Behavioural Processes 163 (2019) 113–120

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Stress-induced flexibility and individuality in female and male zebra finch distance calls

H.A. Soula[a,c,*], D. Carnino[c], M.S.A. Fernandez[c,d], E.C. Perez[d], A.S. Villain[d], C. Vignal[b]

[a] Sorbonne Université, INSERM CRC UMRS 1138, F-75006, Paris, France
[b] Sorbonne Université CNRS, iEESUMR 7618, F-75005, Paris, France
[c] INRIA EPI Beagle, F69621, Villeurbanne, France
[d] NeuroPSI/ENES, Univ.Lyon/UJMUMR-CNRS 9197, F42023, Saint-Etienne, France

ARTICLE INFO

Keywords:
Zebra finch
Machine learning
Stress
Individual signature

ABSTRACT

Vocal recognition is central to the coordination and organization of behavior in pair-bonding species such as zebra finches. Zebra finches' vocalizations are individualized and support acoustic discrimination processes. Physiological states - such as the ones involved in emotional stress - can modify vocal production and consequently the structure of vocalizations. These modifications might signal the state of the caller but also impair individual recognition processes. This may represent a signaling trade-off, especially in contexts where both pieces of information can be critically important, for example when mates use calls to reunite after social isolation. Here we study the impact of a stress on the individual vocal signature in both female and male zebra finch distance calls.

We built a manually curated database of distance calls of several individuals (both females and males) recorded in control and stress conditions. The stress was induced either by social isolation of the bird or using exogenous corticosterone. We developed a machine learning approach to assess the impact of stress on the individual characterization of calls.

We show that while calls' spectral structure is significantly modified by stress, it still allows for the correct classification of calls to the caller. Moreover, we also show that the stress-induced modification of calls' structure is not a 'general feature signal' that can be detected as a 'stress' signal regardless of identity. Thus, female and male zebra finch calls' structure show stress-induced flexibility that stays within the range of individual vocal signatures.

1. Introduction

Vocal recognition or discrimination is central to the coordination and organization of behavior in many social species. The acoustic structure of vocalizations may inform conspecifics on a caller's species, sex, group membership, relatedness or individual identity (Bradbury and Vehrencamp, 2011; Marler and Slabbekoorn, 2004). To allow efficient discrimination, vocalizations must bear stable acoustic signatures (Beecher, 1989). For instance, individualized features of vocalizations make individual vocal recognition possible (e.g. Aubin and Jouventin, 1998, Bee and Gerhardt, 2002; Charrier et al., 2001; Janik et al., 2006; Jouventin and Aubin, 2002; Mathevon et al., 2008).

Beside stable information on the caller, vocalizations may also bear labile and transient pieces of information, like context, motivation or emotion (Morton, 1977, Manser, 2001; Briefer, 2012). Such temporary information can be detected in modifications of the temporal pattern of calling behavior: as e.g. in birds (Kilner and Johnstone, 1997) and frogs (Tobias et al., 2004). Instantaneous information can also be expressed in changes in the vocalization units using for instance call type variations like in the great gerbil (Randall and Rogovin, 2002) or acoustic structure variations as in white-browed scrubwren (Leavesley and Magrath, 2005).

Both types of information, stable and transient, may be encoded in different categories of vocalizations: besides their species-specific and individualized songs, many small passerine species share the 'seet hawk alarm call' in response to a bird of prey in flight, a high-pitched and narrow band pure tone (Marler, 2004). But the two categories of information, stable and transient, can also be encoded in the same vocalizations see for examples African elephants (Soltis et al., 2005), domestic dogs (Yin and McCowan, 2004) or meerkats (Schibler and Manser, 2007). If one vocalization transmits information pertaining to multiple pieces of information, those can still be encoded in different

* Corresponding author.
E-mail address: hedi.soula@upmc.fr (H.A. Soula).

https://doi.org/10.1016/j.beproc.2018.03.018
Received 20 April 2017; Received in revised form 23 January 2018; Accepted 17 March 2018
Available online 23 March 2018
0376-6357/

parameters (segregation hypothesis, Marler). For example, in mongoose the segregation is temporal and within the syllable (Jansen et al., 2012) whereas in baboons it is both in tempo and spectral structure (Rendall, 2003).

Stable and transient pieces of information can also be coded in the same sets of parameters of the vocalization, and in that case there might be a trade-off in the signal's reliability, with one component more accurately communicated than the other (trade-off hypothesis, Briefer et al., 2010). In fallow deer (*Dama dama*), some acoustic features of the groans code for both individual identity and phenotypic quality (age and dominance status), so vocal cues coding for individuality change over years and information about quality is more reliable than identity (Briefer et al., 2010).

This signaling trade-off might be particularly stringent when both pieces of information can be critically important to the receiver. In zebra finches (*Taeniopygia guttata*), birds from life-long pair-bonds and mated partners have highly synchronized activities (Zann, 1996; Mariette and Griffith, 2012). When mates lose visual contact or reunite after a separation, they exchange distance calls, which are highly individualized harmonic stacks (Elie and Theunissen, 2016; Forstmeier et al., 2009; Mouterde et al., 2015) and allow mate discrimination (Vignal et al., 2004, 2008; Hernandez et al., 2016). These calls can also convey information on the physiological stress of the caller: calls produced in social isolation (a stressful situation of negative valence in this species (Remage-Healey et al., 2003; Banerjee and Adkins-Regan, 2011)) or after experimental corticosterone elevation show specific structural changes (Perez et al., 2012). Moreover, these stress-induced calls promote emotional contagion in mates, as females show increased corticosterone levels after hearing the stressed calls of their mate, and not after hearing the stressed calls of another male (Perez et al., 2015). In an operant discrimination task, females grouped the stressed calls and the control calls of the same male in the same category (Perez et al., 2015).

Zebra finch distance calls seem to reliably convey both individual identity and stress, defined here as an emotion of high arousal and negative valence (Briefer, 2012). But how do these two pieces of information impact each other? Does stress modify all individuals' calls the same way, making it a universal superimposed feature allowing any listener to detect stress in the calls of any caller (universality hypothesis)? Or do stress modifications of call structure stay within the caller's acoustic features space, so it is possible only for a familiar listener to detect stress in the vocalizations of a caller (familiarity hypothesis)?

To test these hypotheses, we built a manually curated database of distance calls of female and male zebra finches recorded in control and stress conditions. The stress was induced either by social isolation or by experimental increase of corticosterone (oral ingestion). We developed a machine learning approach using random forests to assess the impact of a stress on the individual vocal signature in both female and male calls.

2. Material and methods

2.1. Dataset

We used recordings of zebra finch distance calls from 4 different experiments (described in Table 1). For the purpose of this article all vocalizations have been pooled and no distinction of the origin of the data has been made. However, all recordings were made using bird from the same aviary - similar housing conditions, daylight cycles and food type and quantities - in addition to similar recording rooms and apparatus. The total number of calls was 1780 control calls (53%) 1548 stress calls (44%) spanning 57 different individuals (see Table 1 for summary).

Most control calls were elicited using a playback of opposite sex calls (female calls for male subjects, male calls for female subjects) and in the presence of an audience of 2 other birds (dataset 1, 2 & 4, see Perez et al. (2012) for details on protocol) with exception for dataset 3 where the stimulus was a physical and visual separation but within hearing distance from partner (see Perez et al. (2015) for details on protocol). The stress condition was either obtained via social isolation (dataset 1, 2 & 4) or after exogenous corticosterone (CORT) ingestion (dataset 3) (see Perez et al. (2012) for details on protocol).

2.2. Data pre-processing - duration extraction

Calls were separated wave files (16bit 44.1 kHz sampling monochannel) that were processed automatically using custom software written in Python. The first step of pre-processing was to compute the amplitude using the spectrograms, which were obtained using an FFT with frequency band of 125 Hz using a Gaussian Filtering. The temporal sampling frequency was set at 1 kHz. Within each wave file, the maximum amplitude time was extracted and a call was defined as the interval around the maximum amplitude time such that the amplitude within this interval remained 10% of the maximum amplitude. This processing allows us to extract a call reliably within a recording by obtaining the same amplitude variation. This also yields a reproducible call length independent of the recording levels defining the duration of the calls.

2.3. Spectrogram matrix

For this analysis, the spectrogram of a call will be viewed as a matrix of acoustic parameters. In order to compare those matrices between calls, we needed to standardize their sizes across the entire dataset. To do so, we first find the longest call of the dataset, which then sets the duration reference. We compute the spectrogram of each call for the same duration as this reference duration by padding with zeros (both in time and frequency) to obtain matrices with identical dimension for every calls (shorted calls were centered). Finally, we selected a frequency band - between 200 Hz and 8 kHz as it contained most of the spectral information for calls and ignored spectral information beyond that interval. In this dataset, the longest call was 888 ms long and the frequency band had 59 points. So each call of the dataset has been described as a series of 59*888 values (parameters) corresponding to its full spectrogram. The so-called spectrogram matrix data which will contain the spectrogram values of all the calls of the dataset is a matrix of 3328 lines (the number of calls) and 52392 columns (59 * 888 the spec). We reduced the dimensionality of this matrix data by using a principal component analysis (PCA). We kept the 20 first components to obtain a SPEC matrix data of 3328 lines and 20 columns (similar to the technique used in Mouterde et al., 2015).

Table 1
Dataset summary.

Set	Sex	# Ind	Stimulus	Control	# Control calls	Stress	# Stress calls
1	F	25	Playback	Audience	528	Isolation	490
2	M	15	Playback	Audience	423	Isolation or Audience + Exogenous CORT	456
3	M	11	Separation	Acoustic contact with partner	677	Acoustic contact with partner + Exogenous CORT	410
4	M	6	Playback	Audience	152	Isolation	192
Total		57			1780		1548

Table 2
Mean (in white) and standard deviation (grey) of acoustic parameters breakdown by male/female and control/stress.

Stress	Sex	Mean Spectrum	SD Spectrum	Skewness Spectrum	Kurtosis Spectrum	Entropy Spectrum	Q1	Q2	Q3
C	F	3759.49	714.31	0.205	5.34	0.675	3319.46	4110.39	4895.38
C	M	3753.91	796.41	0.189	5.34	0.726	3283.61	3687.03	4214.11
S	F	3575.50	697.57	0.238	6.01	0.675	3157.91	4396.82	5703.66
S	M	3826.79	749.71	0.082	5.07	0.714	3371.02	3795.26	4282.05
C	F	601.42	142.34	0.780	3.37	0.075	530.82	2357.76	3116.85
C	M	451.99	171.51	0.829	2.52	0.073	510.42	548.23	588.09
S	F	738.31	142.16	0.876	4.77	0.075	648.34	3303.88	4684.17
S	M	398.15	151.96	0.655	1.96	0.075	434.25	473.77	512.56

Stress	Sex	Mean Time	SDTime	Skew Time	Kurtosis Time	Entropy Time	Amplitude	Mean Saliency	Duration
C	F	0.098	0.051	0.143	2.159	0.972	0.146	1469.02	270.65
C	M	0.071	0.036	0.175	2.162	0.973	0.099	1496.78	134.16
S	F	0.117	0.061	0.141	2.111	0.976	0.159	1375.51	281.75
S	M	0.078	0.041	0.163	2.171	0.975	0.142	1439.25	142.87
C	F	0.048	0.024	0.221	0.191	0.010	0.075	297.29	80.77
C	M	0.031	0.015	0.262	0.203	0.010	0.070	238.42	40.71
S	F	0.082	0.040	0.186	0.163	0.009	0.078	278.93	92.92
S	M	0.032	0.016	0.219	0.168	0.009	0.078	267.98	44.87

2.4. Acoustic parameters matrix

In addition to the SPEC matrix, we selected the following 16 acoustic parameters as acoustic features of a call in a similar way to (Elie and Theunissen, 2016). Eight parameters were computed from the power spectral density (PSD): the average spectral density, its standard deviation, its skewness (order 3 centered moment), its kurtosis (order 4 centered moment), its entropy and the quartiles 1, 2 and 3 (Q1, Q2, Q3 respectively). Five parameters were calculated from the resampled waveform signal (amplitude against time) with a new sample rate of 1000 Hz: the average time, the temporal standard deviation, the skewness, the kurtosis and the entropy. We have added the maximal amplitude and the mean saliency. The saliency is a measure of the significance of the fundamental frequency, that is whether (high saliency) or not (low saliency) the fundamental is easily identified in the mass of frequencies. It can be formulated as the perceptual pitch strength. All these parameters are described in detail in (Elie and Theunissen, 2016) and mean values can be found in Table 2.

Finally, the duration computed from the calls pre-processing step was added to the matrix for a total of 16 parameters to obtain the ACOUS matrix data of 3328 lines and 16 columns. Acoustic parameters were chosen quite arbitrarily and full spectrum may contain noisy and redundant information. For sake of completeness, we added another matrix to the analysis by concatenating the first two to obtain the COMB (for combined) matrix data of 3328 by 36.

2.5. Machine learning analysis

In order to assess the impact of stress on individual discrimination we used machine learning techniques to learn how to discriminate between emotional states or between individuals. To obtain the maximum correct discrimination we used an extremely efficient classification algorithm: the random forest (Breiman, 2001). This provided a lower bound of discrimination capabilities. Random forests are supervised learning algorithm based on consensus of tree classification. Here we trained 100 forests — each one having 1000 trees —to learn a classification task using 75% of the data available (training set) randomly chosen for each forest. We computed the error of classification on the remaining 25% (validating set). The feature selection criterion is the variance reduction. The algorithm is provided in the random forest classifier in the sklearn library for python (Pedregosa et al., 2011). Unless otherwise mentioned, the standard errors were computed using all 100 learning errors. We stress here that each forest used a randomly chosen training and validating data set. We used the average values plus/minus the standard error for the figures. The high number of trials ensures that we obtained a good approximation of the mean. In all cases, mean comparison statistical tests were significant.

3. Results

3.1. Sex differences in calls and impact of stress

By inspecting the first two principal components density computed on the full centered spectrograms (see methods) yields a clear distinction between female and male (Fig. 1A and B) whereas differences are harder to make between control and stress.

3.2. Control/stress differences in calls

We first show that this is possible to distinguish between control and stress calls, no matter the caller's identity: the 'Status' feature taking only two values: Control or Stress. The classification error was computed on the validating data set and the results are displayed on Fig. 2A (for 100 learning trials). The first 3 bars show the results when all calls are used for training and validating (SPEC: 18.37 ± 0.13%/ACOUS: 17.17 ± 0.13%/COMB: 13.75 ± 0.13%). The next two sets of 3 bars are obtained by restricting dataset to male (SPEC: 14.4 ± 0.14/ ACOUS: 17.3 ± 0.15/COMB: 12.6 ± 0.15) and female calls (SPEC: 15.32 ± 0.25/ACOUS: 15.91 ± 0.20/COMB: 12.25 ± 0.24) respectively. Because the task is to classify between control and stress, a pure random selection should yield an error of 50%. The first result is that there is a detectable difference between control and stress calls that can be discriminated in our dataset.

In the first training session most forests did train on exemplars of calls of all individuals in both conditions. We did another experiment where we trained the forest on *75% of the individuals* with the validating dataset being the 25% remaining individuals. In other words, the forests did train on a subset of the population and try to predict on another subset of individuals on which they did not train. Results are displayed on Fig. 2B. As for the previous figure, the first 3 bars show the results when all calls are used for training and validating (SPEC: 42.1 ± 0.5%/ ACOUS: 47.40 ± 0.4% / COMB: 44.4 ± 0.5%). The next two sets of 3 bars are obtained by restricting dataset to male (M SPEC: 42.2 ± 0.6%/M ACOUS: 47.5 ± 0.5%/M COMB: 43 ± 0.5%) and female calls (F SPEC: 36.0 ± 0.9%/ F ACOUS: 45.0 ± 0.8%/ F COMB: 45.2 ± 1.0%) respectively. Note that the random selection has the same average error as above (50%). When learning the emotional state classification on a sub-population to predict the emotional state of

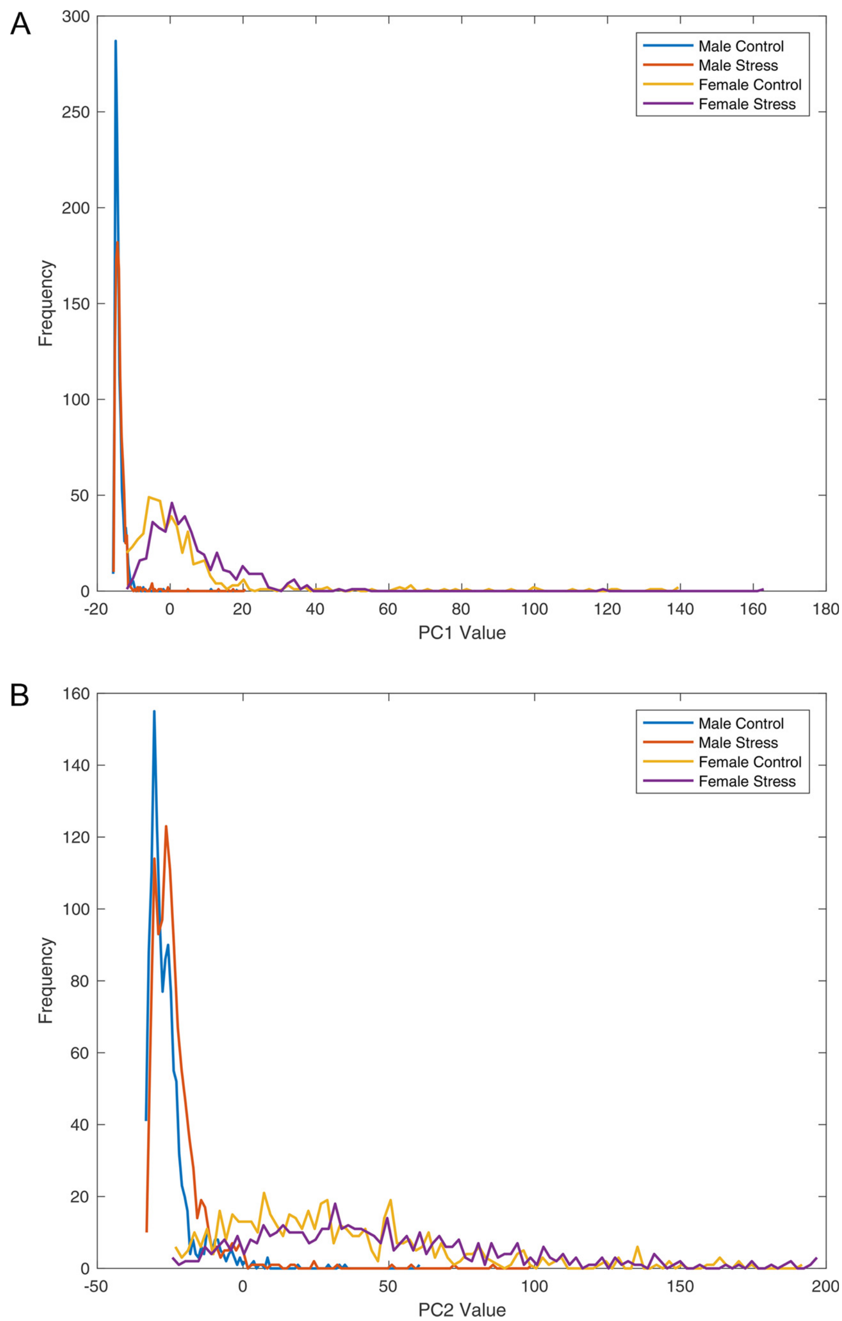

Fig. 1. Density distribution of PC1 A) and PC2 B) according to sex and stress condition. PCA was performed on all the calls.

another (different) subpopulation, the performance of discrimination dramatically crashed and constituted a very small improvement compared to random selection.

To conclude, our analysis shows that control and stress calls are acoustically distinct, but the correct assignment of an emotional state to a caller is efficient only with previous knowledge of control and stress versions of the calls of the same individual.

3.3. Impact of stress on individual discrimination based on calls

Instead of classifying conditions, we train the random forests to classify individuals using their calls. As there are 57 individuals in the dataset (32 males/ 25 females) we trained the random forests to classify the 57 categories. As seen in Fig. 3A, we first trained the random forests on all calls according to the three features datasets which obtained respectively 21.4 ± 0.1% (SPEC), 11.4 ± 0.1% (ACOUS) and 9.5 ± 0.1% (COMB) of error rate. We then broke down according to stress status (see Fig. 4A) yielding C SPEC: 19.4 ± 0.2% C ACOUS 11.6 ± 0.2% and C COMB 9.8 ± 0.2% when restricted to control calls, and S SPEC: 18.4 ± 0.2% S ACOUS 10.4 ± 0.2% and S COMB 8.4 ± 0.2% when restricted to stress. The best results were obtained for COMB dataset which performed equally well (between 8.4% and 9.9% of error) regardless of set choices. Since, in the case of pure random selection the expected error rate is equal to one minus the inverse of the number of individuals. This amounts to 98.3%, 97.1% and 96.8% error rate for all, male and female calls respectively. Therefore, when trying to identify individuals using either their control calls, stress calls or both, the random forest is ten times better than the random selection. There is no impact of the emotional status on the power of discrimination. We can conclude that calls are highly individualized, and emotional state does not impair individual signatures.

Fig. 3B shows the same results as 4A for all calls and displays also

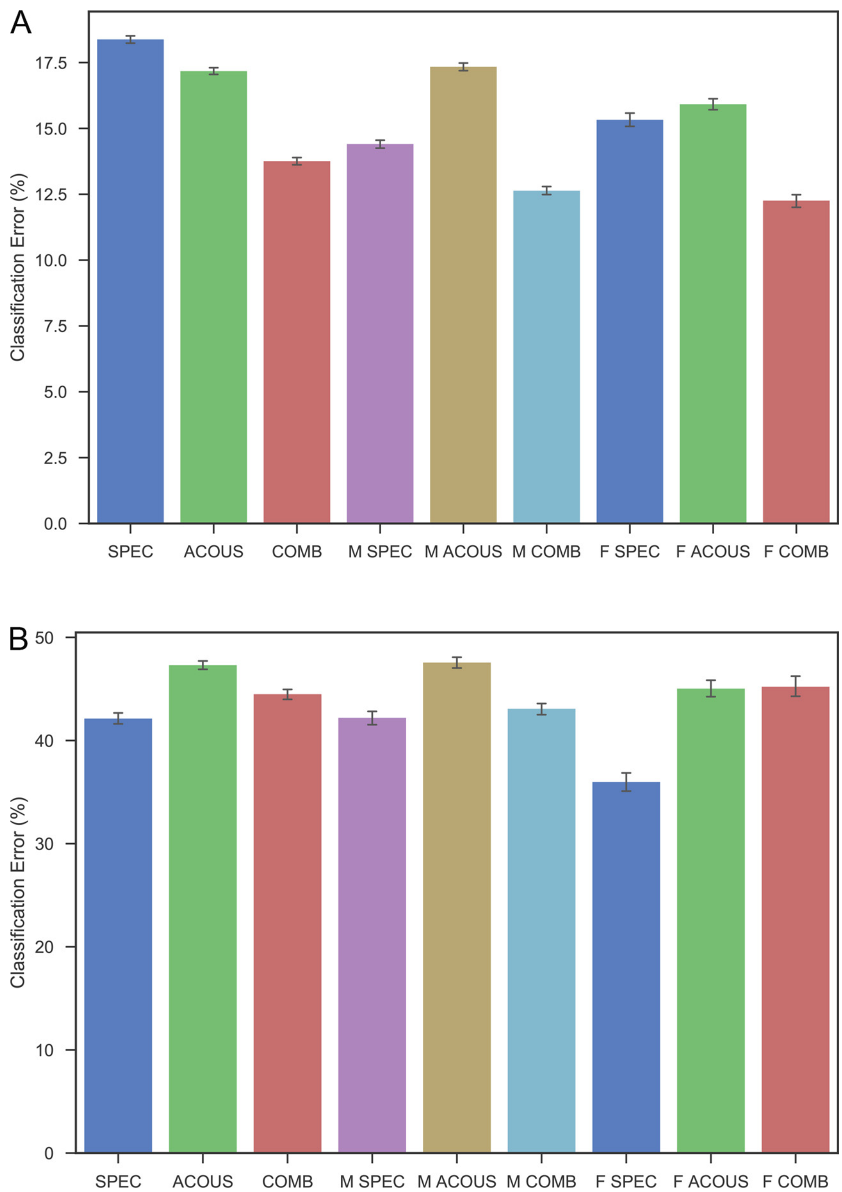

Fig. 2. Learning results on stress with three data matrices: spectrogram (SPEC), acoustical data (ACOUS) and combined (COMB) for all the calls, male calls and female calls separately. A) Control/Stress classification errors performed on all the calls separately using the calls of all the individuals in the training set. Individuals in the training and validating set are the same. B) Control/Stress classification errors with dataset separated by individuals. The training is performed on the calls 75% of the individuals and predicted on the remaining 25%. Individuals tested in the validating data set were different from the ones in the training data set.

the individual classification results when the data set is restricted to male: M SPEC 7.0 ± 0.1% / M ACOUS 8.6 ± 0.1% and M COMB 4.4 ± 0.1% and to female F SPEC 24.3 ± 0.2% F ACOUS 18.3 ± 0.3% and F COMB 13.1 ± 0.25%. This result suggests that sex has an impact on discrimination. We obtain a very low error (4.4%) for male discrimination (for COMB) whereas the best we can obtain is 13.1% for female making female calls harder to discriminate than males'.

Since the previous results were obtained for both control and stress calls, we, for sake of completeness, present the results on Fig. 3C - using acoustic parameters only - of the individual discrimination splitting between sex and emotional status. Male discrimination error state in control condition is 9.26 ± 0.17% whereas in stress condition is 8.4 + ± 0.18%. The female results are 18.6 ± 0.35% for control and 14.5 ± 0.34% for stress. When broken down by sex and emotional status, stress calls are easier to discriminate than control but for both sets discrimination is ten times better than random selection making them easily discriminable. In short, individuality is not lost by modifications due to emotional state.

Individual discrimination using calls of different emotional states

We then tried to predict the identity of the caller using the calls in a given emotional state (e.g. stress) while learning on the other emotional state (e.g. control). Prediction results are displayed on Fig. 3D. The classification errors using all control calls for training and all stress calls for validating ("C- > S" 46.6 ± 0.1%) gets comparably high error rates than the reverse ("S- > C" 51.0 ± 0.1%). On male calls only, the discrimination is better ("C- > S" 35 ± 0.1% and "S- > C" 42.7 ± 0.1%) than on female calls (C- > S" 72.4 ± 0.2% and "S- > C" 71.0 ± 0.2%) but remains high (one call out of three is attributed to the wrong individual). Again random selection error rates are respectively 98.3%, 97.1% and 96.8%. Thus, while still better than the random selection, error rates are high when predicting the identity of a caller using its calls in a given emotional state while learning on calls of another emotional state. Individuality is harder to discriminate in one emotional state if we know the calls of individuals only in the other state.

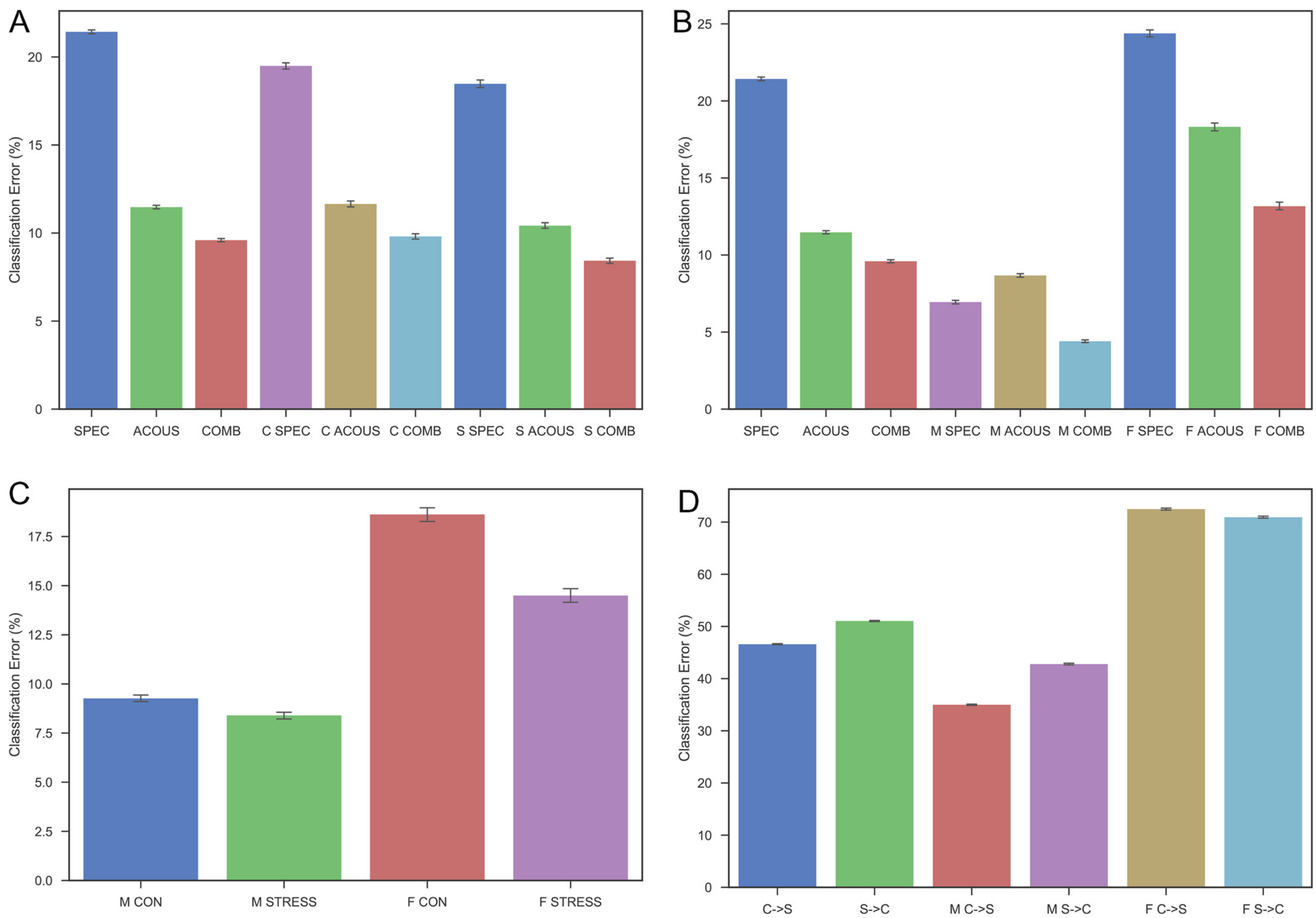

Fig. 3. Individual discrimination results: A) percentage of misclassification according to the data used (SPEC/ACOUS/COMB) for all calls and according to stress status (both sexes) B) percentage of misclassification according to the data used (SPEC/ACOUS/COMB) for all calls and according to sex status (both status) all calls and C) classification errors for the ACOUS data set separating both sexes and stress D) individual classification error using either control (C) or stress (S) data as training set and stress (S) or control (C) as validating separated by sex. This was performed using the ACOUS dataset.

3.4. Discriminating acoustic features

Random forests use features for learning and provide the GINI index to quantify the importance of each feature for the classification. A high GINI value means that the feature is important for the discrimination. Because we used 16 acoustic parameters, the average explanation percentage is 6.25%. We computed the GINI index for all the classification tasks above and display the results as a heatmap on Fig. 4. For the control/stress discrimination, the three features with the highest GINI value are the *amplitude* (9.45%) *duration* (8.41%) and *temporal standard deviation* (SD Time 7.5%). When restricted to female calls, the first three are the *entropy time* (9.70%), *duration* (9.1%) and *standard deviation of spectrum* (SD spectrum 8.3%) whereas for male calls the first three are *amplitude* (10.77%) *duration* (8.15%) and *temporal standard deviation* (SD Time 7.03%). While it differs depending on the sex, duration and amplitude were among the most important variables used to discriminate between control and stress calls.

For individual discrimination tasks, the most discriminant features are the duration (12.14%), temporal standard deviation (SD Time 11.5%) and mean time (10.6%) when considering all calls. When restricted to control calls the most discriminant feature is the temporal standard deviation (11.71%) whereas when restricted to stress calls the most discriminant feature is the duration (13.73%). As expected, important cues for individual discrimination slightly differ between stress and control conditions. When breaking down according to sex, the most discriminant features for male control calls are temporal standard deviation (11.0%), mean time (9.4%) and duration (8%) whereas it is duration (14.6%), temporal standard deviation (13.0%), mean time (11.2%) for male stress calls. In contrast, for female, duration (12%), mean spectrum (10%) and Q1 (8.5%) are the most important for control calls, and duration (14.4%), entropy time (10.5%) and mean spectrum (8.3%) for stress calls. Temporal feature such as temporal standard deviation and duration emerged as the most prominent features for individual discrimination whether we separate the dataset by sex or not.

4. Discussion

Here we confirm that zebra finch distance calls reliably convey both individual identity and stress, and we show for the first time how both pieces of information influence one another.

Using random forests, we show that whatever the sets of parameters (full spectrogram, selected acoustic parameters, or both combined), control and stress calls are discriminated with error rates well below the random selection. In short, it is clearly possible when confronted to several exemplars of control and stress calls of various individuals to predict robustly the emotional state of the caller. This result seems in favor of the universality hypothesis, stating that stress modifies all individuals' calls the same way. However, stress calls can be distinguished from control calls as long as both versions from the same callers have been used for learning. When trying to predict on a *new* (i.e not trained for) sub-population of callers, performances collapse almost to the level of random selection. Therefore, it seems reasonable to conclude here that calls can serve as an emotional carrier but only between familiar

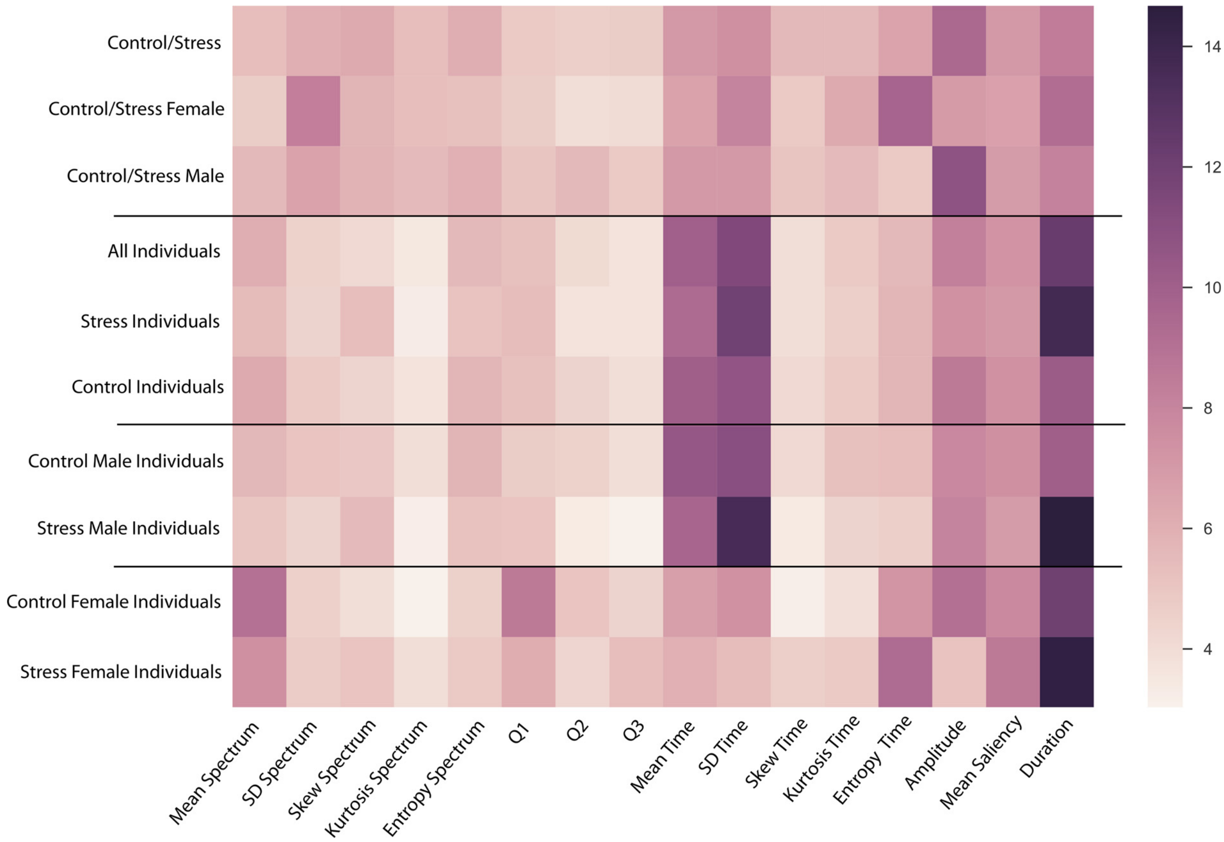

Fig. 4. Heatmap on GINI as percentage of importance for stress/control classification (first three rows) and individual discrimination (remaining rows). All rows sum to 100%.

birds. This result is in favor of the familiarity hypothesis, i.e. it is possible only for a familiar listener to detect stress in the vocalizations of a caller.

How is individual identity maintained in calls through stress? First, we show that it is very easy to recognize an individual when confronted to both versions of calls (both stress and control) as well as restricted to only one emotional status. Indeed, we obtain errors below 10% - compared to a 95% error rate for the random selection. This success is independent on the emotional status. Whereas stress changes the call structure and properties it does not alter identity recognition (as estimated by performance). This is consistent with previously published results on females' perception (Perez et al., 2015), which showed that in an operant discrimination task, females grouped the stressed calls and the control calls of the same male in the same category.

When breaking down by sex, results are even better for male calls (reaching an extremely low 4% error rate for combined data). Results are fair for female calls but it seems their identity is harder to discriminate using their calls. Changes due to emotional state do not change individual vocal signature enough to alter individual discrimination, which remains very efficient even when merging control and stress calls. This is again in favor of the familiarity hypothesis, stating that stress modifications of call structure remain in the caller's individual acoustic space, making discrimination of control and stress calls of one individual only possible for familiar listeners.

In contrast, when trying to predict individual identity on calls of one emotional state with the previous knowledge of the other emotional state only (e.g. predicting identity on stress calls with previous knowledge of control calls only), the error rate increased dramatically. In short, while stress calls retain individual identity information, stress does blur individual signature, and it is more difficult to classify stress calls to the correct caller if only control calls were previously known (and the reverse is true). This suggests individual recognition is more robust for listeners that have heard calls of an individual in both states.

Random forest can provide a value called GINI that indicates for each parameter its importance for the classification. When considering both sexes together, most important parameters for discrimination of control and stress calls were duration, maximum amplitude and standard deviation of the waveform (SD Time). Stress increased calls' duration, maximum amplitude and dispersion in amplitude within call. As shown above, these parameters allow for discrimination between emotional states only if both exemplars of calls are available for training. In other words, while stress calls are usually longer and louder, they are mostly longer and louder than their control counterparts. When considering the sexes separately, most discriminant parameters were the same for males, and stress modifies female call structure by mostly increasing duration, increasing entropy of the waveform (entropy time) and decreasing dispersion in the call's frequency spectrum (spectral standard deviation).

Many studies, including our own (Perez et al., 2012; 2015), reported that stress, as an emotion of high arousal and negative valence, impacts the acoustic structure of vocalizations by increasing average frequency, pitch (F0) and amplitude (reviewed in Briefer, 2012). In the present results, maximum amplitude and duration are the parameters that best discriminate control and stress calls, and both increased with stress. Even though they vary significantly between control and stress, spectrum parameters are not the most discriminant parameters of stress and control calls. This seems counterintuitive but the machine learning learns to separate all control calls from stress calls irrespective of individuality. In a pool of calls made by various individuals the best parameters to separate stress from control is to start with duration and amplitude. Spectral parameters can vary for individual between its control and stress calls but this variation is blurred when pooled with other individuals.

For individual discrimination, the most discriminant parameter is duration when considering only stress calls, but standard deviation of the waveform (temporal standard deviation) when considering control calls only. The fact that salient cues for discrimination are different between stress and control helps explaining the apparent contradiction between the fact that stress does not influence recognition but this recognition is less robust when learning on calls of one emotional state and classifying calls of the other state. Indeed, discrimination does not operate on the same cues.

When considering the sexes separately, most discriminant parameters for individuality in male control calls are duration, mean of the waveform (meantime) and standard deviation of the waveform (temporal standard deviation), whereas in female control calls they are duration, mean of the spectrum and first quartile (Q1).

Even if they use a diversity of parameters, the coding of emotional state (control/stress) and individuality share common call parameters important for acoustic discrimination (duration, standard deviation of the wave form). Consequently, stress calls do not have generic properties which allow any listener to detect stress in the calls of any caller (universality hypothesis). Instead, it is necessary to compare stress calls to control calls of the same caller to identify stress. Therefore, stress is not translated in zebra finch distance calls as a generic feature with clear meaning to everyone (as opposed to e.g. an alarm call). Indeed, to convey stress information, the caller must be known acoustically in both emotional conditions by the receiver. This result is consistent with Perez et al. (2015) in which we showed physiological resonance of female zebra finches in response to the stress calls of their mate.

While this is out of the scope of the present manuscript, it is probably a future venue to examine the differences in both parameters and performance in discrimination depending on the origin or nature of the stress. In the present dataset, we pooled together stress induced by isolation and by corticosterone ingestion. While they share common properties, these two stress conditions can potentially have a different impact and be discriminated. It is entirely possible that various stress typologies lead to differentiated call structure that can be assessed and differentiated by birds.

Finally, because we used very efficient machine learning tools, we could assess information quantity using classification errors and we can conclude that our results represent the potential bottom line of discrimination. We are fully aware that birds do not need to be at the bottom line and that behavioral experiments must be conducted to assess the bird's actual classification power. By using general acoustic parameters as well as full spectrograms, classification was obtained on a comprehensive set of parameters that are fully available to the birds themselves. This very general description of acoustic information is extremely powerful since it allows us to be the most comprehensive and with the fullest discrimination power.

Acknowledgments

We wish to thank Julie Elie and Frederic Theunissen for their help for the acoustic parameters. This work has been partly supported by the CRCNS NSF/ANR grant.

References

Aubin, T., Jouventin, P., 1998. Cocktail–party effect in king penguin colonies. Proc. Biol Sci. 265 (1406), 1665–1673.
Banerjee, S.B., Adkins-Regan, E., 2011. Effect of isolation and conspecific presence in a novel environment on corticosterone concentrations in a social avian species, the zebra finch (Taeniopygia guttata). Horm. Behav. 60 (3), 233–238.
Bee, M.A., Gerhardt, C.H., 2002. Individual voice recognition in a territorial frog (Rana catesbeiana. Proc. Biol Sci. 269 (1499), 1443–1448.
Beecher, M.D., 1989. Signaling systems for individual recognition: an information theory approach. Anim. Behav. 38, 248–261.
Bradbury, J.W., Vehrencamp, S.L., 2011. Principles of Animal Communication, 2nd edition. Sinauer Associates.
Breiman, L., 2001. Random forests. Machine Learning 45 (1), 5–32.
Briefer, E.F., 2012. Vocal expression of emotions in mammals: mechanisms of production and evidence. J. Zool. 288 (1), 1–20.
Briefer, E., Vannoni, E., McElligott, A.G., 2010. Quality prevails over identity in the sexually selected vocalisations of an ageing mammal. BMC Biol. 8 (1), 35.
Charrier, I., Mathevon, N., Jouventin, P., Aubin, T., 2001. Acoustic communication in a black-headed gull colony: how do chicks identify their parents? Ethology 107 (11), 961–974.
Elie, J.E., Theunissen, F.E., 2016. The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. Anim. Cognit. 19, 285–315.
Forstmeier, W., Burger, C., Temnow, K., Derégnaucourt, S., 2009. The genetic basis of zebra finch vocalizations. Evolution 63, 2114–2130.
Hernandez, A.M., Perez, E.C., Mulard, H., Mathevon, N., Vignal, C., 2016. Mate call as a reward: Acoustic communication signals can acquire positive reinforcing values during adulthood in female zebra finches (*Taeniopygia guttata*). J. Comp. Psychol. 130, 36–43.
Janik, V.M., Sayigh, L.S., Wells, R.S., 2006. Signature whistle shape conveys identity information to bottlenose dolphins. PNAS 103 (21), 8293–8297.
Jansen, D.A., Cant, M.A., Manser, M.B., 2012. Segmental concatenation of individual signatures and context cues in banded mongoose (Mungos mungo) close calls. BMC Biol. 10 (97).
Jouventin, P., Aubin, T., 2002. Acoustic systems are adapted to breeding ecologies: individual recognition in nesting penguins. Anim. Behav. 64 (5), 747–757.
Kilner, R., Johnstone, R.A., 1997. Begging the question: are offspring solicitation behaviours signals of need? Trends Ecol. Evol. 12 (1), 11–15.
Leavesley, A.J., Magrath, R.D., 2005. Communicating about danger: urgency alarm calling in a bird. Anim. Behav. 70, 365–373.
Manser, M.B., 2001. The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. Proc. Biol Sci. 268 (1483), 2315–2324.
Mariette, M.M., Griffith, S.C., 2012. Nest visit synchrony is high and correlates with reproductive success in the wild Zebra finch Taeniopygia guttata. J. Avian Biol. 43 (2), 131–140.
Marler, P., 2004. Bird calls: Their potential for behavioral neurobiology. Ann. N. Y. Acad. Sci. 1016, 31–44.
Marler, P., Slabbekoorn, H., 2004. Nature's Music – The Science of Birdsong. Elsevier Academic Press.
Mathevon, N., Aubin, T., Vieillard, J., da Silva, M., Sebe, F., Boscolo, D., 2008. Singing in the Rain forest: how a tropical bird song transfers information. Plos One 3 (2), e1580.
Morton, E.S., 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. Am. Nat. 111, 855–869.
Mouterde, S., Theunissen, F.E., Elie, J.E., Vignal, C., Mathevon, N., 2015. Acoustic communication and sound degradation: how does the individual signature of zebra finch calls transmit over distance? PloS One 9 (7).
Pedregosa, F., et al., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 28252830.
Perez, E.C., Elie, J.E., Soulage, C.O., Soula, H.A., Mathevon, N., Vignal, C., 2012. The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? Horm. Behav. 61 (4), 573–581.
Perez, E.C., Elie, J.E., Boucaud, I.C.A., Crouchet, T., Soulage, C.O., Soula, H.A., Theunissen, F.E., Vignal, C., 2015. Physiological resonance between mates through calls as possible evidence of empathic processes in songbirds. Horm. Behav. 75, 130–141.
Randall, J.A., Rogovin, K.A., 2002. Variation in and meaning of alarm calls in a social desert rodent rhombomys opimus. Ethology 108 (6), 513–527.
Remage-Healey, L., Adkins-Regan, E., Romero, L.M., 2003. Behavioral and adrenocortical responses to mate separation and reunion in the zebra finch. Horm. Behav. 43, 108–114.
Rendall, D., 2003. Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. J. Acoust. Soc. Am. 113 (6), 3390–3402.
Schibler, F., Manser, M.B., 2007. The irrelevance of individual discrimination in meerkat alarm calls. Anim. Behav. 74 (5), 1259–1268.
Soltis, J., Leong, K., Savage, A., 2005. African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. Anim. Behav. 70 (3), 589–599.
Tobias, M.L., Barnard, C., O'Hagan, R., Horng, S.H., Rand, M., Kelley, D.B., 2004. Vocal communication between male Xenopus laevis. Anim. Behav. 67 (2), 353–365.
Vignal, C., Mathevon, N., Mottin, S., 2004. Audience drives male songbird response to partner's voice. Nature 430, 448–451.
Vignal, C., Mathevon, N., Mottin, S., 2008. Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. Behav. Processes 77, 191–198.
Yin, S., McCowan, B., 2004. Barking in domestic dogs: context specificity and individual identification. Anim. Behav. 68, 343–355.
Zann, R.A., 1996. The Zebra Finch: A Synthesis of Field and Laboratory Studies. Oxford University Press, Oxford, pp. 335.