At sea vocal repertoire of a foraging seabird

Andréa Thiebault[1], Isabelle Charrier[2], Pierre Pistorius[1,3], Thierry Aubin[2]

[1]Marine Apex Predator Research Unit, Institute for Coastal and Marine Research, Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa

[2]CNRS UMR 9197, Institut des Neurosciences Paris-Saclay, Université Paris Sud, Orsay, France

[3]DST/NRF Centre of Excellence at the Percy FitzPatrick Institute, Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa

Corresponding author: Andréa Thiebault, Marine Apex Predator Research Unit, Institute for Coastal and Marine Research, Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa. E-mail: andrea.thiebault@gmail.com

Abstract

Seabirds spend most of their time at sea, yet our knowledge of their activities and behaviour is limited due to difficulties of *in-situ* data collection. In particular, we know virtually nothing about their acoustic communication when at sea. We benefited from the recent development of miniaturised audio-recording devices to deployacoustic recorders on breeding Cape gannets *Morus capensis* to study their vocal activity while foraging. Call sequences were recorded on 1718 occasions, from which acoustic variables were measured on calls with good recording quality. A total of 1348 calls from 18 birds were measured in temporal and frequency domains. Each call was assigned to a behavioural context defined acoustically: sitting on the water, flying, taking off or just before diving. Potential discrimination among calls from different contexts was tested using the random forest algorithm. Within each context, individual stereotypy in the calls was assessed per acoustic variable using a measure of potential of individual coding, and as a combination of variables using a similar multivariate analysis. The acoustic structure differed according to the behavioural context (global accuracy of prediction 75 %). Temporal variables (sequence and call duration sequence and) were most important to correctly classify the calls among the four contexts. When considering only two contexts, on the water and in the air (merging flying and diving), frequency and spectral variables (percentage of energy below 1200 Hz and fundamental frequency) were of most importance (accuracy 86 %). A combination of acoustic variables was necessary to discriminate individuals, but calls from all contexts were not strongly individually distinct (accuracy 41 % - 63 %). We provided the first detailed acoustic analysis of a foraging seabird and demonstrated context-specific acoustic structure in its vocalisations at sea. Our results suggest that seabirds use vocal communication to exchange various types of information that likely improves foraging success.

Keywords: bioacoustics, biologging, cape gannets, group foraging, social interactions, random forest

Introduction

Birds interact socially with their conspecifics or other species using mainly visual and acoustic signals. The latter, in particular, fundamentally contributes to vital biological functions such as feeding (Elgar 1986, Stoddard 1988), breeding (Kroodsman et al. 1982), localisation in space (Mathevon et al. 2008), warning about danger (Aubin 1991), or species and individual's identification (Ceugniet et al. 1999, Draganoiu et al. 2006). Studies on acoustic communication have been largely conducted on terrestrial bird species, while for seabirds they have been limited by their elusive lifestyle.

Seabirds constitute a group of more than 300 species classified in five Orders. They show a diversity of morphological and behavioural adaptations developed to exploit the marine environment, from inshore to offshore as well as from the surface to the seabed through the water column. Beyond this diversity, seabirds are often observed gathering in groups (Hoffman et al. 1981, Harrison et al. 1991, Danchin and Wagner 1997, Camphuysen and Webb 1999, Clua and Grosvalet 2001), which suggests that social interactions are important in their respective life-history strategies.

On land, about 95 % of seabird species breed colonially (Danchin and Wagner 1997). The accessibility of these birds during the breeding stage have allowed for the study of vocal communications in several species. Species-specific, individual and sexual coding have been found in the vocal signals of Spheniscidae (Aubin et al. 2000, Aubin and Jouventin 2002), Sulidae (White et al. 1970, Dentressangle et al. 2012, Krull et al. 2012), Laridae (Charrier et al. 2001b, Mulard et al. 2009), Stercorariidae (Charrier et al. 2001a) and Procellariidae (Bourgeois et al. 2007, Curé et al. 2009, 2012) and seem to be common in seabirds in a breeding context. Nevertheless, seabirds spend most of their lives at sea and the current knowledge of their vocal behaviour at sea is extremely poor.

At sea, these birds generally forage on spatially dynamic patchily distributed prey (Russell et al. 1992, Weimerskirch 2007). They need to acquire information on the location of food during each and every foraging trip. Oceanographic fronts (Schneider 1982, Bost et al. 2009, Scales et al. 2014) and prior experience (Davoren et al. 2003, Carroll et al. 2018) can be used to narrow down the search to productive areas. At a finer scale, seabirds cue on conspecifics and other marine predators to locate the presence of prey (Silverman et al. 2004, Tremblay et al. 2014). As a consequence, species from all seabird orders commonly gather at sites where prey resources are concentrated, most often in multi-species associations (Siegfried et al. 1975, Hoffman et al. 1981, Harrison et al. 1991, Camphuysen and Webb 1999, Clua and

Grosvalet 2001). In this context, individuals may benefit from group foraging (Ryan et al. 2012, Lett et al. 2014, Thiebault et al. 2015, McInnes et al. 2017) and use both visual and acoustic cues (Thiebault et al. 2014b, 2016). Despite numerous descriptions of seabird aggregations (mainly from boat observations), studies addressing the mechanisms involved in their social interactions at sea have long been constrained by technical limitations. Since the 1990's, tremendous technological development has improved the ability to track animals beyond our boundaries of perception (Boyd et al. 2004, Kooyman 2004), thereby advancing our knowledge of seabird ecology and behaviour at sea (Ropert-Coudert and Wilson 2005). Nonetheless, studies are limited by the devices used (often measuring a single variable at a time such as the geographical position or the body acceleration) and the (often small) number of individuals that can be equipped simultaneously (Cooke et al. 2004). As a consequence, tracked individuals are often considered moving alone in the seascape. It is only recently that studies quantifying seabirds interactions at sea have gained impetus thanks to the miniaturisation of image recording devices in particular (Tremblay et al. 2014, Berlincourt and Arnould 2014, McInnes et al. 2017). Although visual observation of *in-situ* foraging seabirds have greatly improved interpretation of remote movement data, the study of their vocal communication at sea has only very recently started (Thiebault et al. 2016, Choi et al. 2017).

Here, we aimed to study the acoustic structure of calls emitted at sea by a seabird. We chose the Cape gannet *Morus capensis* because of the known importance of social interactions in the foraging strategies of this particular species (Thiebault et al. 2014a, b, 2015, 2016). The Cape gannet is a monomorphic seabird endemic to southern Africa with adults breeding in dense colonies (Nelson 2005). Individual signature was found in the calls emitted at the colony in their closest related species, the Northern gannet *M. bassanus* (White and White 1970, White 1971) and the Australasian gannet *M. serrator* (Krull et al. 2012). At sea, Cape gannets feed mainly on small pelagic fish, sardines *Sardinops sagax* and anchovies *Engraulis encrasicolus* (Green et al. 2014). They use the central place of the colony to gather information on the feeding ground, tracking the direction from which conspecifics return (Thiebault et al. 2014a). They also use the aggregations of predators as cues of the presence of prey over large distances (Thiebault et al. 2014b). Once aggregated at a fish school, they benefit from the quasi-synchronisation of attacks by the group of predators to capture prey (Thiebault et al. 2015). During those foraging activities, we know that Cape gannets do emit calls, but no detailed acoustic analysis could previously be done due to the poor quality of the sound recorded (Thiebault et al. 2016). In particular, we do not know if the calls emitted in different behavioural contexts have different acoustic structures. To address this question, we benefited

from the recent development of miniaturised audio-recording devices, which have for example recently been used on free-ranging small mammals (Ilany et al. 2013, Couchoux et al. 2015), terrestrial birds (Anisimov et al. 2014, Stowell et al. 2017, Gill et al. 2016) and bats (Greif and Yovel 2019). Here we faced the challenge of recording the vocalisations of wild seabirds when foraging, for which we need recorders small enough to be carried by a flying animal and both shock- and water-proof to resist the dives of a marine animal, still recording high quality vocal activity. We first investigated the acoustic structure of calls in relation to their behavioural context of emission, thereby testing for the existence of a repertoire of different call types in a seabird when foraging. Secondly, we investigated individual-specific signatures in these vocalisations.

Materials and methods

Data collection

Fieldwork took place on Bird Island (Algoa Bay, South Africa) during December 2015. Twenty chick-rearing Cape gannets were deployed with devices to record their behaviour and vocal communication while foraging at sea. Nests were observed unceasingly until a change-over (characterised by a display dance) between two partners was observed. The bird on departure to sea was captured near the nest using a pole with a hook on the end (thus staying as far from the colony as possible). Only one parent (of either sex since the species is monomorphic) was captured per nest and devices were attached for one foraging trip only, while the partner was on the nest guarding the chick. Nests were then monitored every hour from sunrise to sunset, so the deployed birds were captured again soon after their return to the colony and the devices were retrieved. Birds were handled for eight and six minutes on average for the first and second capture respectively. It consisted of attaching devices (using adhesive tape, Tesa, Germany, see Thiebault et al. 2014a), measuring the bird's body mass and marking the bird with a spot on the head for further identification (using sticks for short-term animal marking, Raidex, Germany) for the first capture, and retrieving devices and taking standard measurements (not used in this study except for body mass) for the second capture. Body mass was measured using a spring balance (Pesola, Switzerland; precision 50 g). Devices deployed included acoustic recorders, video cameras, global positioning system devices (GPS) and time-depth recorders (TDR). Acoustic recorders were deployed in combination with a GPS on eight birds (total mass 60 g), a GPS and a video camera on one bird (90 g), or a TDR and a video camera on eleven birds (80 g). Only the data from the acoustic recorders were used in this study. Audio recorders (Edic-mini Tiny+ B80, frequency response 100 Hz – 10 kHz ± 3 dB, 65 dB dynamic range, TS-Market

Ltd., Russia, fitted with a CR2450 battery) were set up to record sound in mono at a 22.05 kHz sampling frequency. They recorded continuously, hence collecting data during the whole foraging trips of the birds (one to two days). Five different units were deployed successively on the 20 individuals. The main challenge for collecting such acoustic data was to ensure high quality recordings on board of a flying and diving bird. To limit disturbance from the wind, we placed the audio recorder on the lower back of the bird, under feathers and facing backward. In addition, a thin layer of foam was added after the first deployment to reduce flow and background noise. To keep the devices sufficiently dry when immersed in the sea water but still ensure sound recordings at a decent amplitude level (avoiding thick waterproof casing), we sealed the microphones in nitrile gloves (amplitude attenuation of 3 dB measured in the laboratory, no modification of the frequency response).

Potential instrument effects

The total mass of devices attached to the birds (60-90 g) corresponded on average to 2.7 % (range 1.1 % - 3.4 %) of their body mass (2350-2975 g) . This is within commonly accepted practice in seabird research (Phillips et al. 2003, Vandenabeele et al. 2012).

The body mass of equipped birds measured before and after their foraging trip was compared in order to assess the capacity of those birds to fulfil the purpose of their foraging trip (finding and capturing food), despite the presence of devices on their back.

In order to further check for potential effect of our devices on the behaviour of the birds, we compared the foraging trip duration between equipped and non-equipped birds, as well as between trips with or without devices for equipped birds. Foraging trip duration has been shown to be a reliable proxy for foraging effort in Cape Gannets (Pichegru et al. 2007, Mullers et al. 2009, Rishworth et al. 2014a) and is regularly used to assess the effects of tagging on seabirds (Wilson et al. 1986, Phillips et al. 2003). In particular, the presence of devices has been shown to increase foraging trip duration (Croll et al. 1991, Ballard et al. 2001, Passos et al. 2010). We monitored the attendance of both partners (every hour from sunrise to sunset) on all the nests where a bird was equipped with devices. On those nests, only the equipped birds were captured twice (for deployment and retrieval of devices) and marked for identification. In addition, we monitored the attendance of both partners on 10 independent nests in the vicinity, where a chick of similar size was reared. On those nests both partners were captured once (duration <1min) and marked differently (left VS right wing) for identification on the nest. The nest monitoring was undertaken in the

shortest time possible (usually within five minutes) and at the largest distance possible from the nests (just close enough to identify markings) in order to limit potential observer effects. The effect of being a control bird, a deployed bird, or a partner of a deployed bird on the duration of trips was tested using linear mixed-effects models (package "lme4" in R software, (R Core Team 2013, Bates et al. 2014)). A null model including only individual birds as a random effect was compared to a model including the group (control, deployed, partner) as a fixed effect in addition to the random effect using an analysis of variance. Among the deployed birds, the effect of carrying devices was tested in a similar way. Residuals of all models were checked graphically for normality and heteroscedasticity. Results are shown as mean ± standard deviation.

Measure of acoustic variables

Sound data were resampled at 12 kHz to increase the accuracy of frequency measurements and because no call was observed to contain energy at frequencies higher than 5 kHz. All the vocalisations were analysed using Avisoft-SASLab Pro (version 5.2.09, Avisoft Bioacoustics, Germany). The spectrogram of each recording (Hamming function, FFT 512 points window size, 75 % overlap) was visualised over a sliding window of 20 sec length to identify and label all the calls. Calls were produced in sequences, and sometimes composed of two parts (with only one occurrence of three parts), identifiable by a drop in amplitude (a call made of two parts is illustrated on Figure 1d). Calls were selected for measurements wherever the quality of the recordings allowed (i.e. low background noise, good signal to noise ratio, no overlap with other vocalisations, no clipping). The background noise was due to wind and water sounds, and its energy was mostly concentrated below 300 Hz. The calls produced at sea by the Cape gannets were tonal sounds, they did not contain noisy components, and they were produced above 300 Hz. We thus assume that there was no bias in selecting calls from low background noise, only an improvement of the acoustic measurements performed.

Selected calls (only the first part if the call was made of two parts) were measured in both temporal and frequency domains. Temporal variables were measured on the oscillogram and included the duration of the sequence (DurSeq, s), the duration of the call measured (DurCall, s) and if more than one call was emitted in the sequence, the rhythm of call emission in the sequence (number of calls divided by duration of the sequence, Rhythm, s^{-1}). Spectral features were extracted from the average energy spectrum displayed between 300 and 5000 Hz and included the fundamental frequency (F0, Hz), the frequency of maximum amplitude (Fmax, Hz), the second (Q50, Hz) and third quartiles (Q75, Hz) of the energy

distribution, and the percentage of energy occurring below 1200 Hz (E1200, %). From the spectrogram, the ascending (between start and inflection point, SlopeAsc, Hz.s^{-1}) and descending (between inflection point and end, SlopeDesc, Hz.s^{-1}) slopes of the call were measured. In addition, the number of calls in the sequence (Ncalls) and the number of parts composing the call (Nparts) were counted.

For multivariate analyses, the set of acoustic variables was tested for multicollinearity at the level of 0.8 (Kuhn 2008) and the variables Q50 and Ncalls were removed to limit potential overestimation of importance for highly correlated variables (Strobl et al. 2007). The variable Rhythm was also removed because it could not be measured on call sequences composed of a single call. So, in the end, nine acoustic variables were included in the multivariate analyses: DurSeq, DurCall, F0, Fmax, Q75, E1200, SlopeAsc, SlopeDesc, Nparts.

Acoustic structure in relation to behavioural context

The context of call emission was defined from our knowledge of the behaviour of the study species, based on previous work with observations from bird-borne video cameras with built-in microphones (Thiebault et al. 2014a, 2016). Four categories were defined: on the water (WAT), taking off (TOF), flying (FLY) and just before diving (DIV). The context of each call sequence was then identified from the sound data, with taking off, landings on the water or diving events clearly audible and identifiable from the spectrograms (Supplementary Figure 1, Supplementary audio recording 1, Supplementary audio recording 2). The acoustic stucture of the calls produced in those different contexts were compared to assess if they are structurally different, and if they really correspond to four different call types.

The distribution of each variable per context was tested for normality using a Shapiro-Wilk test. As none of them were normally distributed, their variance was compared using a Fligner-Killeen test of homogeneity of variance and their median using a Kruskal-Wallis rank sum test.

The acoustic structure of calls emitted in the different behavioural contexts was then compared in a multivariate analysis, using the random forest algorithm (RF) for classification (details below). In addition to an analysis comparing the four groups, calls emitted on the water (WAT) and in the air (combining FLY and DIV) were compared to study potential acoustic differences between the calls emitted in these two different environments. Finally, calls defined as FLY and DIV were compared to study potential finer acoustic differences between calls in a similar environment.

Those analyses were conducted on the entire dataset as a whole, thereby assessing the potential existence of call types in the population. Similar analyses were conducted per individual in order to test for potential individual variability and they led us to the same conclusions on different call types (results not shown).

Individual signature

The analysis of individual signature was conducted per context of emission, including WAT, FLY and DIV. The context TOF was excluded from this analysis because of a small sample size (Supplementary Table 1). The calls measured were assumed to be emitted by the equipped bird. In the contexts FLY and DIV, neighbouring birds must be situated at a minimum of 2 m (average wingspan 170 cm) limiting greatly the risk of recording calls from surrounding birds, except if the birds were flying staggered on top of each other at the moment of calling. In the context WAT, birds can be sitting closer to each other, increasing the risk of recording calls from neighbouring birds. However, given the short active space of the audio recorders, calls emitted by surrounding birds would have been of noticeably lower amplitude (Gill et al. 2016), and thus not selected for acoustic measurements. Only individuals for which at least eight calls were recorded in a given context were kept in the analyses. We assessed the potential of individual coding (PIC) for each acoustic variable and per call context. We used the coefficient of variation (CV): $CV = \{100(SD/Xmean)[1+(1/4n)]\}$, where SD is the standard deviation, Xmean the mean of the sample and n the population sample (Sokal and Rohlf 1995). This coefficient was calculated for all the individuals pooled together (CVbetween) and for each individual (CVwithin). The PIC is then calculated by dividing the CVbetween by the mean of the CVwithin of all the individuals (Robisson et al. 1993). A PIC value greater than 1 means that the intra-individual variability is smaller than the inter-individual variability and so the given variable can be interpreted as individual-specific. In addition, the distribution of each variable per context and per individual was compared using a Kruskal-Wallis rank sum test.

Individual identity can be coded from a combination of variables. The set of acoustic variables was then compared per individual for each context using a multivariate analysis. For consistency, the same set of acoustic variables used in the analysis per behavioural context was used here. The RF was used to classify the acoustic structure of calls per individual within each context of emission.

All those analyses were conducted per context, since the individual signature may vary depending on the context and function of different types of calls. In order to assess whether potential bias resulting from the use of different acoustic units, we also conducted the analyses per context and per acoustic unit. The results obtained were similar and more importantly they led to the same conclusion regarding the individual signature in the different types of calls (results not shown).

Classification procedure

The RF was chosen because it does not require assumptions on the distribution of predictor variables, and for its capacity to deal with imbalanced data (when some classes, contexts in our case, are represented more than others). The RF is based on the classification and regression tree algorithm (Breiman et al. 1984) where a tree is constructed by recursively splitting the dataset into two subsets. RF adds two levels of randomization (Breiman 2001): several trees are grown from bootstrap samples of the dataset (random selection with replacement), and at each node a subset of variables is randomly selected to split the data. The accuracy of prediction is estimated intrinsically using the bootstrap process. For each tree grown, the class of data that were left out of the bootstrap sample (out-of-bag data) are predicted using the constructed tree, so their predicted classes are compared to their correct known classes. The accuracy is then calculated as a proportion of correct classification. In addition we used the indicator "precision" (Altman and Bland 1994) to calculate the number of correct prediction per class, based on the confusion matrix. This accuracy of prediction per class is compared to a prediction by chance, calculated as the number of occurrences in the class divided by the total number of occurrences for a given analysis. This evaluates how much the prediction by RF is better than a random allocation of class based on occurrences. The importance of variables can also be estimated intrinsically in the RF using the bootstrap process. For each tree grown, the values of the variable to be tested are randomly permuted and the classes of the out-of-bag data are predicted and compared to their correct known classes. The difference in accuracy of prediction using the original and the randomised values gives a measure of the importance of the given variable (the larger the difference is, the more important the variable was for prediction).

The RF was run in R software using the package "randomForest" (Liaw et al. 2014). The number of trees to be grown from bootstrap samples of the dataset (parameter "ntree") was set at 1000 for the analysis per contexts, and at 2000 for the analysis per individual. This ensures convergence of the results (Supplementary Figure 2, Supplementary Figure 3) as well as robustness in the measure of variable of importance (Genuer et al. 2008). To set the number of variables to be randomly selected at each node

(parameter "mtry") we used the default value proposed in the R package: the square root of the total number of variables, so three in our case. This setting provides the best results for datasets where the number of individuals is largely bigger than the number of variables (Genuer et al. 2008).

Our dataset is imbalanced, with the context TOF representing 2% of all the observations. We used a mix of over-sampling the minority class and down-sampling the majority classes to deal with the issue. The over-sampling increases the accuracy of prediction of the minority class, but also of the other classes since it allows to keep in more data when down-sampling. The down-sampling ensures that data from the minority class are present at each bootstrap, and thus increases accuracy of prediction for this class. We over-sampled the minority class TOF to twice its size using the package "unbalanced" in R (Dal Pozzolo et al. 2015). We then used the balanced random forest algorithm proposed by (Chen et al. 2004), which down-samples the major observational classes according to the occurrence of the minor class. For the analysis comparing the calls defined as WAT and in the air (FLY&DIV), the imbalance of the dataset was not as extreme with calls in the air representing 40% of the observations, so only the down-sampling technique was used (Chen et al. 2004). The analysis comparing the calls defined as FLY and DIV was not imbalanced. Finally for the analyses on individuality we used the down-sampling method (Chen et al. 2004).

Results

We recorded and analysed a total of 214 h 50 min and 43 s of sound data, from 18 foraging Cape gannets. No data could be retrieved from the audio recorders deployed on the other two individuals (the recorders could not connect to the laptop and software, possibly due to water damage). The number of call sequences recorded for each individual varied between seven and 276 (median 75), irrespective of the duration of the trips and recordings (Supplementary Table 1). In total, vocalisations were recorded on 1718 occasions (957 times when the birds were resting on the water, 30 during taking off, 412 when flying and 319 just before diving). For the latter, 75 % of the calls were emitted 1 s before the birds hit the water (maximum 3.4 s). Call sequences were made up of one to 23 calls emitted in close succession (2.3 ± 1.8). Acoustic variables were measured on a total of 1348 of those calls for which recordings were of good quality (776 when the birds were resting on the water, 32 while taking off, 269 when flying and 271 just before diving).

Potential instrument effects

All but one equipped bird gained weight (mean 267 g, range 0 g - 550 g, mean mass at deployment 2595 g, n=17 equipped birds with retrieved acoustic data, mass not available for one equipped bird) during their foraging trip. Knowing that all birds had time to feed their chick before we captured them for retrieval, those measures can be considered as a low estimate of the real mass gain of the birds. We can thus conclude the deployed birds were able to find food despite being fitted with devices.

The duration of trips undertaken by deployed birds fitted with devices lasted 13.2 ± 6.6 h, ranging from 4 to 25 h (n=18 birds). There was no significant difference between those trips when carrying the devices and the duration of following trips from the same birds without devices (Chi2=0.72, p=0.40). During the whole study period, the duration of trips of deployed birds was very similar to that of control birds: 11.6 ± 2.7 h on average (n=18 birds, 3 to 15 trip durations averaged per bird) vs 11.3 ± 2.7 h on average (n=20 birds, 10 to 19 trip durations averaged per bird, Supplementary Figure 3). The duration of trips was not significantly explained by the categorization of birds as control, deployed or partner of deployed (Chi2=2.05, p=0.36).

Acoustic structure in relation to behavioural context

The distribution of all the acoustic variables per context significantly differed in median and most of them also differed in variance (Table 1).

When including all the four contexts, the RF showed a global accuracy of prediction of 75 %. The indicators of precision per context showed that 88 %, 62 %, 56 % and 66 % of the data predicted as WAT, TOF, FLY or DIV, respectively, were correct. These predictions were 1.5, 31, 2.8 and 3.3 times better than a prediction by chance (Table 2). To discriminate among the four different calls, the temporal variables appeared to be the most important ones. Both the duration of the sequence followed by the duration of the call were clearly ranked before the other variables in the RF measure of variable importance (Figure 2a). Call sequences FLY were the shortest ones (mean 0.47 s), composed mostly of a single call (57% of call sequences) of average duration (mean 0.09 s, Table 1). The calls WAT were slightly shorter (mean 0.07 s) but the sequences emitted were more often composed of more than one call (43 % of single calls, 23 % of double calls) so that their average duration was longer (mean 0.66 s, Table 1). Calls DIV were composed of one, two or three calls in similar proportions (23%, 29%, 24% of call sequences respectively). Those calls were the longest ones (mean 0.12 s), often composed of two

parts (Figure 1). Knowing that only the first part of the call was measured, when relevant, the calls emitted before diving were clearly longer than the calls emitted in all other contexts. The calls TOF were the shortest ones (mean 0.05 s), with more calls emitted in series (median 5 calls per sequence) making them the longest sequences (mean 1.48 s, Table 1).

When comparing the calls on the water (WAT) and in the air (FLY&DIV), the RF showed a global accuracy of prediction of 86 %. The indicator precision shows that 88 % and 83 % of the data predicted as WAT and FLY&DIV, respectively, were correctly classified. These predictions were 1.5 and 2 times better than a prediction by chance (Table 2). Frequency and spectral variables appeared to be the most important variables to discriminate between calls emitted in those two different environments, with the percentage of energy emitted below 1200 Hz and the fundamental frequency of calls ranked first in the measure of variable importance (Figure 2b). Calls WAT were in general lower in frequency than calls emitted in all other contexts (Figure 1), displaying both the lowest F0 and Fmax values and the highest E1200 values accordingly (Table 1).

Classifying the calls FLY and DIV, the RF showed a global accuracy of prediction of 79 %, with 77 % and 81 % of correct classification for the two classes, respectively. These predictions per class were 1.5 and 1.6 times better than a prediction by chance (Table 2). The duration of the sequence was the most important variable to discriminate between those two contexts (Figure 2c), with the sequences DIV longer than the FLY ones (mean 1.00 s vs 0.47 s respectively, Table 1).

Individual signature

The PIC of the acoustic variables was studied on 767 calls from 14 individuals for the context on water, on 256 calls from 11 individuals for the context flying, and on 262 calls from 13 individuals for the context before diving. The values of PIC were low in all behavioural contexts, most of them being around 1 (i.e. the variability per individual is similar to the variability of the population) and only a few slightly higher than 1.1 (Table 3). The F0, the Nparts and the Ncalls in particular were identified as potential cues to discriminate between individuals in all contexts. In addition, the distribution of energy contributed to individual signatures of calls WAT and FLY, with the variables Q75 and E1200, respectively, yielded a PIC higher than 1.1 (Table 3). Almost all distributions of acoustic variables per individual and per context were significantly different according to the Kruskal-Wallis rank sum test.

Since a call is a single unit from which we measured different variables, the potential for individual discrimination is stronger (and more realistic) when considering a combination of all acoustic variables. The RF was able to predict the emitter of a call with an accuracy of 41 %, 63 % and 59 % for calls WAT, FLY and DIV respectively. The percentage of correct classification varied greatly per context and per individual, with values of the indicator precision ranging between 0 and 83 % depending on the individual (Table 4). Individuals for which a larger number of calls were recorded and measured were in general showing higher percentages of correct classification than individuals for which less calls were available. Those values were nonetheless higher than a prediction by chance in all but one case (Table 4), with improvement of prediction ranging between 2.3 and 22.3 depending on individuals and contexts. The F0 was the most important variable to discriminate individuals in contexts FLY and DIV, while the distribution of energy (variables Q75 and E1200) was of most importance in the context WAT (Figure 3).

Discussion

This paper demonstrates the vocal repertoire of a seabird during its foraging trip at sea. This is, to our knowledge, the first study to provide a detailed analysis of the acoustic signals by a bird while at sea, based on data collected from bird-borne audio recording devices. We showed that the acoustic structure of the calls differed according to the behavioural context of emission, namely when the birds were resting on the water, at the moment of taking off, when they were flying, or a few seconds before hitting the water and diving to capture prey. Calls that are structurally different must be conveying different information, suggesting that they would be used for different biological functions. Furthermore, we found that each call type contains sufficient information to distinguish among individuals, although a combination of acoustic features is needed for improved inter-individual discrimination.

The different types of calls emitted by Cape gannets when at sea can be differentiated by both temporal and spectral variables. The calls emitted when taking off were the rarest ones recorded, yet they contained on average the highest number of calls in a sequence (Table 1). The calls emitted before diving were the longest ones, as if the signal was elongated in this context. The repetition or increased duration of a signal increases the chances for the information to be received, which can be crucial when there is a risk of collision between birds (Thiebault et al. 2016). In contrast, spectral variables can be explained by differences in anatomy according to the source-filter theory (Fant 1960, Fitch 1999). The different body posture of the birds when flying or preparing to dive compared to when they are on the water could

explain our results on the comparison of calls emitted in the two environments. The vocal production mechanisms in birds is poorly understood compared to those in terrestrial mammals. It is known that birds can adjust their vocal filter by varying their tracheal length, constricting the syrinx or changing the position of their neck or beak (Nowicki 1987, Riede and Suthers 2009). All these modifications correspond to changes in head movements of the bird. When flying and even more so before diving, the birds stretch their neck which elongates the trachea but reduces its diameter (lumen), hence increasing air flow and pressure in the trachea (Daley and Goller 2004). The increase in air flow could result in the production of higher frequencies. This could thus explain our results showing vocalisations at higher frequencies when the birds are flying or diving compared to when they are on the water (Figure 1, Table 1).

During their foraging trips, the Cape gannets engage in different activities. Our results suggest that they make use of vocal communication for different purposes in each of those contexts. When forming rafts on the water, calls could be used for individuals to maintain contact and group cohesion. Indeed, the fact this context was associated with the lowest F0 and Fmax of all calls emitted by those birds at sea (Table 1) makes them the ones that will transmit over the longest distance. Calls on the water could also intervene during agonistic encounters in a high density group and avoid individuals colliding with each other (Thiebault et al. 2016). When flying, the birds emitted the shortest sequences recorded suggesting that they were involved in a brief interaction between birds. At the colony, calls in flight are always involved in a near-collision event (pers. obs.), birds "honking" at each other (Thiebault et al. 2016). Before diving, calls could also be used to avoid collision. Those calls were the longest ones and the ones emitted at the highest frequency (Table 1). Those acoustic features make the emitter easier to locate. In this context, a collision can be fatal when birds are diving at high speed (Lee and Reddish 1981, Machovsky-Capuska et al. 2011) so the localisation of the emitter of the call could be of higher importance. Since not all dives were preceded by a call, the immediate surrounding of a bird when about to dive could explain the need to produce a call (when crowded) or not (when scattered). Furthermore, since gannets benefit from previous attacks to disturb the fish school and capture prey (Thiebault et al. 2015), a signal informing that a gannet is going to dive may be used by surrounding birds to synchronize their attack and increase their chance of capturing a prey. Gannets initiating a dive are easy to identify from visual observation: they turn upside down and elongate their body shape. But in a large busy flock, acoustic communication may be more efficient than visual communication. This second potential function of the calls emitted before

diving might be an add-on value of a warning signal and could potentially be used differently by individuals depending on theirexperience. Indeed, adult seabirds commonly show higher foraging performances than juveniles (Porter and Sealy 1982, Bertellotti and Yorio 2000, Riotte-Lambert and Weimerskirch 2013), suggesting that they acquire knowledge and skills over consecutive experiences. Given the long life span of the Cape gannets (oldest ringed bird over 30 years old, The South African Bird Ringing Unit, http://safring.adu.org.za/safring_stories.php), we may suppose that breeding birds continue to acquire experience and improve their foraging skills during adulthood (first breeding at three to four years old, (Nelson 2005)). As a consequence the potential use of a before-diving warning signal by a gannet to dive immediately after and increase its chances to capture prey might depend on its experience with group foraging.

Within each context, calls were specific to individuals which can be explained by differences in anatomy based on the source-filter theory (Fant 1960, Fitch 1999). Different individuals may have vocal tracts of slightly different sizes including their syrinx, the "source" from which they emit calls (Greenewalt 1968, Suthers and Zollinger 2008, Riede and Goller 2010, Fitch and Suthers 2016). Accordingly, the F0 was the variable with the highest PIC value in contexts FLY and DIV (Table 2) and the variable of most importance to correctly predict the emitter of the call with the RF (Figure 3). Nonetheless, the PIC calculated per variable were relatively low (maximum of 1.52 for the F0) compared to other seabird species in a breeding context (no other study conducted at sea to compare with): e.g. 4.7 and 4.9 for the F0 in the black-headed gull *Larus ridibundus* and the slender-billed gull *L. genei* respectively (Mathevon et al. 2003), 1.88 for the F0 in the kittiwake *Rissa tridactyla* (Aubin et al. 2007), up to 3.84 and 4.49 for measures on duration of signals and silences in the Yelkouan shearwater *Puffinus yelkouan* and the Cory's shearwater *Calonectris diomedea diomedea* (Curé et al. 2009), or 1.85 for a measure on amplitude modulation in the African penguin *Spheniscus demersus* (Favaro et al. 2015). A combination of variables would then be necessary for individual Cape gannets to identify each other at sea, although still with a risk of confusion. While individual recognition at the nest is of major importance in a breeding context, it might not be crucial for seabirds when foraging at sea, at least as it is suggested by our results. The Cape gannet colony breeding at Bird Island (Aloga Bay) numbers over 90 000 breeding pairs (Department of Environmental Affairs, unpublished data), and individuals leave to forage at various unsynchronised times (Rishworth et al. 2014b). As a consequence, there is little evidence of individuals to repeatedly

forage with the same conspecifics. More probably, birds will opportunistically meet at sea to forage in groups made of individuals present at that time.

Our understanding of acoustic communication among seabirds when at sea is still poor but we have in this study made some major advances. Better knowledge on the function of seabirds' calls when foraging is, however, necessary. This could be addressed using playback experiments for example, observing the behavioural response of birds to different signals (Beer 1970). Together with artificial modification of the acoustic signals (Aubin et al. 2000, Charrier et al. 2001b), this would provide insights into the acoustic communication of those species. Moreover, the marine environment will have to be considered with its own specificities (Larsen and Radford 2018), being wide and open and containing typical sounds from the wind and the waves. The adaptation of acoustic communication in such an environment will also have to be addressed to fully understand the use of this type of communication among other communication channels. Since many seabird species commonly form groups when at sea (i.e. Hoffman et al. 1981, Harrison et al. 1991, Hodges and Woehler 1993, Camphuysen and Webb 1999, Assali et al. 2017), the exploration of this field of research promises to hold much potential.

Acknowledgements

Authors' contributions

AT, TA and IC conceived the ideas and designed the methodology; PP designed fieldwork under a permit from SANParks; AT collected the data; AT analysed the data; AT, IC and TA interpreted the results; AT led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for submission.

Permits

Fieldwork was undertaken under a permit from SANParks: PISTP1238. All experimental protocols were approved by both SANParks and the Nelson Mandela University Ethics Committee (reference: A10-SCI-ZOO-008), and were carried out in accordance with the approved guidelines.

Data accessibility

Data will be published on the Dryad Digital Repository upon acceptance of this manuscript.

References

Altman, D. G. and Bland, J. M. 1994. Diagnostic tests. 1: Sensitivity and specificity. - BMJ 308: 1552.

Anisimov, V. N., Herbst, J. A., Abramchuk, A. N., Latanov, A. V., Hahnloser, R. H. R. and Vyssotski, A. L. 2014. Reconstruction of vocal interactions in a group of small songbirds. - Nature Methods 11(11): 1135-1137.

Assali, C., Bez, N. and Tremblay, Y. 2017. Seabird distribution patterns observed with fishing vessel's radar reveal previously undescribed sub-meso-scale clusters. - Scientific Reports 7: 7364.

Aubin, T. 1991. Why do distress calls evoke interspecific responses? An experimental study applied to some species of birds. - Behavioural Processes 23: 103–111.

Aubin, T. and Jouventin, P. 2002. How to vocally identify kin in a crowd: The penguin model. - In: Peter J. B. Slater, J. S. R., Charles T. Snowdon and Timothy J. Roper (ed), Advances in the Study of Behavior. Academic Press, pp. 243–277.

Aubin, T., Jouventin, P. and Hildebrand, C. 2000. Penguins use the two–voice system to recognize each other. - Proceedings of the Royal Society of London B: Biological Sciences 267: 1081–1087.

Aubin, T., Mathevon, N., Staszewski, V. and Boulinier, T. 2007. Acoustic communication in the Kittiwake Rissa tridactyla: potential cues for sexual and individual signatures in long calls. - Polar Biol 30: 1027–1033.

Ballard, G., Ainley, D. G., Ribic, C. A. and Barton, K. R. 2001. Effect of instrument attachment and other factors on foraging trip duration and nesting success of adélie penguins. - The Condor 103: 481–490.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H. and Dai, B. 2014. lme4 Linear mixed-effects models using Eigen and S4.

Beer, C. G. 1970. Individual recognition of voice in the social behaviour of birds. - In: Advances in the study of behaviour. Academic Press. Daniel S. Lehrman, Robert A. Hinde, Evelyn Shaw, pp. 27–74.

Berlincourt, M. and Arnould, J. P. Y. 2014. At-Sea Associations in Foraging Little Penguins. - Plos One 9: e105065.

Bertellotti, M. and Yorio, P. 2000. Age-related feeding behaviour and foraging efficiency in Kelp Gulls Larus dominicanus attending coastal trawlers in Argentina. - Ardea 88: 207–214.

Bost, C. A., Cotte, C., Bailleul, F., Cherel, Y., Charrassin, J. B., Guinet, C., Ainley, D. G. and Weimerskirch, H. 2009. The importance of oceanographic fronts to marine birds and mammals of the southern oceans. - Journal of Marine Systems 78: 363–376.

Bourgeois, K., Curé, C., Legrand, J., Gómez-Díaz, E., Vidal, E., Aubin, T. and Mathevon, N. 2007. Morphological versus acoustic analysis: what is the most efficient method for sexing yelkouan shearwaters Puffinus yelkouan? - J Ornithol 148: 261–269.

Boyd, I. L., Kato, A. and Ropert-Coudert, Y. 2004. Bio-logging science: sensing beyond the boundaries. - Mem Natl Inst Polar Res Spec 58: 1–14.

Breiman, L. 2001. Random Forests. - Machine Learning 45: 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. 1984. Classification and Regression Trees. - Chapman & Hall.

Camphuysen, C. J. and Webb, A. 1999. Multi-species feeding associations in North Sea seabirds: jointly exploiting a patchy environment. - Ardea 87: 177–198.

Carroll, G., Harcourt, R., Pitcher, B. J., Slip, D. and Jonsen, I. 2018. Recent prey capture experience and dynamic habitat quality mediate short-term foraging site fidelity in a seabird. - Proc. R. Soc. B 285: 20180788.

Ceugniet, M., Aubin, T., Bernard-Laurent, A. and Soyez, D. 1999. Vocal signatures of the rally call of red-legged and rock partridges and of their hybrids. - Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie 322: 887–895.

Charrier, I., Jouventin, P., Mathevon, N. and Aubin, T. 2001a. Individual identity coding depends on call type in the South Polar skua Catharacta maccormicki. - Polar Biol 24: 378–382.

Charrier, I., Mathevon, N., Jouventin, P. and Aubin, T. 2001b. Acoustic Communication in a Black-Headed Gull Colony: How Do Chicks Identify Their Parents? - Ethology 107: 961–974.

Chen, C., Liaw, A. and Breiman, L. 2004. Using random forest to learn imbalanced data. - University of California, Berkeley in press.

Choi, N., Kim, J.-H., Kokubun, N., Park, S., Chung, H. and Lee, W. Y. 2017. Group association and vocal behaviour during foraging trips in Gentoo penguins. - Scientific Reports 7: 7570.

Clua, E. and Grosvalet, F. 2001. Mixed-species feeding aggregation of dolphins, large tunas and seabirds in the Azores. - Aquatic Living Resources 14: 11–18.

Cooke, S. J., Hinch, S. G., Wikelski, M., Andrews, R. D., Kuchel, L. J., Wolcott, T. G. and Butler, P. J. 2004. Biotelemetry: a mechanistic approach to ecology. - Trends in Ecology & Evolution 19: 334-343.

Couchoux, C., Aubert, M., Garant, D. and Réale, D. 2015. Spying on small wildlife sounds using affordable collar-mounted miniature microphones: an innovative method to record individual daylong vocalisations in chipmunks. - Scientific Reports 5: 10118.

Croll, D. A., Osmek, S. D. and Bengtson, J. L. 1991. An Effect of Instrument Attachment on Foraging Trip Duration in Chinstrap Penguins. - The Condor 93: 777–779.

Curé, C., Aubin, T. and Mathevon, N. 2009. Acoustic convergence and divergence in two sympatric burrowing nocturnal seabirds. - Biological Journal of the Linnean Society 96: 115–134.

Curé, C., Mathevon, N., Mundry, R. and Aubin, T. 2012. Acoustic cues used for species recognition can differ between sexes and sibling species: evidence in shearwaters. - Animal Behaviour 84: 239–250.

Dal Pozzolo, A., Caelen, O. and Bontempi, G. 2015. R package "unbalanced."

Daley, M. and Goller, F. 2004. Tracheal length changes during zebra finch song and their possible role in upper vocal tract filtering. - J. Neurobiol. 59: 319–330.

Danchin, É. and Wagner, R. H. 1997. The evolution of coloniality: the emergence of new perspectives. - Trends Ecol Evol 12: 342–347.

Davoren, G. K., Montevecchi, W. A. and Anderson, J. T. 2003. Search Strategies of a Pursuit-Diving Marine Bird and the Persistence of Prey Patches. - Ecological Monographs 73: 463–481.

Dentressangle, F., Aubin, T. and Mathevon, N. 2012. Males use time whereas females prefer harmony: individual call recognition in the dimorphic blue-footed booby. - Animal Behaviour 84: 413–420.

Draganoiu, T. I., Nagle, L., Musseau, R. and Kreutzer, M. 2006. In a songbird, the black redstart, parents use acoustic cues to discriminate between their different fledglings. - Animal Behaviour 71: 1039–1046.

Elgar, M. A. 1986. House sparrows establish foraging flocks by giving chirrup calls if the resources are divisible. - Animal Behaviour 34: 169–174.

Fant, G. 1960. Acoustic theory of speech production. in press.

Favaro, L., Gamba, M., Alfieri, C., Pessani, D. and McElligott, A. G. 2015. Vocal individuality cues in the African penguin (*Spheniscus demersus*): a source-filter theory approach. - Scientific Reports 5: 17255.

Fitch, W. T. 1999. Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. - Journal of Zoology 248: 31–48.

Fitch, W. T. and Suthers, R. A. 2016. Vertebrate Vocal Production: An Introductory Overview. - In: Suthers, R. A. et al. (eds), Vertebrate Sound Production and Acoustic Communication. Springer Handbook of Auditory Research. Springer International Publishing, pp. 1–18.

Genuer, R., Poggi, J.-M. and Tuleau, C. 2008. Random Forests: some methodological insights. - arXiv:0811.3619 in press.

Gill, L. F., D'Amelio, P. B., Adreani, N. M., Sagunsky, H., Gahr, M. C. and Maat, A. 2016. A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. - Methods in Ecology and Evolution 7: 1349-1358.

Green, D. B., Klages, N. T. W., Crawford, R. J. M., Coetzee, J. C., Dyer, B. M., Rishworth, G. M. and Pistorius, P. A. 2014. Dietary change in Cape gannets reflects distributional and demographic shifts in two South African commercial fish stocks. - ICES J. Mar. Sci. in press.

Greenewalt, C. H. 1968. Bird song: Acoustics and physiology. - Smithsonian Institution Press.

Greif, S. and Yovel, Y. 2019. Using on-board sound recordings to infer behaviour of free-moving wild animals. - Journal of Experimental Biology 222: jeb184689.

Harrison, N. M., Whitehouse, M. J., Heinemann, D., Prince, P. A., Hunt, G. L. and Veit, R. R. 1991. Observations of Multispecies Seabird Flocks around South Georgia. - The Auk 108: 801–810.

Hodges, C. L. and Woehler, E. J. 1993. Associations between seabirds and cetaceans in the Australian sector of the Southern Indian Ocean. - Marine Ornithology 22: 205–212.

Hoffman, W., Heinemann, D. and Wiens, J. A. 1981. The Ecology of Seabird Feeding Flocks in Alaska. - The Auk 98: 437–456.

Ilany, A., Barocas, A., Kam, M., Ilany, T. and Geffen, E. 2013. The energy cost of singing in wild rock hyrax males: evidence for an index signal. - Animal Behaviour 85: 995–1001.

Kooyman, G. 2004. Genesis and Evolution of Bio-Logging Devices: 1963-2002. - Mem Nat Inst Polar Res 58: 15–22.

Kroodsman, D., Miller, E. and Ouellet, H. 1982. Acoustic Communication in Birds. - Academic Press.

Krull, C. R., Ranjard, L., Landers, T. J., Ismar, S. M. H., Matthews, J. L. and Hauber, M. E. 2012. Analyses of sex and individual differences in vocalizations of Australasian gannets using a dynamic time warping algorithm. - The Journal of the Acoustical Society of America 132: 1189–1198.

Kuhn, M. 2008. Building predictive models in R using the caret package. - Journal of Statistical Software 28: 1–26.

Larsen, O. N. and Radford, C. 2018. Acoustic Conditions Affecting Sound Communication in Air and Underwater. - In: Slabbekoorn, H. et al. (eds), Effects of Anthropogenic Noise on Animals. Springer Handbook of Auditory Research. Springer New York, pp. 109–144.

Lee, D. N. and Reddish, P. E. 1981. Plummeting gannets: a paradigm of ecological optics. - Nature 293: 293–294.

Lett, C., Semeria, M., Thiebault, A. and Tremblay, Y. 2014. Effects of successive predator attacks on prey aggregations. - Theor Ecol 7: 239–252.

Liaw, A., Wiener, M., Breiman, L. and Cutler, A. 2014. randomForest: Breiman and Cutler's random forests for classification and regression.

Machovsky-Capuska, G. E., Dwyer, S. L., Alley, M. R., Stockin, K. A. and Raubenheimer, D. 2011. Evidence for fatal collisions and kleptoparasitism while plunge-diving in Gannets. - Ibis 153: 631–635.

Mathevon, N., Charrier, I. and Jouventin, P. 2003. Potential for individual recognition in acoustic signals: a comparative study of two gulls with different nesting patterns. - Comptes Rendus Biologies 326: 329–337.

Mathevon, N., Aubin, T., Vielliard, J., Silva, M.-L. da, Sebe, F. and Boscolo, D. 2008. Singing in the Rain Forest: How a Tropical Bird Song Transfers Information. - PLOS ONE 3: e1580.

McInnes, A. M., McGeorge, C., Ginsberg, S., Pichegru, L. and Pistorius, P. A. 2017. Group foraging increases foraging efficiency in a piscivorous diver, the African penguin. - Royal Society Open Science 4: 170918.

Mulard, H., Aubin, T., White, J. F., Wagner, R. H. and Danchin, É. 2009. Voice variance may signify ongoing divergence among black-legged kittiwake populations. - Biological Journal of the Linnean Society 97: 289–297.

Mullers, R. H. E., Navarro, R. A., Daan, S., Tinbergen, J. M. and Meijer, H. A. J. 2009. Energetic costs of foraging in breeding Cape gannets Morus capensis. - Mar Ecol Prog Ser 393: 161–171.

Nelson, J. B. 2005. Pelicans, cormorants, and their relatives: The Pelecaniformes. - Oxford University Press, New York, New York.

Nowicki, S. 1987. Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. - Nature 325: 53–55.

Passos, C., Navarro, J., Giudici, A. and González-Solís, J. 2010. Effects of Extra Mass on the Pelagic Behavior of a Seabird - Efectos del Peso Suplementario en el Comportamiento Pelágico de un Ave Marina. - The Auk 127: 100–107.

Phillips, R. A., Xavier, J. C., Croxall, J. P. and Burger, A. E. 2003. Effects of satellite transmitters on albatrosses and petrels. - The Auk 120: 1082–1090.

Pichegru, L., Ryan, P. G., Lingen, C. D. van der, Coetzee, J., RopertCoudert, Y. and Grémillet, D. 2007. Foraging behaviour and energetics of Cape gannets Morus capensis feeding on live prey and fishery discards in the Benguela upwelling system. - Mar Ecol Prog Ser 350: 127–136.

Porter, J. M. and Sealy, S. G. 1982. Dynamics of Seabird Multispecies Feeding Flocks: Age-Related Feeding Behaviour. - Behaviour 81: 91–109.

R Core Team 2013. R : A Language and Environment for Statistical Computing.

Riede, T. and Suthers, R. A. 2009. Vocal tract motor patterns and resonance during constant frequency song: the white-throated sparrow. - J Comp Physiol A Neuroethol Sens Neural Behav Physiol 195: 183–192.

Riede, T. and Goller, F. 2010. Functional morphology of the sound-generating labia in the syrinx of two songbird species. - J Anat 216: 23–36.

Riotte-Lambert, L. and Weimerskirch, H. 2013. Do naive juvenile seabirds forage differently from adults? - Proc. R. Soc. B 280: 20131434.

Rishworth, G. M., Tremblay, Y., Green, D. B. and Pistorius, P. A. 2014a. An automated approach towards measuring time-activity budgets in colonial seabirds. - Methods Ecol Evol 5: 854–863.

Rishworth, G. M., Tremblay, Y., Green, D. B., Connan, M. and Pistorius, P. A. 2014b. Drivers of Time-Activity Budget Variability during Breeding in a Pelagic Seabird. - PLoS ONE 9: e116544.

Robisson, P., Aubin, T. and Bremond, J.-C. 1993. Individuality in the Voice of the Emperor Penguin Aptenodytes forsteri: Adaptation to a Noisy Environment. - Ethology 94: 279–290.

Ropert-Coudert, Y. and Wilson, R. P. 2005. Trends and perspectives in animal-attached remote sensing. - Front Ecol Environ 3: 437–444.

Russell, R. W., Hunt, G. L., Coyle, K. O. and Cooney, R. T. 1992. Foraging in a fractal environment: Spatial patterns in a marine predator-prey system. - Landscape Ecol 7: 195–209.

Ryan, P. G., Edwards, L. and Pichegru, L. 2012. African Penguins Spheniscus demersus, Bait Balls and the Allee Effect. - Ardea 100: 89–94.

Scales, K. L., Miller, P. I., Hawkes, L. A., Ingram, S. N., Sims, D. W. and Votier, S. C. 2014. On the Front Line: frontal zones as priority at-sea conservation areas for mobile marine vertebrates. - J Appl Ecol 51: 1575–1583.

Schneider, D. 1982. Fronts and seabird aggregations in the southeastern Bering Sea. - Marine ecology progress series. Oldendorf 10: 101–103.

Siegfried, W. R., Frost, P. G. H., Kinahan, J. B. and Cooper, J. 1975. Social behaviour of Jackass penguins at sea. - African Zoology 10: 86–100.

Silverman, E. D., Veit, R. R. and Nevitt, G. 2004. Nearest neighbors as foraging cues: information transfer in a patchy environment. - Mar Ecol Prog Ser 277: 25–36.

Sokal, R. and Rohlf, F. 1995. Biometry : the principles and practice of statistics in biological research.

Stoddard, P. K. 1988. The bugs call of the cliff swallow: A rare food signal in a colonially nesting bird species. - Condor 90: 714–715.

Stowell, D., Benetos, E. and Gill, L. F. 2017. On-Bird Sound Recordings: Automatic Acoustic Recognition of Activities and Contexts. - IEEE/ACM Transactions on Audio, Speech, and Language Processing 25: 1193–1206.

Strobl, C., Boulesteix, A.-L., Zeileis, A. and Hothorn, T. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. - BMC Bioinformatics 8: 25.

Sueur, J., Aubin, T. and Simonis, C. 2008. Seewave : a free modular tool for sound analysis and synthesis.

Suthers, R. A. and Zollinger, S. A. 2008. From brain to song: The vocal organ and vocal tract. - In: Neuroscience of birdsong. H. P. Zeigler & P. Marler. Cambridge, UK: Cambridge University Press, pp. 78–98.

Thiebault, A., Mullers, R., Pistorius, P., Meza-Torres, M. A., Dubroca, L., Green, D. and Tremblay, Y. 2014a. From colony to first patch: Processes of prey searching and social information in Cape Gannets. - Auk 131: 595–609.

Thiebault, A., Mullers, R. H. E., Pistorius, P. A. and Tremblay, Y. 2014b. Local enhancement in a seabird: reaction distances and foraging consequence of predator aggregations. - Behavioral Ecology 25: 1302–1310.

Thiebault, A., Semeria, M., Lett, C. and Tremblay, Y. 2015. How to capture fish in a school? Effect of successive predator attacks on seabird feeding success. - J Anim Ecol 85: 157–167.

Thiebault, A., Pistorius, P., Mullers, R. and Tremblay, Y. 2016. Seabird acoustic communication at sea: a new perspective using bio-logging devices. - Scientific Reports 6: srep30972.

Tremblay, Y., Thiebault, A., Mullers, R. and Pistorius, P. 2014. Bird-Borne Video-Cameras Show That Seabird Movement Patterns Relate to Previously Unrevealed Proximate Environment, Not Prey. - Plos One 9: e88424.

Vandenabeele, S. P., Shepard, E. L., Grogan, A. and Wilson, R. P. 2012. When three per cent may not be three per cent; device-equipped seabirds experience variable flight constraints. - Marine Biology 159: 1–14.

Weimerskirch, H. 2007. Are seabirds foraging for unpredictable resources? - Deep-Sea Res Pt II 54: 211–223.

White, S. J. 1971. Selective responsiveness by the gannet (Sula bassana) to played-back class. - Animal Behaviour 19: 125–131.

White, S. J. and White, R. E. C. 1970. Individual Voice Production in Gannets. - Behaviour 37: 40–54.

White, S. J., White, R. E. C. and Thorpe, W. H. 1970. Acoustic Basis for Individual Recognition by Voice in the Gannet. - Nature 225: 1156–1158.

Wilson, R. P., Grant, W. S. and Duffy, D. C. 1986. Recording Devices on Free-Ranging Marine Animals: Does Measurement Affect Foraging Performance? - Ecology 67: 1091–1093.

Figure Legends

Figure 1. Illustration of calls emitted by two different individuals of Cape gannets in the four behavioural contexts: (a,e) WAT = on water, (b,f) TOF = taking off, (c,g) FLY = flying, (d,h) DIV = before diving. Figure designed using the 'seewave' package (Sueur et al. 2008) in R software.
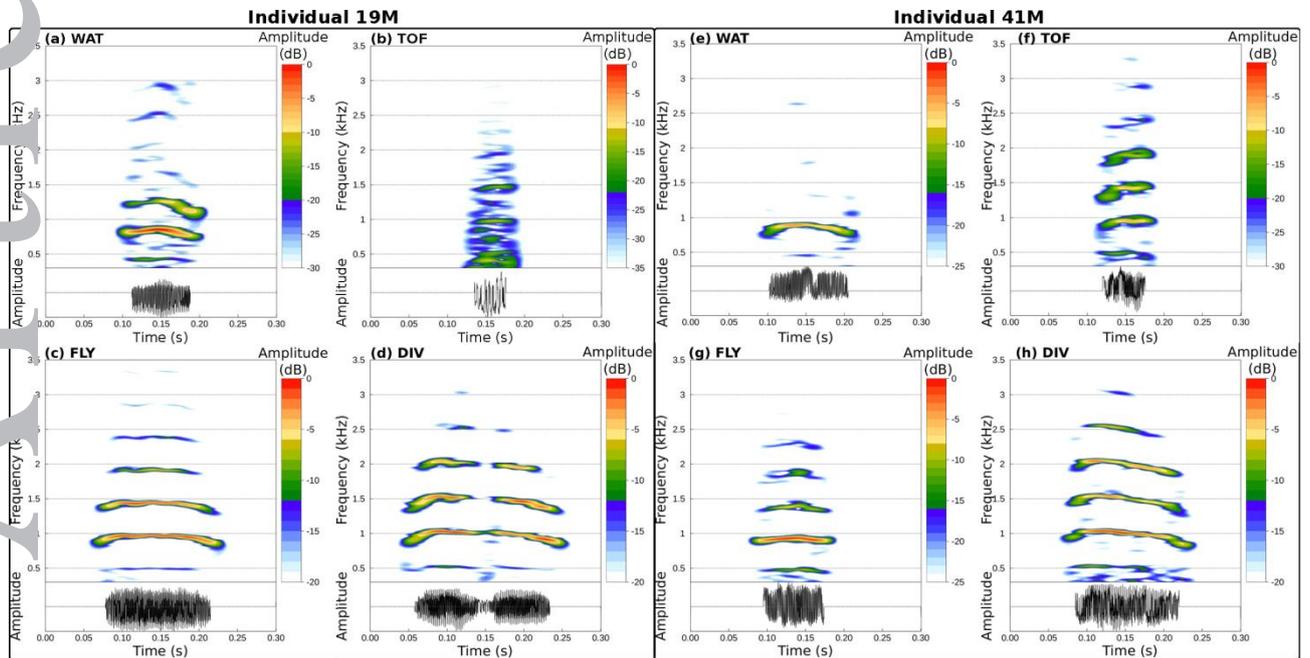
Figure 2. Ranking of importance of acoustic variables (top to bottom) calculated as a mean decrease in accuracy in the random forest algorithm, for the three analyses conducted on behavioural contexts: (a) comparing calls from all the four behavioural contexts observed, (b) comparing calls on the water and in the air, (c) comparing calls in the air. Behavioural contexts of call emission: WAT = on water, TOF = taking off, FLY = flying, DIV = before diving. Acoustic variables: DurSeq = duration of the sequence (s), DurCall = duration of the call measured (s), F0 = fundamental frequency (Hz), Fmax = frequency of maximum amplitude (Hz), Q75 = third quartile of the energy distribution (Hz), E1200 = percentage of energy occurring below 1200 Hz (%), SlopeAsc = ascending slope of the call (Hz.s^{-1}), SlopeDesc = descending slope of the call (Hz.s^{-1}), Nparts = number of parts composing the call.
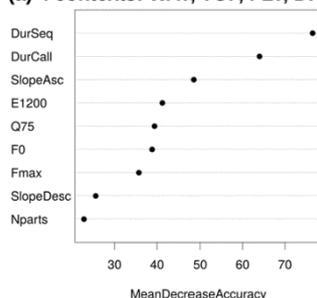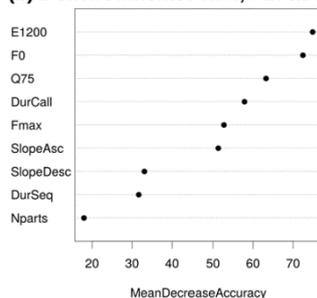
Figure 3. Ranking of importance of acoustic variables (top to bottom) calculated as a mean decrease in accuracy in the random forest algorithm, for the analyses conducted on individual signature for three behavioural contexts: (a) on calls emitted on water, (b) on calls emitted when flying (c) on calls emitted just before diving. Acoustic variables: DurSeq = duration of the sequence (s), DurCall = duration of the call measured (s), F0 = fundamental frequency (Hz), Fmax = frequency of maximum amplitude (Hz), Q75 = third quartile of the energy distribution (Hz), E1200 = percentage of energy occurring below 1200 Hz (%), SlopeAsc = ascending slope of the call (Hz.s^{-1}), SlopeDesc = descending slope of the call (Hz.s^{-1}), Nparts = number of parts composing the call.

(a) Context WAT.

(b) Context FLY.

(c) Context DIV.
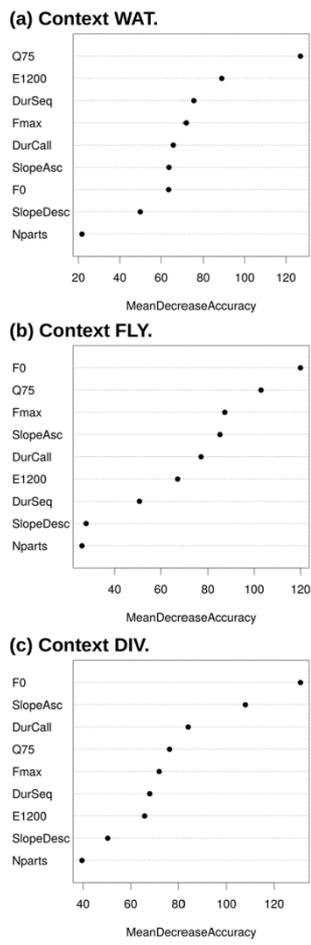
Table Legends

Table 1. Summary of distribution of acoustic variables measured on calls per behavioural context: WAT = on water, TOF = taking off, FLY = flying, DIV = before diving. Difference in variance of distribution per context was evaluated using the Fligner-Killeen test of homogeneity of variance. Difference in median of distribution per context was evaluated using the Kruskal-Wallis rank sum test. Acoustic variables: DurSeq = duration of the sequence (s), DurCall = duration of the call measured (s), Rhythm = number of calls divided by duration of the sequence (s^{-1}), F0 = fundamental frequency (Hz), Fmax = frequency of maximum amplitude (Hz), Q50 = second quartile of the energy distribution (Hz), Q75 = third quartile of the energy distribution (Hz), E1200 = percentage of energy occurring below 1200 Hz (%), SlopeAsc = ascending slope of the call ($Hz.s^{-1}$), SlopeDesc = descending slope of the call ($Hz.s^{-1}$), Ncalls = number of calls in the sequence, Nparts = number of parts composing the call. N = number of calls, SD = standard deviation, NS = non-significant.

Acoustic variables	WAT			TOF			FLY			DIV			Fligner-Killeen	Kruskal-Wallis
	N	Mean ± SD	Range	N	Mean ± SD	Range	N	Mean ± SD	Range	N	Mean ± SD	Range	p-value	p-value
DurSeq (s)	776	0.66±0.66	[0.04 7.12]	32	1.48±1.06	[0.06 3.25]	269	0.47±0.80	[0.04 4.96]	271	1.00±0.80	[0.07 4.18]	<0.001	<0.001
DurCall (s)	776	0.07±0.02	[0.02 0.24]	32	0.05±0.01	[0.04 0.08]	269	0.09±0.04	[0.03 0.24]	271	0.12±0.05	[0.04 0.27]	<0.001	<0.001
Rhythm (s-1)	599	5.6±1.0	[1.2 8.6]	30	5.3±1.1	[3.8 7.5]	139	5.1±1.2	[3.2 9.1]	238	3.9± 1.1	[1.6 7.1]	NS	<0.001
F0 (Hz)	776	388±58	[211 650]	32	418±53	[311 510]	269	446±53	[263 546]	271	454±45	[295 554]	<0.05	<0.001
Fmax (Hz)	776	843±286	[360 2140]	32	903±312	[320 2000]	269	1035±273	[457 2070]	271	1008±284	[433 2210]	NS	<0.001
Q50 (Hz)	776	1101±231	[600 1910]	32	1197±317	[720 2020]	269	1339±184	[820 2000]	271	1350±181	[826 2160]	<0.001	<0.001
Q75 (Hz)	776	1754±318	[1030 2830]	32	2094±390	[1300 2940]	269	2083±295	[1370 3090]	271	2062±279	[1458 2917]	<0.01	<0.001

E1200 (%)	776	54±14	[15 83]	32	47±15	[23 74]	269	42±9	[21 67]	271	43±9	[15 69]	<0.001	<0.001
SlopeAsc (Hz.s-1)	776	20.8±20.1	[-35.0 115.0]	32	33.6±26.7	[-10.0 105.0]	269	31.7±25.1	[-35.0 104.9]	271	39.2±27.1	[-25.0 114.9]	<0.001	<0.001
SlopeDesc (Hz.s-1)	776	-23.4±28.3	[-140.0 70.0]	32	-9.6±19.3	[-57.0 20.0]	269	-35.4±36.9	[-200.1 59.9]	271	-49.1±37.8	[-165.1 60.0]	<0.001	<0.001
Ncalls	776	3.5±2.7	[1 23]	32	7.2±4.3	[1 15]	269	2.3±3.0	[1 19]	271	3.4±2.1	[1 13]	<0.001	<0.001
Nparts	776	1.2±0.4	[1 3]	32	1.1±0.2	[1 2]	269	1.2±0.4	[1 2]	271	1.3±0.5	[1 2]	<0.001	<0.001

Table 2. Comparison of the accuracy of prediction of the behavioural context of emission of calls, obtained by chance (number of calls for the given context divided by total number of calls in the given analysis) or using the random forest algorithm (indicator precision). Behavioural contexts of call emission: WAT = on water, TOF = taking off, FLY = flying, DIV = before diving.

Analysis	Context	Number of calls	Prediction (%)		Improvement of prediction with RF
			by chance	with RF	
4 contexts	WAT	776	58	88	1.5
	TOF	32	2	62	31.0
	FLY	269	20	56	2.8
	DIV	271	20	66	3.3
2 environments	WAT	776	59	88	1.5
	FLY&DIV	540	41	83	2.0
2 aerial contexts	FLY	269	50	77	1.5
	DIV	271	50	81	1.6

Table 3. Potential of individual coding (PIC) calculated for each acoustic variable and per behavioural context: WAT = on water, FLY = flying, DIV = before diving. Difference in median of distribution per context and per individual was evaluated using the Kruskal-Wallis rank sum test. Acoustic variables: DurSeq = duration of the sequence (s), DurCall = duration of the call measured (s), Rhythm = number of calls divided by duration of the sequence (s^{-1}), F0 = fundamental frequency (Hz), Fmax = frequency of maximum amplitude (Hz), Q50 = second quartile of the energy distribution (Hz), Q75 = third quartile of the energy distribution (Hz), E1200 = percentage of energy occurring below 1200 Hz (%), SlopeAsc = ascending slope of the call ($Hz.s^{-1}$), SlopeDesc = descending slope of the call ($Hz.s^{-1}$), Ncalls = number of calls in the sequence, Nparts = number of parts composing the call. Ni = number of individuals, Nc = number of calls. NS = non-significant.

Acoustic variables	WAT (Ni = 14; Nc = 767)		FLY (Ni = 11; Nc = 256)		DIV (Ni = 13; Nc = 262)	
	PIC	Kruskal-Wallis p-value	PIC	Kruskal-Wallis p-value	PIC	Kruskal-Wallis p-value
DurSeq	1.03	<0.001	1.00	<0.001	1.06	<0.001
DurCall	1.03	<0.001	1.01	<0.001	0.98	<0.001
F0	1.14	<0.001	1.27	<0.001	1.52	<0.001
Fmax	0.86	<0.001	1.06	<0.05	1.08	<0.001
Q50	1.08	<0.001	1.02	<0.001	1.05	<0.001
Q75	1.27	<0.001	0.90	<0.001	1.00	<0.001
E1200	0.92	<0.001	1.14	<0.001	1.09	<0.001
SlopeAsc	0.89	NS	0.99	<0.05	1.02	<0.001
SlopeDesc	1.01	<0.001	0.92	<0.01	0.92	<0.001
Ncalls	1.17	<0.001	1.10	<0.001	1.30	<0.001
Nparts	1.42	<0.001	1.19	<0.001	1.35	<0.001

Table 4. Comparison of the accuracy of prediction of the emitter of a call (individual) per behavioural context, obtained by chance (number of calls for the given context divided by total number of calls in the given analysis) or using the random forest algorithm (indicator precision). Behavioural contexts of call emission: WAT = on water, TOF = taking off, FLY = flying, DIV = before diving.

Context	Individual	Number of calls	Prediction (%)		Improvement of prediction with RF
			by chance	with RF	
WAT	17M	136	18	58	3.2
	19M	12	2	14	7.0
	24M	11	1	21	21.0
	34M	29	4	33	8.3
	38M	93	12	57	4.8
	39M	17	2	16	8.0
	41M	17	2	23	11.5
	42M	86	11	55	5.0
	43M	11	1	7	7.0
	45M	129	17	53	3.1
	46M	62	8	48	6.0
	47M	8	1	0	0.0
	48M	69	9	51	5.7
	49M	87	11	25	2.3
FLY	19M	47	18	67	3.7
	38M	31	12	41	3.4
	39M	27	11	76	6.9
	41M	26	10	39	3.9
	42M	17	7	47	6.7
	43M	20	8	47	5.9
	45M	14	5	38	7.6
	46M	28	11	71	6.5
	47M	8	3	50	16.7
	48M	29	11	70	6.4
	49M	9	4	67	16.8
DIV	17M	8	3	67	22.3
	19M	47	18	83	4.6
	38M	55	21	83	4.0
	39M	21	8	66	8.3
	40M	10	4	55	13.8
	41M	14	5	43	8.6
	42M	17	6	58	9.7
	43M	8	3	18	6.0
	45M	14	5	50	10.0
	46M	10	4	43	10.8
	47M	9	3	50	16.7
	48M	32	12	76	6.3
	49M	17	6	50	8.3