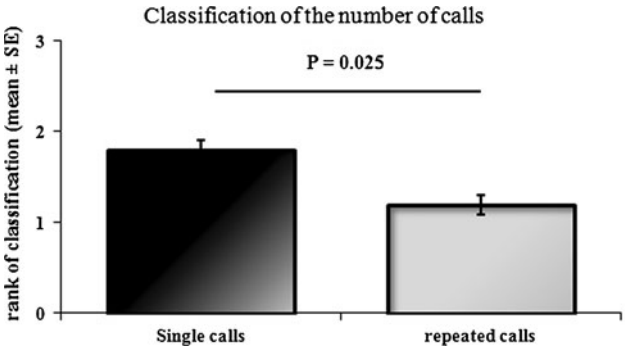


Fig. 4 Mean rank of classification of the number of each type of call produced by the females. *Different color* indicates significant difference

more often higher for the positive interaction and the silence than for the other types of interaction. The mean ± SE number of single calls per minute was 6.73 ± 1.91 for the PI, 5.99 ± 2.53 for the NI, 5.62 ± 1.10 for the S, 3.95 ± 1.65 for the PSI, and 3.65 ± 1.35 for the NSI interaction. On the other hand, we did not find any significant effect of the interactions for the number of repeated calls ($\chi^2 = 6.17$, $df = 4$, $p = 0.187$, $n = 15$; Fig. 5b). The mean ± SE number of repeated calls was 3.73 ± 1.42 for the PI, 1.91 ± 0.94 for the NI, 3.27 ± 1.13 for the S, 2.49 ± 1.07 for the PSI, and 1.99 ± 1.14 for the NSI interaction.

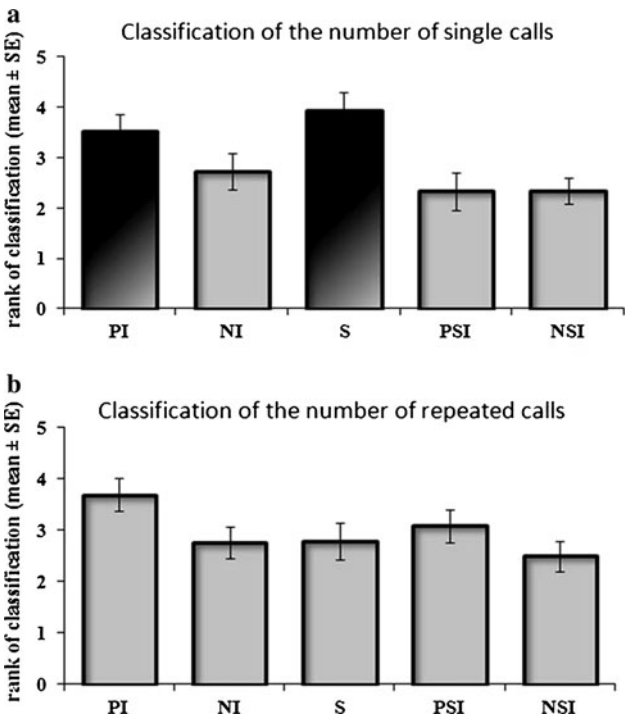


Fig. 5 Mean ± SE rank of classification of the number of calls produced by the females for each type of interaction. *Different color* indicates significant difference. *PI* positive interaction; *NI* negative interaction; *S* silence; *PSI* synthetic positive interaction with synthetic sounds; *NSI* synthetic negative interaction with synthetic sounds. **a** single calls, **b** repeated calls

Finally, for single and repeated calls, we found that the total number of vocalizations was not randomly distributed among the days ($\chi^2 = 16.99$, $df = 9$, $p = 0.049$, $n = 15$ followed by SNK: p = NS, $n = 15$ for the single calls and $\chi^2 = 18.36$, $df = 9$, $p = 0.03$, $n = 15$ followed by SNK: p = NS, $n = 15$ for the repeated calls). Nevertheless, in both cases, post hoc analyses revealed that no significant difference existed between the days. We concluded that the day order had little impact on the number of produced calls.

Experiment 2

In this experiment, birds produced on average 8.9 vocalizations (±3.74) per minute when exposed to the PI procedure and 3.13 vocalizations (±1.89) when exposed to the playback procedure. The Wilcoxon signed-rank test revealed that the birds produced significantly more calls in the PI than in the playback procedure ($W = -28$, $p = 0.016$, $n = 7$). Nevertheless, we did not find significant differences in the number of single ($W = -15$, $p = 0.063$, $n = 7$) or repeated calls ($W = -3$, $p = 0.813$, $n = 7$) between the two situations, probably due to weak statistical power (0.30 in both cases).

Discussion

This article presents a new approach to studying bird vocal communication. Our proposal is to put the bird in a closed-loop interaction with a computer (artificial interaction). The fundamental novelty of this research is that the system, consisting of a recording device, a computer, audio analysis software, and a playback device, is an active component of the experiment. In earlier attempts at using automatic audio analysis techniques, the analysis was only used *after* the experiment itself to process the signals recorded to help the researchers interpret the results (Derégnaucourt et al. 2005; Giret et al. in press; Tchernichovski et al. 2004).

In the implementation of the framework described herein, the bird is recorded and the computer uses signal-processing techniques to analyze its vocalizations in real time to produce an adapted response. Several experimental protocols can be implemented in this framework, depending on various parameters, e.g., the type of response the computer plays to the bird, the delays before the response, and the type of signal played by the computer. This framework is a generalization of traditional playback techniques and allows the implementation of complex experiments that are not conceivable otherwise, although

Author's personal copy

unfortunately some visual aspects of communication such as wing-flapping (West and King 1988) cannot be taken into account.

Comparison of several interaction scenarios (e.g., positive and negative interaction, with real or synthesized calls) has shown that the birds react differently to different situations (Fig. 5). More particularly, comparison of the frequency of single call production in (1) a positive interaction context and (2) a regular stimulation context proved that the birds are actually engaged in the interaction (see experiment 2).

These results are of great importance, as they legitimate artificial interaction as a tool to study bird vocal communication and pave the way for fully automated experiments that could not be implemented before. The progress in sound analysis techniques allows the design of complex experiments in which the computer interacts with several birds to study their social behavior. This requires individual bird identification techniques, which we are currently developing. Early results are encouraging, with 80% accuracy on the 15 birds used in this study.

The results presented herein (experiment 1) show that female single call rates are heterogeneously distributed among the different types of interaction. The single call rate was significantly higher for females in positive interaction (PI) or silence (Sil) than in negative interaction (NI), or positive (PSI) or negative synthetic interaction (NSI). It is not surprising that isolated birds (Sil) have a high rate of contact calls, as the canary is a social species, living in flocks (Newton 1972). In such species, individuals out of earshot produce louder and more calls (Marler 2004), since isolated individuals try to communicate with others with single calls. Nevertheless, it must be emphasized that this situation is abnormal and probably occurs exceptionally in the wild. For this reason, it must be taken into account that our method is not designed to compete with field observations of real bird flocks, which technology cannot easily replace.

Moreover, because the number of single calls was higher in each situation than the number of repeated calls, the computer produced more single calls in PI and more repeated calls in NI.

These two observations are consistent with the hypothesis of Mulligan and Oslen (1969) suggesting that single calls are contact calls.

On the other hand, Nagle et al. (2002) revealed that female single-call production is influenced by male songs and is therefore an aspect of female sexual behavior. Our experiments show that playback of calls produced by females influenced female single-call production. This observation suggests that the function of single calls is not limited to sexual responses as established by Nagle et al. (2002) but is also involved in maintenance of flock cohesion (Okanoya and Dooling 1991).

Both types of synthetic calls (PSI and NSI) and negative interaction (NI) induced a low rate of bird single calls. This result shows that, despite the similarity of these artificial calls to natural ones, females are able to discriminate between them. Unknown noises may be stressful for the birds and thus inhibit them in PSI and NSI. Similarly, the birds remain silent when hearing a majority of repeated calls. According to Mulligan and Olsen (1969), repeated calls reflect bird excitement; it is probable that they serve as a signal of danger.

In conclusion, birds producing a high rate of contact calls are more likely to initiate social interaction. The fact that the playback procedure produced a lower rate of bird single calls showed that interaction is a key aspect of bird vocal behavior. Birds are stimulated by the responses of the computer more than by regular playback. This observation is a strong indicator that single calls relate to first-order intentionality of the bird, as defined by Emery (2005). On the other hand, production of repeated calls may be perceived as a signal of danger from the receiver's point of view.

Our study shows that computer-based closed-loop interaction with birds is possible and effective. This method allows the creation and study of artificial communication situations. We are currently designing new experiments using this framework to study social communication involving more than two birds. To do so, the framework is being extended with additional features such as automatic individual identification.

Acknowledgments A.L. is funded by the Institut Emilie Du Châtelet, which supports research on women, sex, and gender. We want to thank N. Giret and Z. Skandrani for participation in the early stages of this work, and C. Desaleux and P. Groué for taking care of the birds. We also thank the referees for their helpful comments and advice, and Katy Masuga for her help with the English translation. All stages of this experiment were carried out under agreement number 92-244 (L.N.), delivered by the French Ministry for Agriculture and Fisheries, and comply with the current laws of France. The authors declare that they have no conflicts of interest.

References

Aldridge HDJN, Rautenbach IL (1987) Morphology, echolocation and resource partitioning in insectivorous bats. J Anim Ecol 56:763–778

Amrhein V, Erne N (2006) Dawn singing reflects past territorial challenges in the winter wren. Anim Behav 71:1075–1080

Amy M, Monbureau M, Durand C, Gomez D, Théry M, Leboucher G (2008) Female canary mate preferences: differential use of information from two types of male-male interaction. Anim Behav 76:971–982

Author's personal copy

Beecher MD, Beecher IM, Hahn S (1981) Parent-offspring recognition in bank swallows (*Riparia riparia*): II. Development and acoustic basis. Anim Behav 29:95–101

Bradbury JW (2003) Vocal communication in wild parrots. In: De Waal FBM, Tyack PL (eds) Animal society complexity: intelligenceculture and individualized societies. Harvard University Press, Cambridge, pp 293–316

Brumm H, Todt D (2004) Male-male vocal interactions and the adjustment of song amplitude in a territorial bird. Anim Behav 67:281–286

Castellano S (2009) Unreliable preferences, reliable choice and sexual selection in leks. Anim Behav 77:225–232

Catchpole CK, Slater PJB (1995) Bird song, biological themes and variations. Cambridge University Press, Cambridge

Cheney DL, Seyfarth RM (1990) How monkeys see the world—inside the mind of another species Behaviour. University of Chicago Press, Chicago

Clutton-Brock TH, Albon SD (1979) The roaring of red deer and the evolution of honest advertising. Behaviour 69:145–170

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. Nature 433:710–716

Draganoiu TI, Nagle L, Musseau R, Kreutzer M (2006) In a songbird, the black redstart, parent use acoustic cues to discriminate between their different fledglings. Anim Behav 71:1039–1046

Emery NJ (2005) The evolution of social cognition. In: Easton A, Emery NJ (eds) The cognitive neuroscience of social behaviour. Psychology, New York, pp 7.1–7.98

Evans CS, Evans L, Marler P (1993) On the meaning of alarm calls: functional reference in an avian vocal system. Anim Behav 46:23–38

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans: common problems and diverse solutions. University of Chicago Press, Chicago

Giret N, Roy P, Albert A, Pachet F, Kreutzer M, Bovet D (in press) Finding acoustic features for parrots vocalizations: the features generation approach. JASA

Grodzinski U, Lotem A (2007) The adaptative value of parental responsiveness to nestling begging. Proc R Soc B 274:2449–2456

Hauser MD (1997) The evolution of communication. MIT Press, Bradford

Johnstone RA (1995) Sexual selection, honest advertisement and the handicap principle: reviewing the evidence. Biol Rev 70:1–65

Klump GM, Shalter MD (1984) Acoustic behaviour of birds and mammals in the predator context. I. Factors affecting the structure of alarm signals. II. The functional significance and evolution of alarm signals. Zeit Tierpsychol 66:189–226

Kroodsma DE (2004) Vocal behavior. In: Podulka S, Rohrbaugh RWJ, Bonney R (eds) Handbook of bird biology. Princeton University Press, New York, pp 7.1–7.98

Lu L, Zhang H-J, Jiang H (2002) Content analysis for audio classification and segmentation. IEEE Trans Speech Audio Process 10:504–516

Marler P (2004) Bird calls: a cornucopia for communication. In: Marler P, Slabbekoorn H (eds) Nature's music. Elsiever Academic, San Diego, pp 132–177

Marler P, Dufty A, Pickert R (1986) Vocal communication in the domestic chicken. II. Is a sender sensitive to the presence and nature of a receiver? Anim Behav 34:194–198

Maynards Smith J, Harper D (2003) Animal signals. Oxford University Press, Oxford

McGregor PK, Dabelsteen T (1996) Communication networks. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Comstock, New York, pp 409–425

Mennill DJ, Ratcliffe LM (2004) Overlapping and matching in the song contests of black-capped chickadees. Anim Behav 67:441–450

Mennill DJ, Boag PT, Ratcliffe LM (2003) The reproductive choice of eavesdropping female black-capped chickadees, *Poecile atricapillus*. Naturwissenschaften 90:577–582

Mitchell Masters W (1979) Insect disturbance stridulation: its defensive role. Behav Ecol Soc 5:187–200

Molnár C, Kaplan F, Roy P, Pachet F, Pongrácz P, Dóka A, Miklósi Á (2008) Classification of dog barks: a machine learning approach. Anim Cogn 11:389–400

Mulligan JA, Olsen KC (1969) Communication in canary courtship calls. In: Hinde RA (ed) Bird vocalisation. Cambridge University Press, Cambridge, pp 165–184

Nagle L, Kreutzer M, Vallet E (2002) Adult female canaries respond to male song by calling. Ethology 108:463–472

Naguib M (2005) Singing interactions in songbirds: implications for social relations and territorial settlement. In: McGregor PK (ed) Animal communication networks. Cambridge University Press, Cambridge, pp 300–319

Newton I (1972) Finches. Collins, London

Okanoya K, Dooling RJ (1991) Detection of species-specific calls in noise by zebra finches *Poephila guttata* and budgerigars *Melopsittacus undulatus*: time or frequency domain? Bioacoustic 3:163–172

Pachet F, Roy P (2009) Analytical features: a knowledge-based approach to audio feature generation. EURASIP J Audio Speech Music Process 1:1–23

Smith WJ (1977) The behaviour of communication—an ethological approach. Harvard University Press, Cambridge

Sprau P, Schmidt R, Roth T, Amrhein V, Naguib M (2010) Effects of rapid broadband trills on responses to song overlapping in nightingales. Ethology 116:300–308

Struhsaker TT (1967) Auditory communication among Vervet monkeys (*Cercopithecus aethiops*). In: Altmann SA (ed) Social communication among primates. University of Chicago Press, Chicago

Tchernichovski O, Lints TJ, Derégnaucourt S, Cimenser A, Mitra PP (2004) Studying the song development process, rationale and methods. Ann NY Acad Sci 1016:348–363

Vallet E, Kreutzer M (1995) Female canaries are sexually responsive to special song phrases. Anim Behav 49:1603–1610

Vallet E, Beme I, Kreutzer M (1998) Two-note syllables in canary songs elicit high levels of sexual display. Anim Behav 55:291–297

Vignal C, Mathevon N, Mottin S (2004) Audience drives male songbird response to partner's choice. Nature 430:448–451

West MJ, King AP (1988) Female visual displays affect the development of male song in the cowbird. Nature 334:244–246

Witten IH, Eibe F (2005) Data mining: practical machine learning tools and techniques, 2nd edn. Morgan Kaufmann, San Francisco

Wyndham E (1980) Diurnal cycle, behaviour, and social organization of the budgerigar (*Melopsittacus undulatus*). Emu 80:25–33