Universidade de Lisboa
Faculdade de Ciências
Departamento de Informática

Automatic rcognition of vocalizations in the Lusitanian toadfish (Halobatrachus didactylus): acoustic rhythms and interactions

Manuel Vieira

Dissertação
Mestrado em Bioinformática e Biologia Computacional
Especialização: Biologia Computacional

2013

Universidade de Lisboa
Faculdade de Ciências
Departamento de Informática

Automatic rcognition of vocalizations in the Lusitanian toadfish (Halobatrachus didactylus): acoustic rhythms and interactions

Manuel Vieira

Dissertação
Mestrado em Bioinformática e Biologia Computacional
Especialização: Biologia Computacional

Paulo J. Fonseca e Carlos Teixeira

2013

ACKNOWLEDGEMENTS

My sincerest thanks to Paulo Fonseca, Clara Amorim and Carlos Teixeira for their continuous support, for always showing a contagious enthusiasm and for teaching me what I needed to develop this thesis (and more!).

Furthermore, I especially appreciate the opportunity given by Paulo Fonseca and Clara Amorim to be present and to participate in the IBAC Brasil conference where I learned much more about the world of bioacoustics.

I am equally thankful for being made part of the work group *Behavioural Physiology and Communication*, where I could be comfortable and be intellectually stimulated. Additionally, I am very thankful to Joana Vincente and Daniel Alves, who were always ready to help; and also to the various people who assisted me in the collection of all the data.

Last but not least, I am deeply thankful for the support of my brother and sister and all of my friends during the execution of this thesis and all other times of my life.

RESUMO

Muitos peixes comunicam através de vocalizações, sendo crucial para o sucesso reprodutor dessas espécies. Para estudar comunicação acústica é necessário discriminar as vocalizações que ocorrem ao longo do tempo, e para alem disso é essencial saber que individuo produz cada som. Este processo em experiencias controladas de curta duração é geralmente moroso e irrepetível, tornando a analise de gravações continuas de vários meses no habitat natural impossível de realizar sem um programa que automatize o processo. Aqui apresenta-se a utilização dos modelos escondidos de Markov que permitem a detecção e reconhecimento de vocalizações de peixes. Usando as vocalizações do xarroco (*Halobatrachus didactylus*) os nossos resultados mostram um sistema capaz de reconhecer eficientemente as vocalizações de acasalamento (sirenes; > 95 %), mas também discriminar os indivíduos através das vocalizações produzidas (> 90 %). Este método para reconhecimento de vocalizações produzidas pelo xarroco foi utilizado para monitorizar a actividade acústica deste peixe no intertidal e no infralitoral. Investigou-se a relação da actividade acústica com a presença de luz, e as flutuações da temperatura e da altura de maré. Observou-se que a actividade na zona intertidal foi muito sujeita ao efeito das marés, não existindo influencia aparente de ciclos circadianos. No infralitoral não se observa influencia das marés, e observa-se um ciclo circadiano negativamente correlacionado com a variação da temperatura com ainda um pico de actividade ao amanhecer. Nos períodos de maior actividade estudou-se as interacções acústicas entre machos e observou-se que geralmente eles evitam sobrepor vocalizações, excepto em alguns períodos mais curtos onde um individuo parece tentar sobrepor entre 3 a 6 vocalizações doutro indivíduo.

ABSTRACT

The study of acoustic communication often requires not only the recognition of species specific acoustic signals but also the identification of individual subjects, all this in a complex acoustic background. Moreover, when very long recordings are to be used, automatic methods can be invaluable tools to extract the relevant biological information. Here we present an automatic pattern recognition methodology based on the Hidden Markov Model that allows accurate detection and recognition of fish vocalizations. Our results show a relevant performance of the system that is capable not only to detect the mating vocalizations (boatwhistles) but also allows individual identification of male Lusitanian toadfish (*Halobatrachus didactylus*). The system was applied to a stream of round-the-clock recordings to monitor toadfish acoustic activity in two locations in a natural habitat in the Tagus estuary: intertidal nests only exposed in spring tides and infralitoral areas (minimum 2.5 m water level). We investigated the relation of the vocalizations' pattern with habitat parameters such as light, water level and temperature fluctuations. The vocalization patterns differed in the intertidal relative to infralitoral ones. In the periods with higher concentration of vocal activity we observed signal interactions where males in a chorus either alternated or synchronized their vocalizations (boatwhistles).

ÍNDICE

INTRODUCTION

The ability to receive information sent by others individuals and then react accordantly, i.e. communication, is one of the most important features to the prevalence of organisms (Hauser, 1997; Bradbury and Vehrencamp, 1998). This is particularly important in individuals with sexual reproduction, seeing as in such cases only with communication can the individuals reproduce (Bradbury and Vehrencamp, 1998).

The information transmitted in the process of communication can be sent through 1) visual signals like the colour shifts in the tilapia skins (Volpato *et al.*, 2004) or the bee dances (Dornhaus and Chittka, 1999), 2) acoustic signals which are most common between anurans, birds and mammals (Bradbury and Vehrencamp, 1998), or 3) even by chemicals signals which are used by almost all organisms (Johansson and Jones, 2007).

The use of acoustic signals is a frequently used form of communication between humans. As a consequence there are several methods to characterize sounds and methods of automatic speech recognition. In bioacoustics (the study of animals vocalizations/sounds), the adaptation of such methods is still a novelty, although is recognized that some procedures are crucial to this field such as the automatic recognition systems to analyse long term recordings.

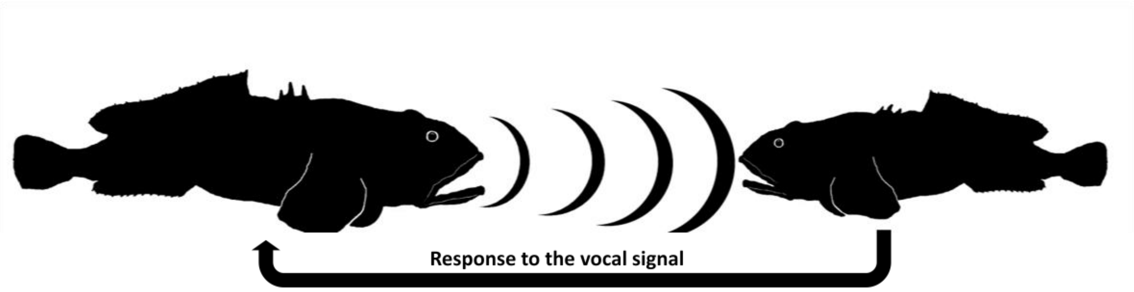

Figure 1 In the Lusitanian toadfish the acoustic communication is very important and is determinant of male reproductive success (R. O. Vasconcelos et al., 2011a).

Although not famous by their acoustic behaviour, fish are probably the largest sound producing vertebrate group, using often acoustic signals during mating and territorial defence (Ladich, 2004). This work attempted for the first time to adapt an automatic recognition system to rapidly analyse vocal activity of fish. We used the Lusitanian toadfish (*Halobatrachus didactylus*), a teleost fish of the Batrachoidiforms order known to have a extensive acoustic repertoire and which has been studied as a great model to understand fish bioacoustical behaviour (Amorim, Simões, & Fonseca, 2008; Figure 1). Additionally, using the automatic recognition system, we also studied the acoustic rhythms and interactions of the breeding males using round-the-clock recordings. Such information is crucial to understand the acoustic behaviour of this species and was only possible using an automatic recognition system.

CHAPTER 1

AUTOMATIC CALL RECOGNITION AND INDIVIDUAL IDENTIFICATION OF FISH VOCALIZATIONS

Automatic call recognition and individual identification of fish vocalizations

Manuel Vieira a)

Departamento de Biologia Animal and Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa. Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

Paulo J. Fonseca b)

Departamento de Biologia Animal and Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa. Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

M. Clara P. Amorim c)

I.S.P.A. Instituto Universitário, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal

Carlos Teixeira d)

Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016

Lisboa, Portugal

Abstract

Many species communicate through acoustic signals that can fulfill several functions from mediating agonistic interactions to allow mate finding, and may provide cues for sexual selection. The study of acoustic communication often requires not only the recognition of species specific acoustic signals but also the identification of individual subjects, all this in a complex acoustic background. Moreover, when very long recordings are to be used, automatic methods can be invaluable tools to extract the relevant biological information. Here we present an automatic pattern recognition methodology based on the Hidden Markov Model that allows accurate detection and recognition of fish vocalizations. Our results show a relevant performance of the system that is capable not only to detect the mating vocalizations (boatwhistles) but also allows individual identification of male Lusitanian toadfish (*Halobatrachus didactylus*) reaching an identification rate of 95.5 ± 0.3%. The system was applied to a stream of round-the-clock recordings of toadfish vocalizations on their natural estuarine habitat and allowed precise detection of mate advertising sounds (boatwhistles).

INTRODUCTION

Fish often use acoustic signals during mating and territorial defense and are probably the largest sound producing vertebrate group (Ladich, 2004). The family Batrachoididae includes several species that have become good models for acoustic communication studies, such as the Lusitanian toadfish (*Halobatrachus didactylus;* Bass and McKibben, 2003; Vasconcelos et al., 2012). This species has an unusually large acoustic signal repertoire that includes boatwhistles, croaks, double croaks, grunts and other less frequent sound combinations (see Amorim et al., 2008, for details of the vocalizations). During the breeding season in Portugal the species can be found in estuarine shallow waters, often presenting high turbidity, where breeding males occupy nests under rocks and produce boatwhistles to attract females (R. O. Vasconcelos et al., 2011a). The advertisement boatwhistle, the most frequent vocalization in this species, is a highly stereotyped low-frequency sound with a duration ranging from 400 ms to 1200 ms and a dominant frequency between 50 Hz and 200 Hz (M. C. P. Amorim & Vasconcelos, 2008; M.C.P. Amorim et al., 2008).

Although much information regarding this species' vocalizations exists, including some of its emitting contexts and social functions, the difficulties in analyzing multiple round-the-clock recordings has limited the capabilities to infer details of the toadfish's acoustic behavior during the mating season in the natural habitat, such as vocal rhythms, acoustic social interactions and variability, possible effects of environmental parameters, anthropogenic noise, etc.

Several approaches have been reported to study extensive bioacoustic recordings. The simplest and most common are automatic detection methods that make use of, for example, energy thresholds or a matched filter to classify the vocalizations, usually followed by common multivariate statistical analysis procedures (e.g. discriminant function analysis or linear discriminant analysis). Recently, more robust methods using machine learning, such as Gaussian Mixture Models (GMMs; Reynolds and Rose, 1995), Artificial Neural Networks (ANN; Lippmann and Gold, 1987; Yu and Oh, 1997) and Hidden Markov Models (HMMs; Baker, 1975; Jelinek, 1976; Jelinek et al., 1975; Rabiner, 1989a; Young and Bloothooft, 1997), have been reported to successfully recognize and classify animals' vocalizations. Table I refers to application examples with sounds of different animals such as insects, birds, amphibians and mammals, but to the best of our knowledge no application to fish sounds exists. The HMM is the most used tool to make automatic speech recognition (ASR) systems (M Frikha, Hamida, & Lahiani, 2011), mostly because it is fast and accurate. The ANN has been more often used recently because of its non-linear approach, that does not assume the independence of each state as requested by the HMM, but needs more computational requirements, and does not have the capability of the HMM to accommodate time sequences (Mondher Frikha & Hamida, 2012). There are also hybrid approaches (Hagen & Morris, 2005; Neto, Almeida, Hochberg, & Martins, 1995).

Hidden Markov Model (HMM; Rabiner, 1989b) based approaches successfully allowed to identify species and individuals, to examine vocal repertoires and to classify vocalizations according to social context or behavior (Table I). HMM based systems can be a powerful option because HMM is a statistical tool that can model both temporal and spectral variations of vocalizations. In short, these systems in a first phase cluster discrete representations of several vocalizations forming a collection of references patterns (markov models). Then each unknown pattern will be compared to each reference pattern, and the most similar reference pattern is chosen. The result is a transformation of the acoustic signal into a string, i.e. a sequence of characters easy to process representing the sequence of patterns in the signal. With the success of these systems in human speech recognition, many improvements were added to Hidden Markov Models systems such as Hidden Markov Models toolkit (HTK; Young et al., 2002), easing the development of automatic speech recognition systems.

In the present work we use the HTK and apply several features common in speech recognition to achieve an efficient automatic recognition system. This system is capable to deal with extensive recordings allowing recognition and classification of the Lusitanian toadfish stochastic vocalizations. The system also aimed to discriminate among vocalizing individuals since a previous study demonstrated the existence of inter-individual differences in boatwhistles during short periods of time (< 10 min) (Amorim and Vasconcelos, 2008; Amorim et al. 2011).

DATA COLLECTION

We recorded the vocalizations of adult territorial males during the breeding season (May to July 2012). The males spontaneously occupied concrete artificial hemicylindrical nests, capped at one end, which we deployed in the Tagus estuary (Air Force Base 6, Montijo, Portugal; 38º42'N, 8º58'W). These nests, positioned at 2 m from each other in a row, had a perfurated metal plaque in the front of the nest with an opening large enough to allow females or small prey (e.g. crabs) to enter the nest but

Examples of automated recognition applications in bioacoustics.

Class		Objective	System	Feature	Reference
Insects		S	ANN	TDSC	(D. Chesmore, 2008)
	Orthoptera	S	ANN	TDSC	(E. . Chesmore, 2001)
		S	ANN	TDSC	(E. D. Chesmore & Ohya, 2004)
	Cicadas	S	GMM	LFCC	(Potamitis, 2007)
Birds		S	ANN	LPC	(McIlraith & Card, 1995)
		S	ANN		(Mills, 1995)
		C	DTW		(Anderson, Dave, & Margoliash, 1996)
		C	LDA	MFC, LPC	(Lee, Lee, & Huang, 2006)
		S	HMM	[a]	(Chou, Lee, & Ni, 2007)
		S,C	ANN	MFC	(Chou, Liu, & Cai, 2008)
		S	HMM	MFC	(Trifa & Kirschel, 2008)
		S	LDA,DT,SVM	[a]	(Acevedo & Corrada-Bravo, 2009)
		C	HMM	LPC	(Chu & Blumstein, 2011)
	Corncrake	I	ANN	[a]	(Terry & McGregor, 2002)
	Norwegian Ortolan Bunting	I, C	HMM	MFC	(Trawicki, 2005)
Amphibians		S,I	ANN	WPD	(Yen & Fu, 2001)
		S	LDA,DT,SVM	[a]	(Acevedo & Corrada-Bravo, 2009)
Mammals	Cetaceans	S	ANN	SCF	(Potter, Mellinger, & Clark, 1994)
		C	ANN	[b]	(Murray, Mercado, & Roitblat, 1998)
		C	ANN	[c]	(Schaar, 2007)
		C	HMM	LPC, MFC	(Pace, White, & Adam, 2012)
	Deer	I	ANN	MFC	(Reby, André-Obrecht, Galinier, Farinas, & Cargnelutti, 2006)
	Bats	S	ANN	[a]	(Parsons & Jones, 2000)
		S	ANN	[a]	(Parsons, 2001)
	Pigs (stress calls)	C	ANN	LPC	(Schön, Puppe, & Manteuffel, 2001)
	Sea lions (females)	I	ANN	[d]	(Campbell, Gisiner, & David, 2000)
	Elephants	I, C	HMM	MFC, PLP	(P. J. Clemins, Johnson, Leong, & Savage, 2005)
	Cows	C	HMM	MFC	(Jahns, 2008a)
	Primates	C	ANN	LPC	(Pozzi, Gamba, & Giacoma, 2010)

ANN, artificial neural network; GMM, generalized method of moments; LFCC, linear frequency cepstral coefficients; LDA, linear discriminant analysis; DT, decision tree; SVM, support vector machine; DTW, dynamic time warping; MFC, Mel-frequency cepstral; LPC, linear prediction cepstral; S, species recognition, I, individual identification; C, call type recognition; TDSC, Time domain signal coding; SCF, spectogramcorrelator filter; WPD, wavelet packet decomposition; [a], vector composed of call variables; [b], each vocalization was characterized by its simultaneous modulations in duty cycle and peak frequency; [c], features were selected using a local discriminant basis; [d], each call was represented by an average logarithmic spectrum on the backpropagation network input layer.

that prevented the larger subject males from escaping. The males' vocalizations were recorded with custom-made hydrophones (Fonseca & Maia Alves, 2011) placed next to each experimental nest in mid-lateral position and about 10 cm above the substrate. This arrangement ensured unequivocal recording of the sounds of the nest-holder male and so its individual signal identity throughout the recordings. Notice that the concrete nests did not noticeably influence sound transmission and that the sounds produced by males occupying adjacent nests arrived much attenuated relative to the nest-holder vocalizations. The signal from each hydrophone was recorded to a 16 channel standalone data logger (Measurement Computing Corporation LGR-5325, Norton, Virginia, USA, 16 bits resolution, 4 kHz sampling rate).

The data set consisted of 13 days round-the-clock simultaneous recordings of 16 nest-holders. The boatwhistles recorded in each channel (nest) were manually selected, classified and cut with the aid of a matched filter function available in Ishmael 1.0 (Mellinger, 2002). Each individualized vocalization preceded by 0.5 s was stored in a ca. 2 seconds separate file. Vocalizations of 13 males were considered for the analyses, including a total of 32559 boatwhistles, 23 croaks, 24 double croaks and 77 grunts. Due to the scarcity of the instances collected for croaks, double croaks and grunts, we additionally used at least 6 instances for each of these signal types, obtained from our previous toadfish recordings (July 2001 to September 2002 (M. C. P. Amorim, Vasconcelos, Marques, & Almada, 2006).

METHODS

This section describes the machine learning methods used for the automatic search of the toadfish vocalization types in the sound recordings. Fig.1 summarizes the main stages of the signal recognition phase which include the feature extraction and the aligning of the obtained feature vectors with several previously trained HMM models

SIGNAL PROCESSING

The first stage of an automatic recognition system is the features extraction from the acoustic signal. This procedure imposes the adjustment of several parameters such as window size, frequency range, and selection of relevant features (see below), to obtain the vocalizations' properties that meet the requirements of the recognition system. Based on the information available from the Lusitanian toadfish vocalizations a set of tests allowed establishing a window size of 32 milliseconds modulated by a 50% overlapping Hamming window as appropriate. The window size was lengthened to account for the lower fundamental frequencies of the Lusitanian toadfish vocalizations.

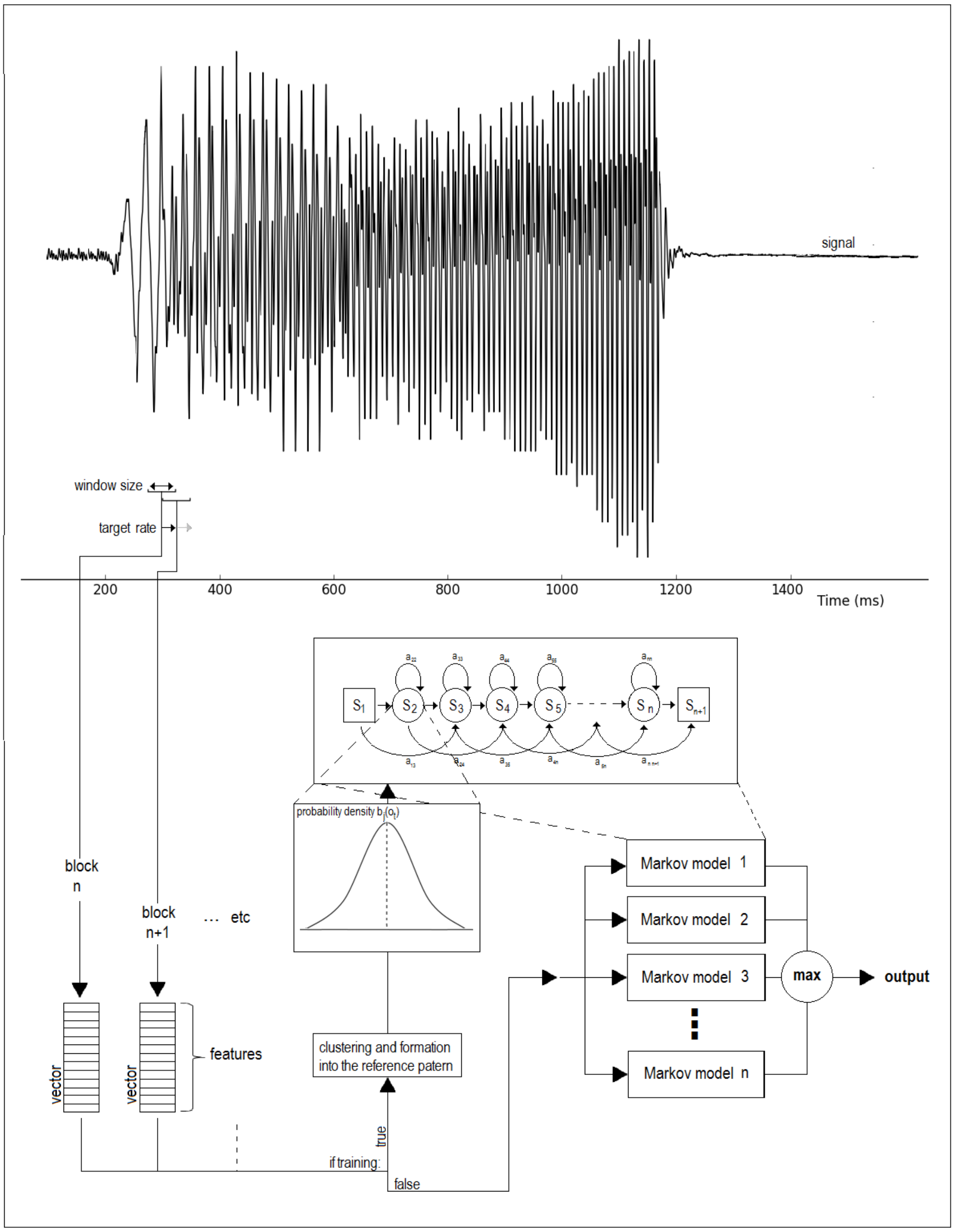

Fig. 1. Workflow of the HMM recognition system using the HTK (diagram based on Young et al., 2006). The signal represents an oscilogram of a boatwhistle; the window size and the target size are represented in the same time scale. S - states of the markov model, being S_1 the firts state and S_{n+1} the last state of the model.

To extract vocalization features many different methods are reported in the literature. For example, in speech recognition some features extraction methods are based on the specific known deconstruction of the signal accomplished by the human ear, or on the way the signal is produced (many of these mentioned in Table I). In this study we used several combinations of relevant signal features: energy, cepstrum, Mel-frequency cepstral (MFC), perceptual linear prediction (PLP), delta and acceleration coefficients, also used in previous bioacoustics studies (Table I; see Davis, 1990 and Young et al., 2002, for details in these signal features coefficients). Preliminary individual identification experiments using training and testing data including all the 13 toadfish referred before produced most relevant results with three feature sets: (1) MFC with delta, acceleration and energy; (2) MFC with delta, acceleration and cepstrum; and (3) PLP with delta, acceleration and energy.

To improve the system's accuracy we also tested different frequency ranges adjusted to the spectrum of the vocalizations recorded in the natural habitat. These tests considered different low (0, 10, 20, 30, 40 and 50 Hz) and high (300, 500, 700, 1000, 2000 Hz) frequency cut-offs. The best recognition results were obtained with a 20 Hz to 2000 Hz frequency range that led to preliminary results of 95% identification rate for feature set 1, a 96% for set 2, and a 94% for set 3. However the identification rate improvement by adding information extending the upper frequency filter from 1000 Hz to 2000 Hz was 3.4% at maximum.

THE HMM STRUCTURE TIME ALIGNMENT

The second stage of the system described in Fig.1 aims to align the obtained feature vectors with several previously trained HMM models. Each alignment, performed with a Viterbi algorithm (Forney, 1973), allows the computation of a likelihood score. The final goal is to find the HMM that presents the highest likelihood score. At least one HMM model must be trained for each type of signal pattern to be found.

We used a linear topology where each model has several states that can transit to itself, to the next state or to the following one, except in the initial and final states; the initial sate could not transit to itself, and the final state has no transitions (see the model in Fig.1). Previous knowledge of the average duration of each type of vocalization was considered for deciding the HMM number of states to be used. Thus, we considered 14 states for the boatwhistle model, 5 states for the croak model, 7 states for the double croak model, and 14 states for the grunt train model (note that, as referred before and illustrated in Fig.1, each state, except the first and the last ones, can have transitions to itself). We further defined a model with 5 states for modeling background noise and an extra model with 5 states to all non-biological patterns with high energy (e.g. small wave splash). The transition probabilities and the observation probability densities of each state are estimated with the Baum-Welch algorithm using previously labeled patterns. We used the start and end points of each model alignment to estimate the duration of each vocalization.

The whole approach was developed using the Hidden Markov Model Toolkit (HTK, University of Cambridge, UK). The HTK is a group of modules written in C to create automatic recognition systems for human speech analyses (S Young et al., 2002).

AUTOMATIC RECOGNITION SYSTEMS

We prepared two automatic recognition systems, one to identify boatwhistle vocalizations of individual fish and another to identify call types. The first experiments (Section IV A: Individual Identification) aimed to recognize one individual among several vocalizing toadfish. Each individual was represented by one model that integrated its boatwhistles. Since testing the recognition system only once is insufficient to evaluate its performance, a resampling method was used based on a random subsampling validation (Efron, 1981). We used a training set of 40 boatwhistles resampled from all data, while the remaining boatwhistles were used for the test. The test included 100 repetitions. Each repetition considered different training and testing sets both obtained by randomly resampling the whole boatwhistles data set. These experiments were repeated for each feature set mentioned in Section III-A. For the subsequent experiments we only used the feature set that presented the highest individual identification scores. To establish the training set size we made two experiments with training sets ranging from 3 to 40 boatwhisles considering 100 repetitions. In the first experiment we considered the whole recordings and selected the training set randomly, while in the other we restricted the selection to the first 50 boatwhistles produced by each fish.

In the call type identification experiments (Section IV B: Call type Identification) we used the same data as for the experiments in Section A as well as additional data obtained in previous seasons and kept in our sound archive (M. C. P. Amorim et al., 2006). A total of 19123 boatwhistles, 23 croaks, 25 double croaks and 77 grunts were used. From these 9 were randomly resampled to include in the training set for each vocalization (also considering 100 repetitions). Fish used for the training set were also included in the testing set. To test the ability of the system to recognize boatwhistles of individual fish not included in the training data, we also tested 504 boatwhistles from 4 other individuals with the previously trained HMM boatwhistle models.

RESULTS

INDIVIDUAL IDENTIFICATION

All HMM models representing boawhistles of each individual fish obtained from the different feature sets mentioned in Section III-A presented an identification rate significantly higher than the one expected from a random classification. Table II presents the confusion matrix computed for the experiment using MFC with acceleration, delta and energy coefficients. The HMM system using this feature set resulted in a global identification rate of 95.0 ± 0.4 % (mean ± standard deviation), the lowest from the three considered feature sets. Table III shows the result of the HMM classification based on feature set 2 that used the cepstrum coefficient instead of energy. This approach generated the highest identification rate and achieved an average of 95.5 ± 0.3 %. The computation on PLP features (Table IV) generated an identification rate averaging 95.2 ± 0.3 %. Identification rate means of all individuals were significantly different among the three sets (ANOVA, $p < 0.001$). The overall correct classification of each individual's boatwhistles ranged between 63.8 and 100 % according to the confusion matrices obtained from 100 repetitions. The mean identification rate per individual using the three sets of features ranged from 82 % to 100 % with the exception of fish no. 2, where the value was below 70%.

The use of a totally random training set always produced higher mean identification rates (Fig. 2). In fact, training the model based on only seven boatwhistles randomly sampled within the whole data set resulted in at least 90% of correct identifications. However, if the training set was restricted to a random selection among the first 50 boatwhistles produced by each fish, more vocalizations were needed to achieve the same results (11 vocalizations). Incrementing the number of vocalizations used for training the model resulted in asymptotic improvement of the mean identification rates along with decreased standard deviation.

TABLE I. Confusion matrix from the hidden Markov model classification computed using the coefficients MFC,energy, delta and acceleration considering a frequency rangefrom 20 Hz to 2000 Hz. The data set included 32559 boatwhistles from thirteen male Lusitanian toadfishesproduced during 12 days. The model was trained with 40 randomly selected boatwhistles, and tested with the remaining ones. Results are averages from 100 repetitions. 95.0 ± 0.4 % of the tested boatwhistles were correctly classified.

	Predicted group membership													
Fish	1	2	3	4	5	6	7	8	9	10	11	12	13	% correct
1	404	0	0	0	0	0	3	0	0	0	0	0	0	99.3
2	11	98	3	7	4	0	0	0	17	2	0	0	5	66.7
3	0	0	1058	3	10	12	4	0	1	2	5	7	1	95.9
4	0	0	1	1522	8	0	0	0	0	113	1	6	0	92.2
5	1	0	6	3	1506	0	0	0	0	3	0	21	12	97.0
6	0	0	1	0	0	1543	1	0	0	0	2	11	0	99.0
7	17	1	2	1	1	39	684	8	9	1	60	1	3	82.7
8	0	0	2	1	1	0	32	605	0	1	4	0	0	93.7
9	23	0	2	1	2	6	5	0	474	0	1	0	2	91.9
10	0	0	7	74	9	0	0	0	0	1480	1	0	0	94.2
11	0	0	0	0	0	0	4	1	0	9	1538	0	0	99.1
12	0	0	0	0	12	0	1	0	0	1	0	1544	12	98.3
13	0	0	0	0	5	5	0	0	0	0	0	18	1571	98.2
total														95.0 ± 0.4

CALL TYPE IDENTIFICATION

The call type identification system achieved a high identification rate when considering boatwhistle recognition, and a low for other vocalizations. A mean identification rate 98.4 % was always obtained in the identification of boatwhistles. In contrast, the other vocalizations were poorly recognized by the system, with correct identification rates below 10%. Some mistakes in the classification of grunts were due to misidentifications associated with the last 100 ms of the boatwhistles. Using the 504 boatwhistles of other individuals not used for training the models, we obtained an identification rate of 99 %. The remainder 1 % was associated with some overlaps of boatwhistles produced simultaneously by different fish, which were classified by the system as a single boatwhistle.

TABLE II. Confusion matrix from the hidden Markov model classification computed using the coefficients MFC, cepstral, delta and acceleration considering a frequency rangefrom 20 Hz to 2000 Hz. The data set included 32559 boatwhistles from thirteen male Lusitanian toadfishesproduced during 12 days. The model was trained with 40 randomly selected boatwhistles, and tested with the remaining ones. Results are averages from 100 repetitions. 95.5 ± 0.3% of tested boatwhistles were correctly classified.

	Predicted group membership													
Fish	1	2	3	4	5	6	7	8	9	10	11	12	13	% correct
1	398	0	0	0	0	0	0	0	0	0	0	0	0	100
2	14	101	4	1	1	3	0	0	23	0	0	1	7	65.2
3	0	1	1065	1	1	13	0	0	1	0	2	6	1	97.6
4	0	0	0	1514	4	0	0	0	0	136	1	8	2	90.9
5	0	0	5	6	1527	0	0	0	0	2	1	4	1	98.8
6	0	0	1	0	0	1542	0	0	0	0	1	10	0	99.2
7	18	0	2	2	9	33	704	2	3	1	36	1	1	86.7
8	0	0	0	0	3	0	20	609	0	0	3	1	0	95.8
9	26	0	0	0	0	10	2	0	475	0	1	0	3	91.9
10	1	0	8	86	5	1	8	2	0	1465	2	2	7	92.3
11	0	0	0	0	0	0	5	0	1	9	1567	0	0	99.1
12	0	0	0	0	7	0	0	0	0	0	0	1554	10	98.9
13	0	0	0	0	0	7	0	0	0	0	0	18	1576	98.4
total														95.5 ± 0.3

DURATION OF THE BOATWHISTLES

The duration of the boatwhistles estimated from our two recognition systems (identification of boatwhistles through individual identification and identification of call type across individuals) was compared with the duration measured by the user(Fig.2). The duration estimated using the call type identification system was very similar to the values measured manually, turning this system into a powerful tool to assess signal durations in large data sets. However, the individual identification system proved less adequate for the evaluation of the boawhistle duration since the estimates were in this case significantly different from the measurements (t-test; p<0.001).

TABLE III. Confusion matrix from the hidden Markov model classification computed using the coefficients PLP,energy, delta and acceleration considering a frequency rangefrom 20 Hz to 2000 Hz. The data set included 32559 boatwhistles from thirteen male Lusitanian toadfishesproduced during 12 days. The model was trained with 40 randomly selected boatwhistles, and tested with the remaining ones. Results are averages from 100 repetitions. 95.2 ± 0.3% of tested boatwhistles were correctly classified.

	Predicted group membership													
Fish	1	2	3	4	5	6	7	8	9	10	11	12	13	% correct
1	397	0	0	0	0	0	0	0	0	0	0	0	0	100
2	13	102	5	2	1	3	1	0	23	0	0	2	8	63.8
3	0	0	1063	1	1	14	0	0	1	0	2	4	0	97.9
4	0	0	0	1514	4	0	0	0	0	141	1	9	2	90.6
5	0	0	5	7	1526	0	0	1	0	3	2	7	0	98.4
6	0	0	1	0	0	1539	0	0	0	0	1	11	0	99.2
7	17	0	1	2	6	33	696	3	2	0	47	1	1	86.0
8	0	0	0	1	3	0	23	607	0	1	4	1	0	94.8
9	28	0	0	0	0	12	3	0	475	0	1	0	3	91.0
10	1	0	11	84	7	1	10	2	0	1460	3	2	8	91.9
11	0	0	0	0	0	0	6	1	0	9	1552	0	0	99.0
12	0	0	0	1	7	0	0	0	0	0	0	1552	8	99.0
13	0	0	0	0	0	6	0	0	0	0	0	18	1578	98.5
total														95.2 ± 0.3

DISCUSSION

Our general goal was to develop a tool to study the vocalization activity of adult Lusitanian toadfish breeding males. This objective requires the identification of the several vocalization types and their assignment to individual males. The analysis of round-the-clock sound recordings during the breeding season, that lasts about three months, represented an unaffordable human task. Therefore, this task required the use of state of the art automatic machine learning techniques for automatic recognition of the calls. Recent major advances in automatic speech recognition have enabled the automated analysis of bioacoustic signals (Table 1), but to the best of our knowledge the current work presents the first results of the application of an automatic recognition system to distinguish fish acoustic signals. The results of the application of HMM-based

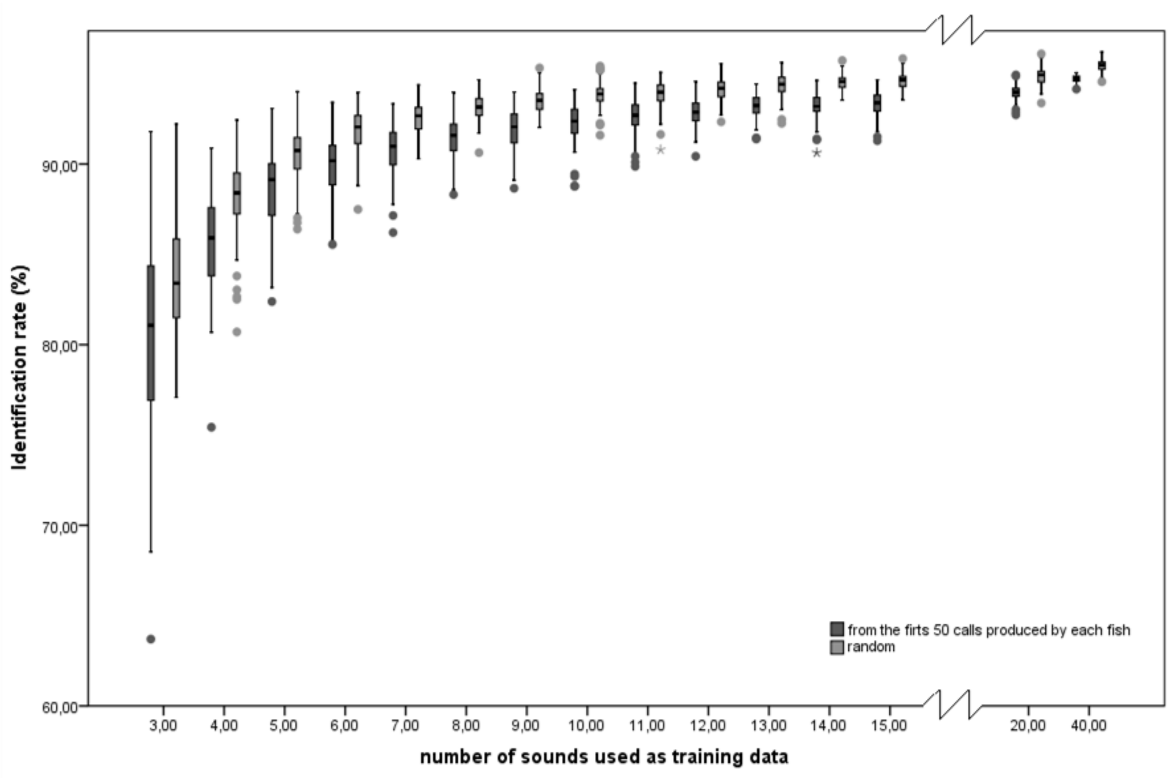

Fig. 2. Boxplots representing the relation between the accuracy of the recognition system and the size of the training data set used (each boxplot represent 100 repetitions). Dark gray symbols represent results using training sets restricted to the firsts 50 boatwhistles produced by each fish. Lighter gray symbols represent the experiment considering a training set randomly selected.

recognition systems using MFC or PLP to extract suitable features from fish vocalizations showed a good performance allowing to correctly classify the majority (> 90%) of each toadfish boatwhistles.

Tests using several frequency bandwidths revealed that the essential signal information lies within 20 and 1000 Hz, a frequency range encompassing the spectral components in Lusitanian toadfish vocalizations (M. C. P. Amorim & Vasconcelos, 2008). Below 20 Hz the recording was dominated by background noise. Preliminary results (data not shown) considering a 20 - 250 Hz frequency band to train the recognition system for individual discrimination achieved an identification rate between 72% and 88%. According to (Raquel O Vasconcelos, Sisneros, Amorim, & Fonseca, 2011), this frequency range corresponds to the range of higher Lusitanian toadfish

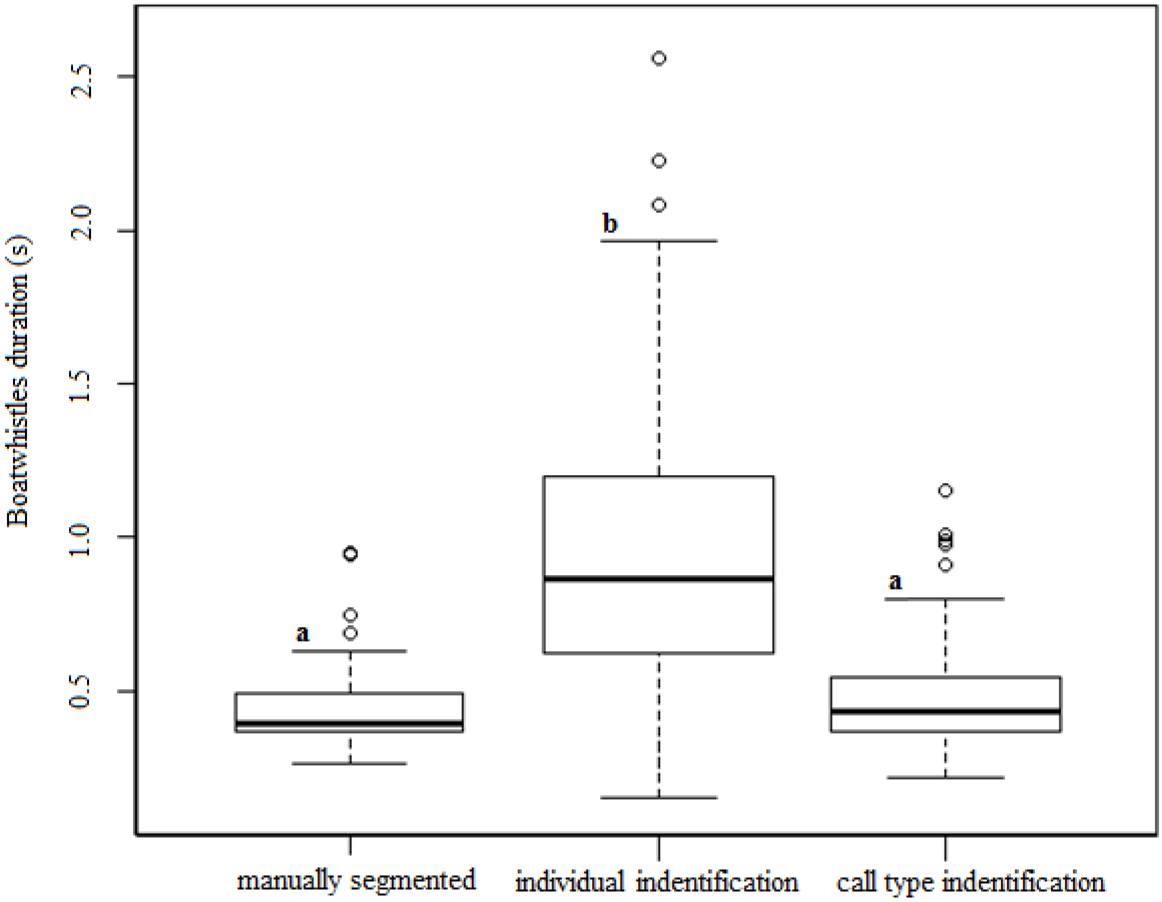

Fig. 3. Duration in seconds of 65 boatwhistles estimated manually, by the individual identification system and by the calll type identification system computed on the MFC with cepstral, delta and acceleration coefficients with a bandpass filter from 20 Hz to 2000 Hz. Different letters represent pairwise significant differences ($p < 0.001$) using t-student tests.

hearing sensitivity. Moreover, auditory evoked potentials showed that the amplitude modulation and spectral components of boatwhistles were well represented in the brain (Raquel O. Vasconcelos, Simões, Almada, Fonseca, & Amorim, 2010). Altogether, these results suggest that the boatwhistles should have enough information to allow males to distinguish among other individuals vocalizing nearby in their natural habitat. Additional experiments indicated that our recognition system (with energy normalization) could correctly classify boatwhistles produced by fish from a distance up to 4 m, but identification was impaired at 6 m. Although very preliminary, such information may allow estimating signal recognition distances in the natural habitat if combined with studies on call propagation and discrimination by fish of calls affected by propagation. These studies could be conducted using either physiological and/or behavioral approaches.

By selecting the more suitable parameters to include in the feature set to be used in the HMM model we successfully identified the vocalizations of individual males (> 90 %). This method reached values at least similar to or even higher than the identification rates observed in automatic recognition approaches with mammals' (Campbell et al., 2000; P. Clemins, 2005; Reby et al., 1997, 2006) and birds' (Terry & McGregor, 2002; Trawicki, 2005) vocalizations. The comparison, however, is not straightforward not only because these investigations were based on recordings involving a smaller number of animals singing simultaneously, which generally improves recognition (as exemplified in Trawicki, 2005), but also different authors focused on different animal models (Table I), used different quality recordings (e.g. Clemins et al. (2005),recorded captive elephants with microphones held by collars ensuring the acquisition of very good quality non-degraded sounds), and applied different recognition algorithms (Table 1).

One important issue of automatic sound recognition systems is the choice of data to be used in training the models (Reby et al., 1997; Tao, Johnson, & Osiejuk, 2008; S Young et al., 2002). As depicted in Fig.2, a larger number of vocalizations available for training the model usually results in improvement of the model's recognition. Thus, a trade-off exists between the effort to manually extract the vocalizations and the accuracy of the system. Our experiments also point to an improvement of the models' recognition by including in the training set vocalizations sampled within the whole data set.

Our call type recognition system was effective to identify the boatwhistles, the most salient vocalization of male Lusitanian toadfish, but was not able to correctly classify other vocalization types such as croaks, double-croaks and grunts. Indeed, the identification rate of these call types was too low (< 10%) for considering their automatic monitoring. Some heterogeneity in the recognition rates of different call types of a species is commonly reported (E. D. Chesmore & Ohya, 2004; Jahns, 2008b; Kogan & Margoliash, 1998; Parsons & Jones, 2000; Schön et al., 2001; Trawicki, 2005). Several reasons may be responsible for the low identification accuracy of sounds (S Young et al., 2002). The low occurrence of croaks, double croaks and grunts was the most likely factor, since it did not allow to properly adjust the model parameters. In contrast to the boatwhistles, the pulsated nature of these sounds could have influenced misclassification. In fact, Young et al. (2002) point out that if the parameters are adjusted to discriminate some sound characteristics, the system could overlook other sound types.

A main parameter in the characterization of a species' vocalizations is the duration of sounds. Our system for call type recognition appeared to give an excellent estimation of the duration of boatwhistles. However, caution is recommended since errors can be expected caused by overlaps of vocalizations produced by different individuals (e.g.), misclassification of parts of boatwhistles, or unrecognized atypical boatwhistles due to much shorter or longer duration, noise interference, etc. These are not problems easy to solve, for example Stowell et al. (2013) attempted to solve the problem of call overlap in birds and Zhang and Gatica-Perez (2005), tried to decreased recognition error by using an additional algorithm based in Bayes statistics to recognize less frequent sounds.

Here we demonstrate the usefulness of using HMM based automatic recognition systems to extensively analyze toadfish sound recordings, and this is likely extensive to other fish species. Future work using this system will allow assessing subtleties of the Lusitanian toadfish vocal behaviour in its natural habitat, which requires very long round-the-clock sound recordings, such as vocalization rhythms, vocal interactions among fish, etc. Moreover, adjustments on this system allowing more subtle discriminations of vocalizations could for example permit to infer the dynamics of agonistic interactions throughout the breeding season, by sorting advertisement from agonistic boatwhistles which differ in dominant frequency and amplitude modulation (Raquel O. Vasconcelos et al., 2010). Also, by increasing the data available we may improve the ability to recognize the less frequent vocalizations, which may have a relevant role in this species communication. In summary, this recognition system could be an important tool in studies where analysis of very large recordings is required and could likely be expanded to various fish species.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology Foundation (project PTDC/MAR/118767/2010), Portugal.

REFERENCES

Acevedo, M., and Corrada-Bravo, C. (2009). “Automated classification of bird and amphibian calls using machine learning: A comparison of methods,” Ecological Informatics, 4, 206–214.

Amorim, M. C. P., Simões, J. M., and Fonseca, P. J. (2008). “Acoustic communication in the Lusitanian toadfish, Halobatrachus didactylus: evidence for an unusual large vocal repertoire,” Journal of the Marine Biological Association of the UK, 88, 1069–1073.

Amorim, M. C. P., and Vasconcelos, R. O. (2008). “Variability in the mating calls of the Lusitanian toadfish Halobatrachus didactylus : cues for potential individual recognition,” Journal of Fish Biology, 73, 1267–1283.

Amorim, M. C. P., Vasconcelos, R. O., Marques, J. F., and Almada, F. (2006). “Seasonal variation of sound production in the Lusitanian toadfish Halobatrachus didactylus,” Journal of Fish Biology, 69, 1892–1899.

Anderson, S., Dave, A., and Margoliash, D. (1996). “Template based automatic recognition of birdsong syllables from continuous recordings,” Journal of the Acoustical Society of America,.

Baker, J. (1975). “The DRAGON system--An overview,” Acoustics, Speech and Signal Processing, IEEE Transactions on, 23, 24–29.

Bass, A., and McKibben, J. (2003). “Neural mechanisms and behaviors for acoustic communication in teleost fish,” Progress in neurobiology, 69, 1–26.

Campbell, G., Gisiner, R., and David, A. (2000). “Acoustic identification of female Steller sea lions (Eumetopias jubatus),” The Journal of the Acoustical Society of America,.

Chesmore, D. (2008). “Automated bioacoustic identification of insects for phytosanitory and ecological applications,” Computational Bioacoustics for Assessing Biodiversity, 234, 59–72.

Chesmore, E. . (2001). “Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals,” Applied Acoustics, 62, 1359–1374.

Chesmore, E. D., and Ohya, E. (2004). “Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition,” Bulletin of Entomological Research, 94, 319–330.

Chou, C., Lee, C., and Ni, H. (2007). “Bird species recognition by comparing the HMMs of the syllables,” Innovative Computing, Information and Control, 2007. ICICIC’07. Second International Conference on, 143.

Chou, C., Liu, P., and Cai, B. (2008). “On the Studies of Syllable Segmentation and Improving MFCCs for Automatic Birdsong Recognition,” Asia-Pacific Services Computing Conference, 2008. APSCC’08 IEEE, 745–750.

Chu, W., and Blumstein, D. (2011). “Noise robust bird song detection using syllable pattern-based hidden Markov models,” Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, 345–348.

Clemins, P. (2005). Automatic classification of animal vocalizations Marquette University.

Clemins, P. J., Johnson, M. T., Leong, K. M., and Savage, A. (2005). “Automatic classification and speaker identification of African elephant (Loxodonta africana) vocalizations,” The Journal of the Acoustical Society of America, 117, 956.

Davis, R. (1990). “Perceptual linear predictive (PLP) analysis of speech,” The Journal of the Acoustical Society of America,.

Efron, B. (1981). Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods, Biometrika, Vol. 68.

Fonseca, P. J., and Maia Alves, J. (2011). Electret capsule hydrophone: a new underwater sound detector, Patent Application PT105, 933.

Forney, G. (1973). “The viterbi algorithm,” Proceedings of the IEEE, 6, 268 – 278.

Frikha, M., and Hamida, A. Ben (2012). “A Comparitive Survey of ANN and Hybrid HMM/ANN Architectures for Robust Speech Recognition,” American Journal of Intelligent Systems, 2, 1–8.

Frikha, M., Hamida, A., and Lahiani, M. (2011). “Hidden Markov models (HMMs) isolated word recognizer with the optimization of acoustical analysis and modeling techniques,” International Journal of the Physical Sciences, 6, 5064–5074.

Hagen, A., and Morris, A. (2005). “Recent advances in the multi-stream HMM/ANN hybrid approach to noise robust ASR,” Computer Speech & Language, 19, 3–30.

Jahns, G. (2008). “Call recognition to identify cow conditions—A call-recogniser translating calls to text,” Computers and Electronics in Agriculture, 62, 54–58.

Jahns, G. (2008). “Call recognition to identify cow conditions—A call-recogniser translating calls to text,” Computers and Electronics in Agriculture, 62, 54–58.

Jelinek, F. (1976). “Continuous speech recognition by statistical methods,” Proceedings of the IEEE, 64, 532–556.

Jelinek, F., Bahl, L., and Mercer, R. (1975). “Design of a linguistic statistical decoder for the recognition of continuous speech,” Information Theory, IEEE Transactions on, 21, 250–256.

Kogan, J. A., and Margoliash, D. (1998). “Markov models : A comparative study,” 103, 2185–2196.

Ladich, F. (2004). “Sound production and acoustic communication,” The Senses of Fish, Springer Netherlands, pp. 210–230.

Lee, C., Lee, Y., and Huang, R. (2006). “Automatic recognition of bird songs using cepstral coefficients,” Journal of Information Technology and Applications, 1, 17–23.

Lippmann, R., and Gold, B. (1987). “Neural classifiers useful for speech recognition,” 1st International Conference on Neural Networks, IEEE,.

McIlraith, A., and Card, H. (1995). “Birdsong recognition with DSP and neural networks,” WESCANEX 95. Communications, Power, and Computing. Conference Proceedings., 409–414.

Mellinger, D. K. (2002). Ishmael 1 . 0 User ’ s Guide,.

Mills, H. (1995). “Automatic detection and classification of nocturnal migrant bird calls,” The Journal of the Acoustical Society of America, 97, 3370.

Murray, S., Mercado, E., and Roitblat, H. (1998). “The neural network classification of false killer whale (Pseudorca crassidens) vocalizations,” The Journal of the Acoustical Society of America, 104, 3626.

Neto, J., Almeida, L., Hochberg, M., and Martins, C. (1995). “Speaker-adaptation for hybrid HMM-ANN continuous speech recognition system,” Proc. of the European Conference on Speech Comm. and Tech., Madrid, 2171–2174.

Pace, F., White, P., and Adam, O. (2012). “Hidden Markov Modeling for humpback whale (Megaptera Novaeanglie) call classification,” Proceedings of Meetings on Acoustics, 070046.

Parsons, S. (2001). “Identification of New Zealand bats (Chalinolobus tuberculatus and Mystacina tuberculata) in flight from analysis of echolocation calls by artificial neural networks,” Journal of Zoology, 253, 447–456.

Parsons, S., and Jones, G. (2000). “Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks,” Journal of Experimental Biology, 203, 2641–56.

Potamitis, I. (2007). “Automatic acoustic identification of crickets and cicadas,” Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on, 1–4.

Potter, J., Mellinger, D., and Clark, C. (1994). “Marine mammal call discrimination using artificial neural networks,” The Journal of the Acoustical society of America, 96, 1255.

Pozzi, L., Gamba, M., and Giacoma, C. (2010). “The use of Artificial Neural Networks to classify primate vocalizations: a pilot study on black lemurs,” American Journal of Primatology, 72, 337–348.

Rabiner, L. R. (1989). “A tutorial on hidden Markov models and selected applications in speech recognition,” Proceedings of the IEEE, 77, 257–286.

Rabiner, L. R. (1989). “A tutorial on hidden Markov models and selected applications in speech recognition,” Proceedings of the IEEE, 77, 257–286.

Reby, D., André-Obrecht, R., Galinier, A., Farinas, J., and Cargnelutti, B. (2006). “Cepstral coefficients and hidden Markov models reveal idiosyncratic voice characteristics in red deer (Cervus elaphus) stags,” The Journal of the Acoustical Society of America, 120, 4080.

Reby, D., Lek, S., Dimopoulos, I., Joachim, J., Lauga, J., and Aulagnier, S. (1997). “Artificial neural networks as a classification method in the behavioural sciences,” Behavioural Processes, 40, 35–43.

Reynolds, D., and Rose, R. (1995). “Robust text-independent speaker identification using Gaussian mixture speaker models,” IEEE Trans. on Speech and Audio Proc., 3, 72–83.

Schaar, M. van der (2007). “Neural network-based sperm whale click classification,” Journal of the Marine Biological Association of the United Kingdom, 87, 35.

Schön, P., Puppe, B., and Manteuffel, G. (2001). “Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (Sus scrofa),” The Journal of the Acoustical Society of America, 110, 1425.

Stowell, D., Saso, M., Bodana, J., and Plumbley, M. D. (2013). “Improved multiple birdsong tracking with distribution derivative method and markov renewal process clustering,” Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 468–472.

Tao, J., Johnson, M. T., and Osiejuk, T. S. (2008). “Acoustic model adaptation for ortolan bunting (Emberiza hortulana L) song-type classification,” The Journal of the Acoustical Society of America, 123, 1582–90.

Terry, A., and McGregor, P. (2002). “Census and monitoring based on individually identifiable vocalizations: the role of neural networks,” Animal Conservation, 5, 103–111.

Trawicki, M. (2005). “Automatic song-type classification and speaker identification of Norwegian Ortolan Bunting (Emberiza hortulana) vocalizations,” Machine Learning for Signal Processing, 2005 IEEE Workshop on., 277–282.

Trifa, V., and Kirschel, A. (2008). “Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models,” The Journal of the Acoustical Society of America, 123, 2424.

Vasconcelos, R. O., Carrico, R., Ramos, A., Modesto, T., Fonseca, P. J., and Amorim, M. C. P. (2011). “Vocal behavior predicts reproductive success in a teleost fish,” Behavioral Ecology, 23, 375–383.

Vasconcelos, R. O., Simões, J. M., Almada, V. C., Fonseca, P. J., and Amorim, M. C. P. (2010). “Vocal Behavior During Territorial Intrusions in the Lusitanian Toadfish: Boatwhistles Also Function as Territorial ‘˜Keep-Out’ Signals,” Ethology, 116, 155–165.

Vasconcelos, R. O., Sisneros, J. a, Amorim, M. C. P., and Fonseca, P. J. (2011). “Auditory saccular sensitivity of the vocal Lusitanian toadfish: low frequency tuning allows acoustic communication throughout the year,” Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology, 197, 903–13.

Yen, G., and Fu, Q. (2001). “Automatic frog calls monitoring system: a machine learning approach,” International Journal of Computational Intelligence and Applications, 1, 165–186.

Young, S., and Bloothooft, G. (1997). Corpus-based methods in language and speech processing, Kluwer Academic Publishers.

Young, S., Evermann, G., and Gales, M. (2006). The HTK book version 3.4, Cambridge University Press.

Young, S., Evermann, G., Kershaw, D., and Moore, G. (2002). The HTK book (for HTK version 3.2), Cambridge University Press.

Yu, H., and Oh, Y. (1997). “A neural network for 500 vocabulary word spotting using acoustic sub-word units,” Proc. Int. Conf. on Acoustic Speech and Signal Proces sing, 3277–3280.

Zhang, D., and Gatica-Perez, D. (2005). “Semi-supervised adapted hmms for unusual event detection,” Computer Vision and Pattern Recognition, 611 – 618.

CHAPTER 2

PATTERNS OF MALE VOCALIZATION ACTIVITY AND ACOUSTIC INTERACTIONS IN THE LUSITANIAN TOADFISH (*HALOBATRACHUS DIDACTYLUS*)

Patterns of male vocalization activity and acoustic interactions in the Lusitanian toadfish (*Halobatrachus didactylus*)

Manuel Vieira a)

Departamento de Biologia Animal and Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa. Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

M. Clara P. Amorim c)

I.S.P.A. Instituto Universitário, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal

Carlos Teixeira d)

Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016

Lisboa, Portugal

Paulo J. Fonseca b)

Departamento de Biologia Animal and Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa. Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

Abstract

Biological rhythms rule many aspects of life, often allowing organisms to anticipate and cope with fluctuations of their internal and external environments. Environmental fluctuations in coastal and estuarine habitats include tidal water level, currents, cyclic variations in turbidity, temperature, salinity and other physicochemical parameters, and day/night related fluctuations. Animals must cope with these fluctuations through physiological and behavioural adaptations. The Lusitanian toadfish breeds in coastal waters and form aggregations. The males use acoustic signals (mostly boatwhistles) for mate attraction and to interact with other males. We used round-the-clock recordings to monitor vocal activity in three locations in a natural habitat in the Tagus estuary: intertidal nests only exposed in spring low tides, subtidal and infralitoral areas (minimum 2.5 m water level). We investigated the relation of the vocalization pattern with habitat parameters such as light, water level and temperature fluctuations. The vocalization patterns differed in the intertidal and subtidal areas relative to the infralitoral. In the periods with higher vocal activity we observed signal interactions where neighbouring males in a chorus either alternated or synchronized boatwhistle emission.

INTRODUCTION

In many species the males form aggregations during the breeding season and vocalize to attract females. In this context males may show chorusing behaviour by vocalizing at the same time. Several advantages for this behaviour were proposed, including improvement in detection of suitable breeding areas (e.g. Muller, 1998), enhancement of mate localization (e.g. Wells 1977), reduction of mate assessment costs and decreased individual predation risk (e.g. Ryan et al. 1981). The major disadvantages being increased competition and the difficulty to discriminate each individual from its neighbours (Gerhardt & Huber, 2002).

Most animals exhibit discontinuous singing, often restricting their acoustic activity to certain periods usually conditioned by environmental (e.g. circadian or tidal rhythms) or biological (e.g. predation risk) constraints. In addition, when vocalising in chorus males have been found to show fine-scale vocal interaction with males such as vocal facilitation or inhibition by the neighbours' sound activity or other patterns relevant to male-male interactions (Gerhardt and Huber, 2002). Indeed many vocal interactions related to male-male assessment or the spacing of individuals in a chorus involve synchrony or alternation of acoustic signals (Gerhardt and Huber, 2002). However, maintenance of strict synchrony or regular alternation is not common in interacting animals. Thus, synchrony has been defined as the overlapping of vocalizations more frequently than expected by chance, and alternation as overlaps occurring less frequently than expected by chance (Greenfield, 1994).

In some species, neighbours actively try to adjust their calling rates to each others' with or without overlapping, and usually leading and a follower male are found (e.g. Naguib 1999). Several explanations exist to the existence of these vocal interactions (alternation and synchrony) but the ability to attract a female is the most common (Gerhardt & Huber, 2002). Although several descriptions of chorusing behaviour exist, the majority involving studies with birds, anurans and insects, information on fish is scarce (Bradbury and Vehrencamp, 1998; Jordão et al. 2012).

In this work we studied patterns of vocal activity in the Lusitanian toadfish (*Halobatrachus didactylus*), a fish that use acoustic signals (mostly boatwhistles; see Amorim et al. 2008 for details of the vocalizations) during mating and territorial defence. Previous studies of the Lusitanian toadfish chorusing behaviour showed that the vocal activity is affected by the social environment. Effects such as increased individual calling rates during chorusing, changes in boatwhistles' duration, adjustments of the calling rate depending on the neighbour vocal activity, and inhibition or facilitation in the presence of other individuals vocalizations' were found (Amorim et al., 2011; Jordão et al., 2012; Vasconcelos et al., 2011). This study describes the vocal rhythms and the interactions of Lusitanian toadfish males in the natural habitat during the breeding season.

MATERIAL AND METHODS

Data collection

We recorded the vocalizations of adult territorial males during the breeding season (May to July 2012 and 2013). The males spontaneously occupied concrete artificial hemicylindrical nests, capped at one end, which we deployed in the Tagus estuary (Air Force Base 6, Montijo, Portugal; 38º42'N, 8º58'W). The males' vocalizations in the lower limit of the intertidal zone were recorded with 16 custom-made hydrophones (Fonseca & Maia Alves, 2011) placed next to each nest in mid-lateral position and about 10 cm above the substrate. These nests were positioned 2 m apart in a line along the shore. In the infralittoral zone we monitored in a similar way two nests placed ca.8 m apart. This arrangement ensured that the sounds of the nest holder male were recorded. In the intertidal zone no other suitable nesting sites were available in the study area. Therefore all vocalizations were produced by fish occupying our artificial nests. Moreover, the sounds produced by males occupying adjacent nests arrived considerably attenuated (over 6 dB). This arrangement allowed ascertaining signal identity throughout the recordings. In contrast, the two infralittoral nests were placed in an area with unmapped natural nesting sites from where other males vocalized. Thus, these two hydrophones recorded sounds from several other fish than the one occupying the deployed nest.

The signal from each hydrophone was recorded by a 16 channel stand alone data logger (Measurement Computing Corporation LGR-5325, Norton, Virginia, USA, 16 bits resolution, 4 kHz sampling rate). Temperature and tidal range were measured with a HOBO® U20 Water Level Data Logger (U20-001-01; Onset Computer Corp., MA.; Range: 0 to 30.6 m; Accuracy: ±0.05% FS, 1.5 cm) positioned into a nest in the intertidal zone at 5 cm above the substrate. We also recorded relative light intensity at the intertidal nests by using a photocell which output was connected to the same data logger used to record the hydrophones.

Signal processing

In the intertidal zone four periods of about two weeks of round-the-clock simultaneous recordings of the 16 nest holders were analyzed. At the infralittoral zone we analyzed 1 month of continuous recordings. From each channel (nest) recording the beginning and the end of each boatwhistle was automatically recognized using an automatic recognition system for call type detection using hidden markov models (HMM) (see Vieira et al. (submitted) for details). The features considered in the automatic recognition system were the combination of cepstrum, Mel-frequency cepstral (MFC), delta and acceleration coefficients of the frequency range between 20 and 500 Hz. This frequency range encompasses most of the spectrum of the toadfish vocalizations and proved appropriate. The system for the recognition of vocalizations was developed using the Hidden Markov Model Toolkit (HTK, University of Cambridge, UK; Young et al., 2002).

DATA EVALUATION

Acoustic rhythms

To investigate the existence of sound production rhythms we measured the overall mean acoustic activity per hour and searched for patterns of vocal activity along the day/night cycles, tidal cycles and temperature fluctuations. Since light intensity may affect behaviour (Davis, 1962), we also compared the vocal activity with the relative light intensity reaching the substrate. The evaluation of possible rhythmic activity was studied both in lower intertidal and infralittoral areas. Using the recordings at the intertidal zone we also compared individual vocal activity with the overall mean acoustic activity patterns. Since vocal activity was strongly variable among individual fish, such comparison allowed to verify how well were individual patterns represented by the overall mean of vocal patterns.

Acoustic interactions

To describe male-male acoustic interactions we focused on pairwise interactions. To assess such vocal behaviour we measured the boatwhistle periods during continuous singing. Continuous singing bouts were defined as vocalization sequences where the boatwhistle periods were shorter than 30 seconds. In these situations we selected all triplets *ABA*, with *A* and *B* representing boatwhistles of two different individuals. For each triplet we transformed the period into a phase angle by dividing the A-to-B by the A-to-A time periods and multiplied the result by 360º (see Fig.9B for a graphic representation). Phase angles are an adequate representation to assess gradients of synchrony to alternation between singing individuals. To study leader or follower behaviours in the male-male acoustic interactions we looked at the association of the period of the boatwhistles of one individual (e.g. interval A-A) with the time lag from a boatwhistle of the same individual to the next boatwhistle of the other individual (e.g. interval A-B; Wickler and Seibt, 1974).

Statistical analysis

Statistical comparisons among mean acoustic activity were performed with Kruskal-Wallis one-way analysis of variance (ANOVA) on Ranks or Mann-Whitney Rank Sum tests. Spearman Rank Order Correlation analysis was used to test for correlations between the mean acoustic activity and water level and temperature.

Statistical analysis of acoustic interactions was made using circular statistic methods. We computed the mean phase angles of boatwhistles from the two fish under analysis and determined the vector strength of the interaction. Then we used the Rayleigh test to assess the significance of the interaction by comparing the vector strength value with a vector strength based on random phase relationships.

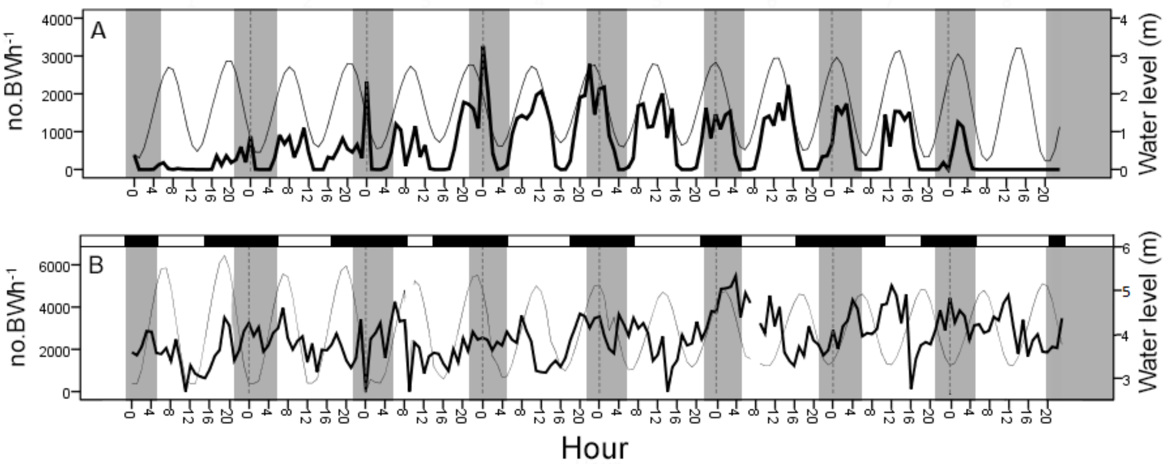

Fig. 1. Overall acoustic activity represented by the number of boatwhistles (line in bold) in the subtidal (A) and infralittoral zones (B). The thin line represent the water level in meters.

RESULTS

Temperature, tidal range and light

Water level and temperature were measure in the nests located at the intertidal zone. Since the site where we recorded fish vocalizing in the infralittoral zone was nearby (ca. 400 m), the water level and temperature values should show similar variations. Water temperature was mostly affected by day-night condition. The water temperature ranged between 11ºC and 27ºC, with daily variations ranging from 1.4ºC to 12.8ºC. Figs. 3 and 4 represent the mean daily variation of water temperature in 2012 and 2013 summers, respectively. Both years present a similar pattern with lower temperatures occurring at dawn and higher temperatures in the afternoon. Notice that the variation of the mean water level is constant during the day if averaging is long enough. Figs. 6 and 7 show the average water level of numerous 12 hour tidal cycles. As expected the water level varies according to a sinusoidal function. Maximum and minimum water levels in the intertidal and infralittoral plots during the recordings were respectively 3.34 and 0.20 m in the intertidal zone, and 5.89 and 2.56 m in the infralittoral zone. Note that recordings in the intertidal plot were only considered for analysis in days where the nests were never exposed to air (neap tides) Light intensity was only recorded in 2013. These data integrate the daylight with the water turbidity, since at night no light reaches the sensor (Fig.1).

Acoustic Rhythms

Were identified 206845 boatwhistles in the intertidal nests from 64 breeding males (42 days with removal of all individuals in each high tide), and 1290673 boatwhistles in the infralitoral zone (26 days without direct human interference). Fig.1 show the overall activity in the intertidal and infralittoral zones through 8 days, with the water level through these days.

Focusing in acoustic activity at the intertidal zone the fish are influenced by the tides (Fig.1). Having no activity at water levels lower that 28 cm. In Fig.2 is represented in detail each individual vocal activity in contrast to the overall vocal activity. Sound production was strongly variable among individual fish but all individuals show the same pattern influenced by the low tides. Fish produced boatwhistles both at day and night, and there was no difference between the absolute numbers of boatwhistles produced (Kruskal-Wallis 1-way ANOVA for ranks of 1 hour, $p = 0.15$). Nonetheless the calling rate during the night is significantly higher (Mann-Whitney-Wilcoxon test, $p > 0.05$; Fig.1 and 3). The daily variation appear to have a lower acoustic activity during the morning and in the afternoon (Fig.3). No significant correlation between the vocalizations and the water temperature was found (Spearman Correlation; $R^2 = 0.027$, $p = 0.4$; Fig.5A). To perceive the tidal cycle influence in Fig. 6 is represented the overall mean vocal activity into the 12 hour tidal cycle (Kruskal-Wallis 1-way ANOVA for ranks of 1 hour, $p < 0.001$), showing an obvious influence of the low water level into the fish vocal activity.

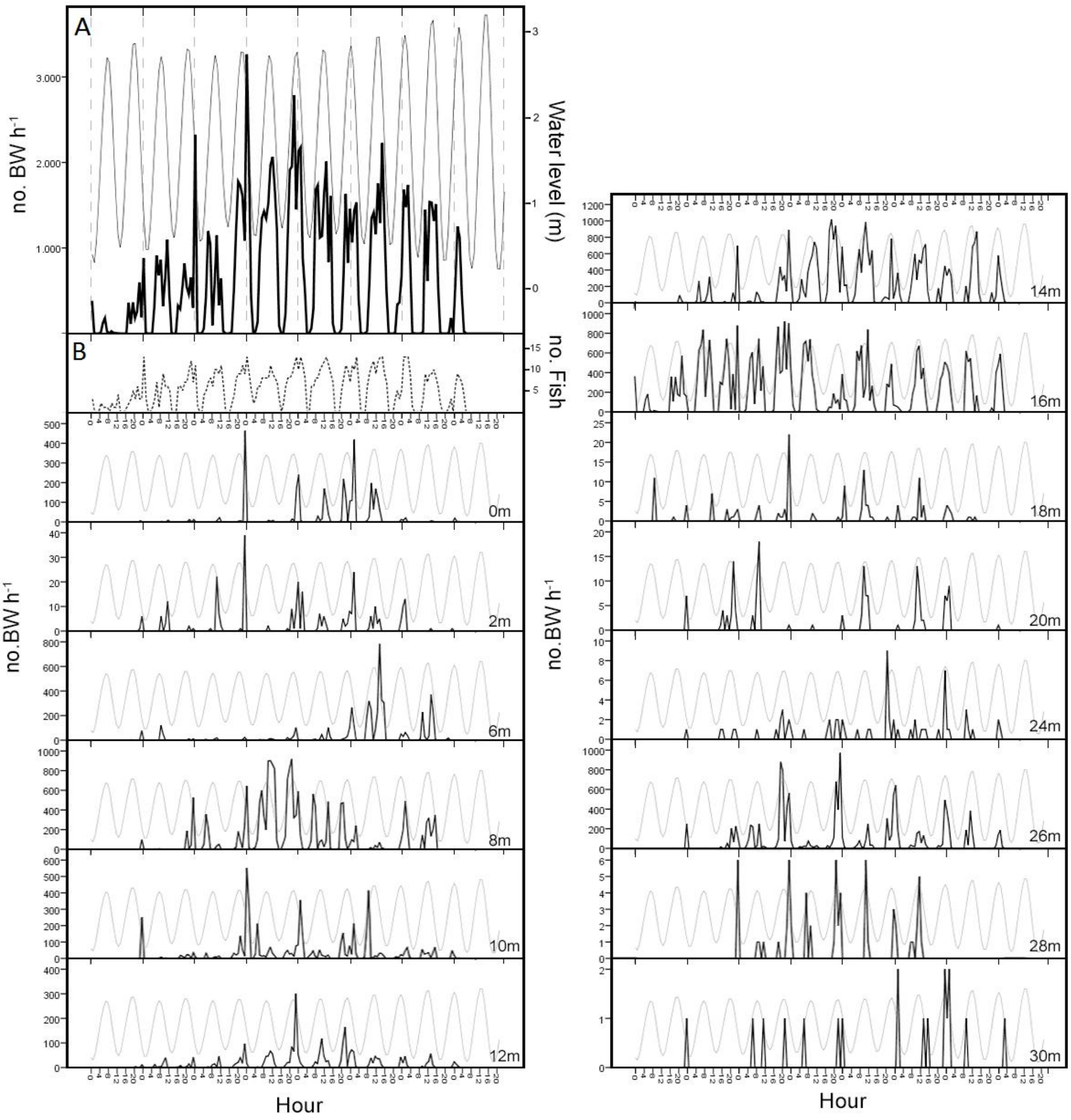

Fig. 2. Acoustic activity in the subtidal zone (number of boatwhistles per hour represented by the line in bold). (A) Overall activity, same graphic as the respresented in figure 1; (B) number of fish producing sounds at each hour; the others graphics represent the acoustic activity of each individual at the 14 distances from 0 to 30 meters (the 2 no represented had no vocal activity). The thin line represent the water level in meters.

In the infralitoral zone the fish also produced boatwhistles both at day and night (Fig.1), but with significant differences in mean overall acoustic activity per hour (Kruskal-Wallis 1-way ANOVA for ranks of 1 hour, $p > 0.001$, Fig.4). With higher calling rates during the night (Mann-Whitney-Wilcoxon test, $p > 0.05$; Fig.4). The highest acoustic activity was during the dawning period (between 6 and 8 a.m.; Fig.4). There was a negative correlation between the averaged number of boatwhistles per hour and the temperature (Spearmen Correlation; $R^2 = 0.414$, $p < 0.01$; Fig.5B). No apparent relation between tidal rhythms and vocal activity was found in the infralittoral zone (Fig.7; Kruskal-Wallis 1-way ANOVA for ranks of 1 hour, $p = 0.99$).

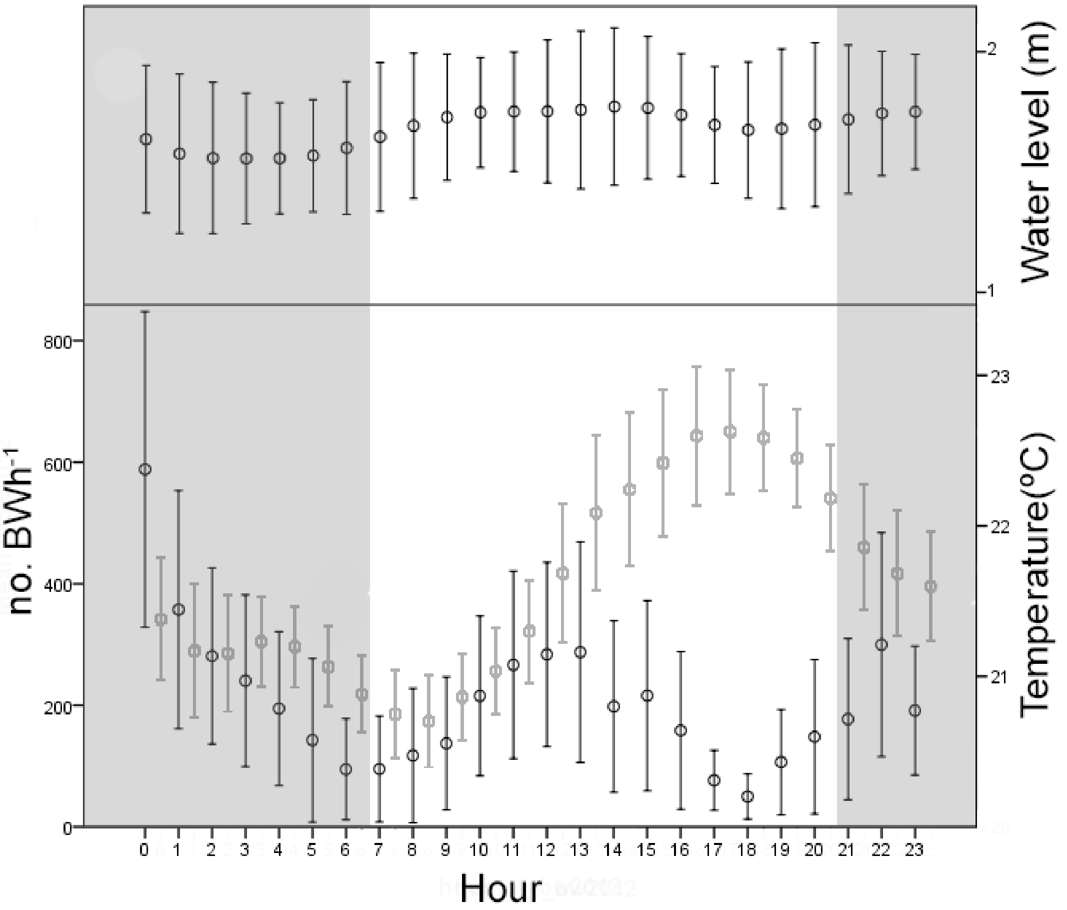

Fig. 3. Mean number of boatwhistles (BW; black bars), temperature (gray bars) and water level by each hour of the day in the subtidal zone. Gray area represent the night. Error bars represent 95% of the confidence interval.

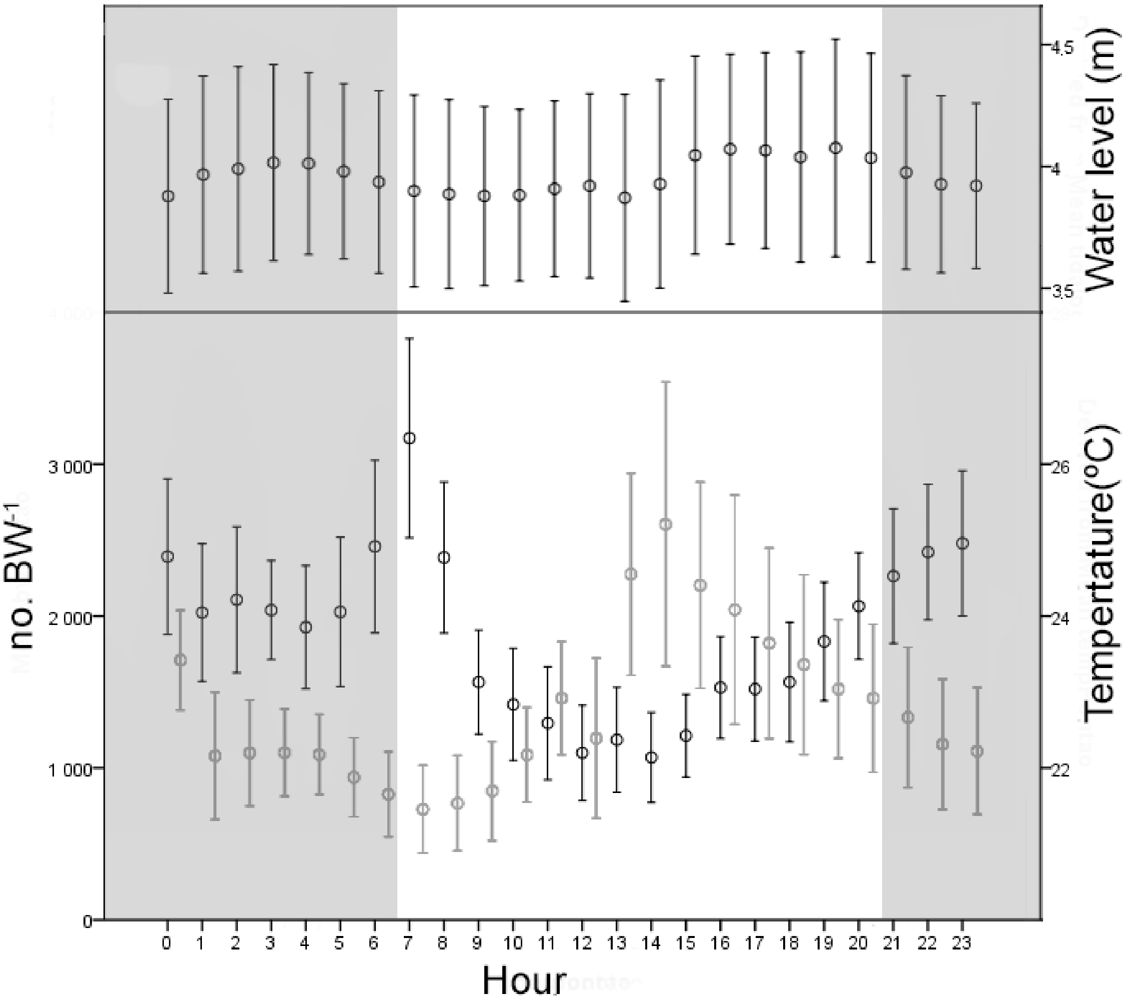

Fig. 4. Mean number of boatwhistles (BW; black bars), temperature (gray bars) and water level by each hour of the day in the infrallitoral zone. Gray area represent the night. Error bars represent 95% of the confidence interval.

Acoustic Interactions

While analyzing the sound recorded on the intertidal zone we saw many cases of alternation and synchrony as represented in the oscillograms in the Fig.8. Alternation is the most frequent pattern. It usually occurred when two or more individuals were singing at the same time. On the other hand, synchrony was usually observed when a fish would produce 3 to 6 boatwhistles overlapping the calls of another fish that was already singing. Fig.9 A represent 30 minutes of simultaneous recordings with boatwhistles of four fish, with of almost 4 minutes in detail. In the oscilogram is represented the recording is an fish in the first nest (zero meters), and then one at 2, 12 and 14 meters. Only these 4 fish were vocalizing during this hour. Fig.9 C show the phase angle of boatwhistles from these 4 individuals during 1 hour (containing the halfhour represented in Fig.9 A). None of the pairwise interactions represented in Fig.9 were considered random (Rayleigh test; $p < 0.05$). The mean phase angle ranged between 157º and 180º, having a visible peak in almost all the represented histograms. In some of the histograms there are a peak at 0º, representing the overlap of calls that are less frequent. The interactions between closer males appear more accurate, with a smaller standard deviation.

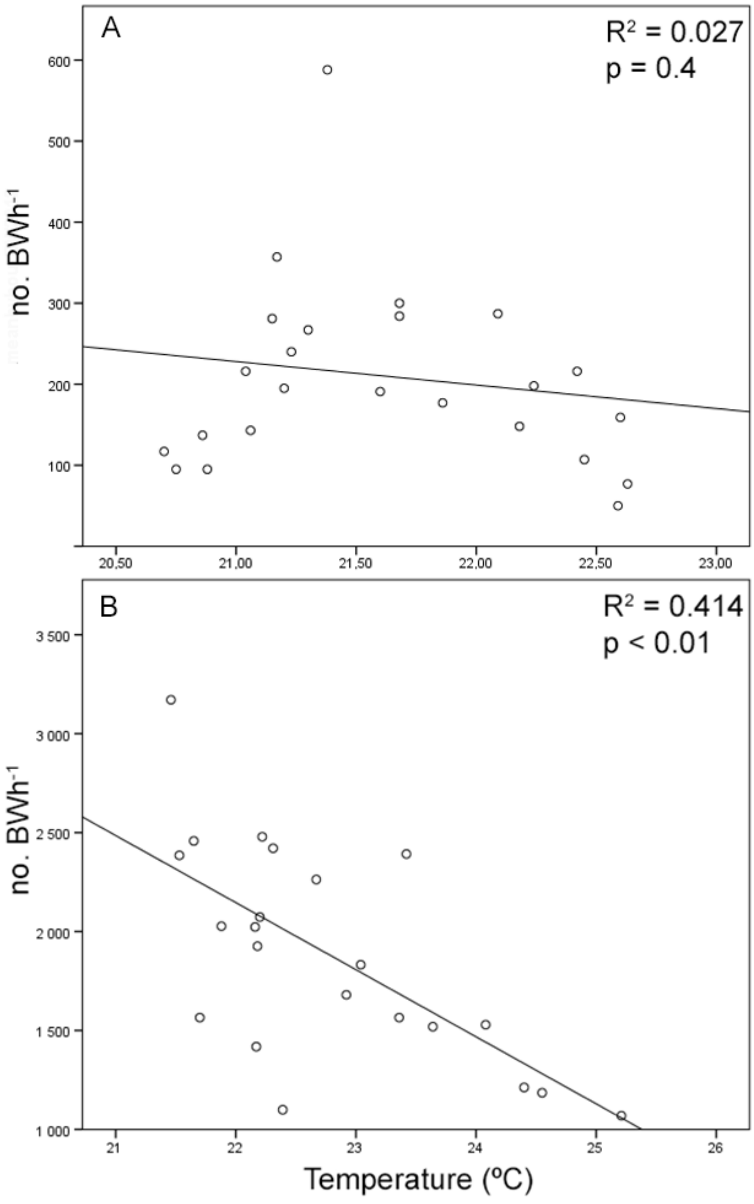

Fig. 5. Scatter plot representing the influence of the temperature in the acoustic activity trough the day (data represented in the Fig. 3 and 4). In (A) intertidal zone and in (B) infralittoral zone. The values of p and R^2 resulted of a Spearman's correlation analysis.

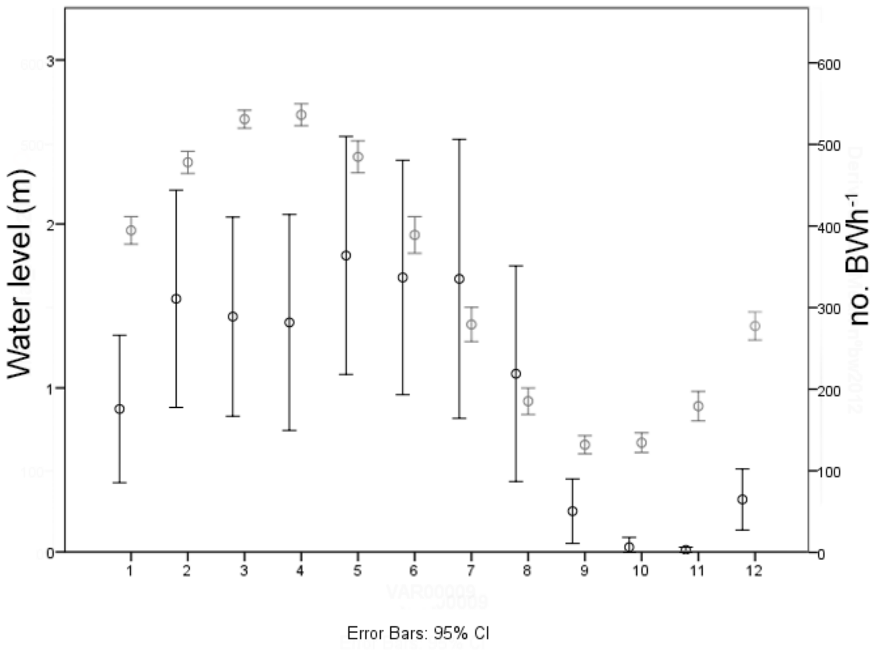

Fig. 6. Mean number of boatwhistles (BW; black bars) and water level (gray bars) by each hour of a 12-hour tidal cicle in the intertidal zone. Error bars represent 95% of the confidence interval.

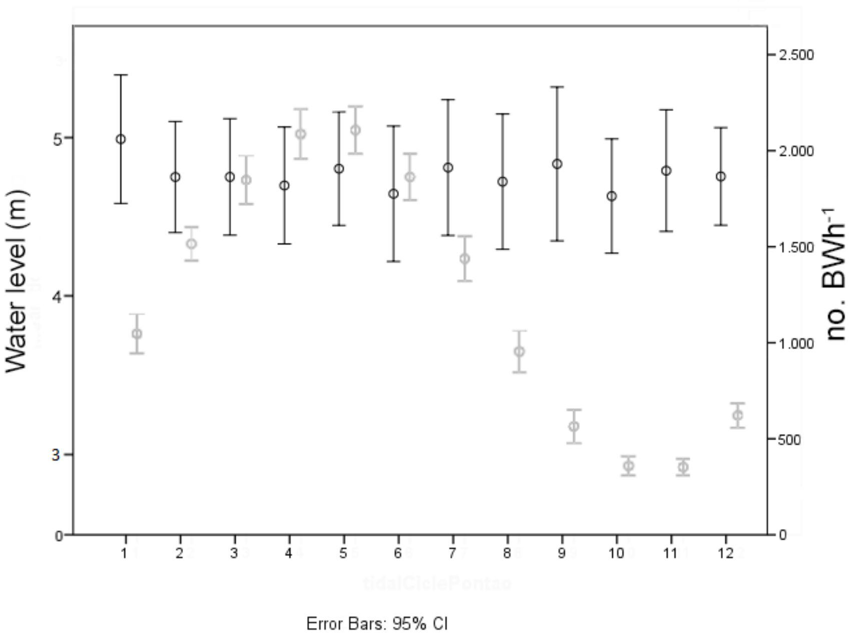

Fig. 7. Mean number of boatwhistles (BW; black bars) and water level (gray bars) by each hour of a 12-hour tidal cicle in the infrallitoral zone. Error bars represent 95% of the confidence interval.

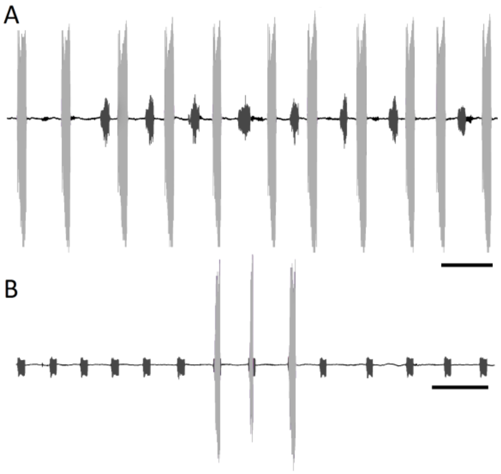

Fig. 8. Examples of the mosts common fine-patterns, (A) alternation and (B) synchrony, in choruzing behavior of a pair of Lusitanian toadfish individuals. Each color (ligther and darker gray) represent a individual, been the lighter gray the individual closer to the hydrophone. The black scale-bar represent 5 seconds.

To have an broader idea of all pairwise interactions in the Fig.10 is represented the number of pairwise interactions with a non-random distribution (Rayleigh test; $p < 0.05$), and random distribution. Fig.10 show that the cases considered random are less frequent and the ones with fewer number of interactions. The mean near 180 is the most frequent, confirming that alternation is the more recurrent type of acoustic interaction in these fish

The Fig.11 shows an representative example of the timing between the calls of two individuals positioned at 2 m from each other. There is an increase of the time between the calls of both individuals with the increase of the calling period of each individual.

DISCUSSION

Our goal was to study the acoustic activity of the Lusitanian toadfish in nature. With this objective in mind we used acoustic recordings of fish both at the intertidal and infrallitoral nests to describe the rhythms while acoustic interactions between males were studied with fish vocalizing in the intertidal nests. Notice that all analyzed recordings in the intertidal zone were made during neap tides, when all nests were always submerged.

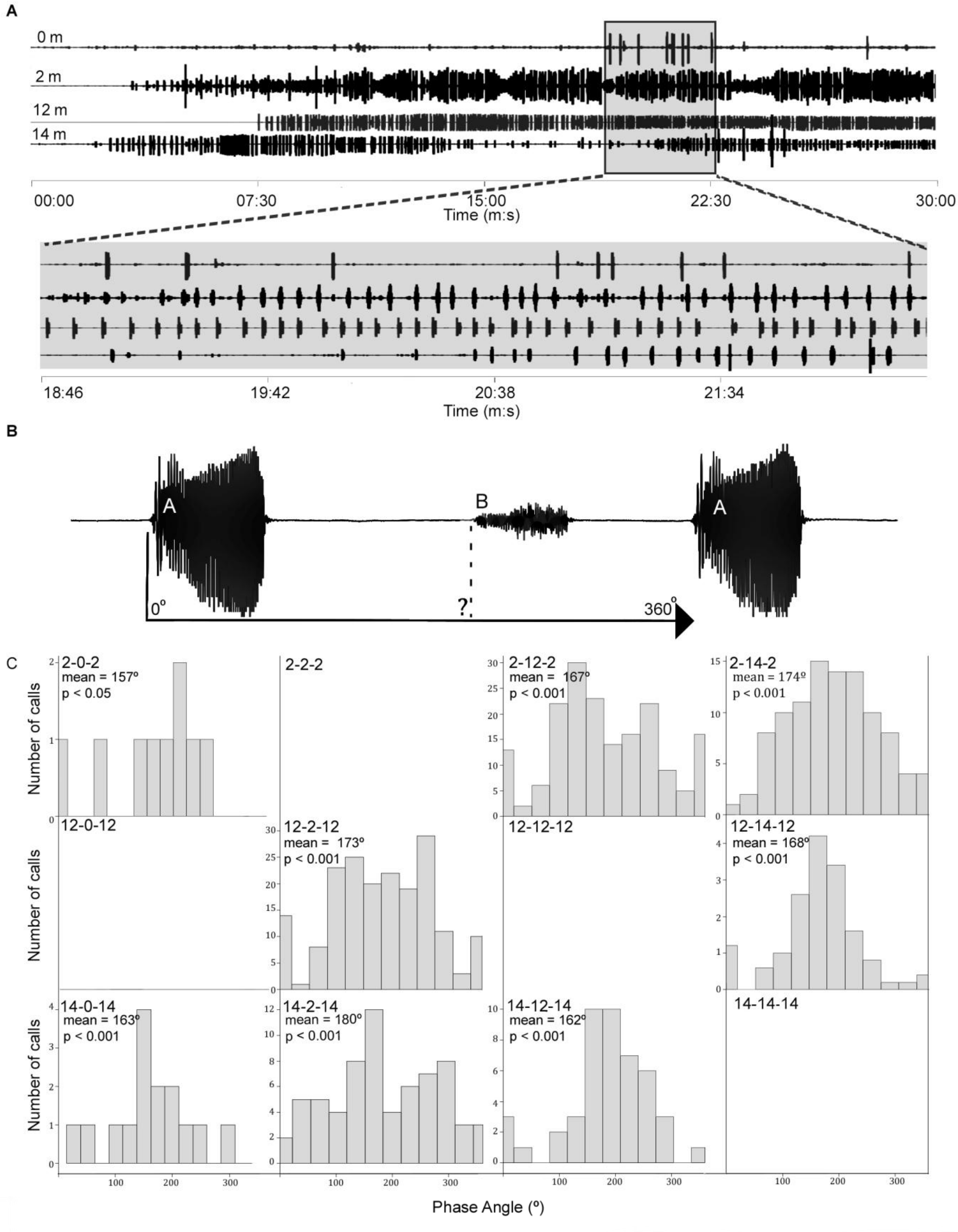

Fig. 9. Duration in seconds of 65 boatwhistles estimated manually, by the individual identification system and by the calll type identification system computed on the MFC with cepstral, delta and acceleration coefficients with a filter compression range of 20 Hz to 2000 Hz. Different letters represent pairwise significant differences ($p < 0.001$) using t-student tests.

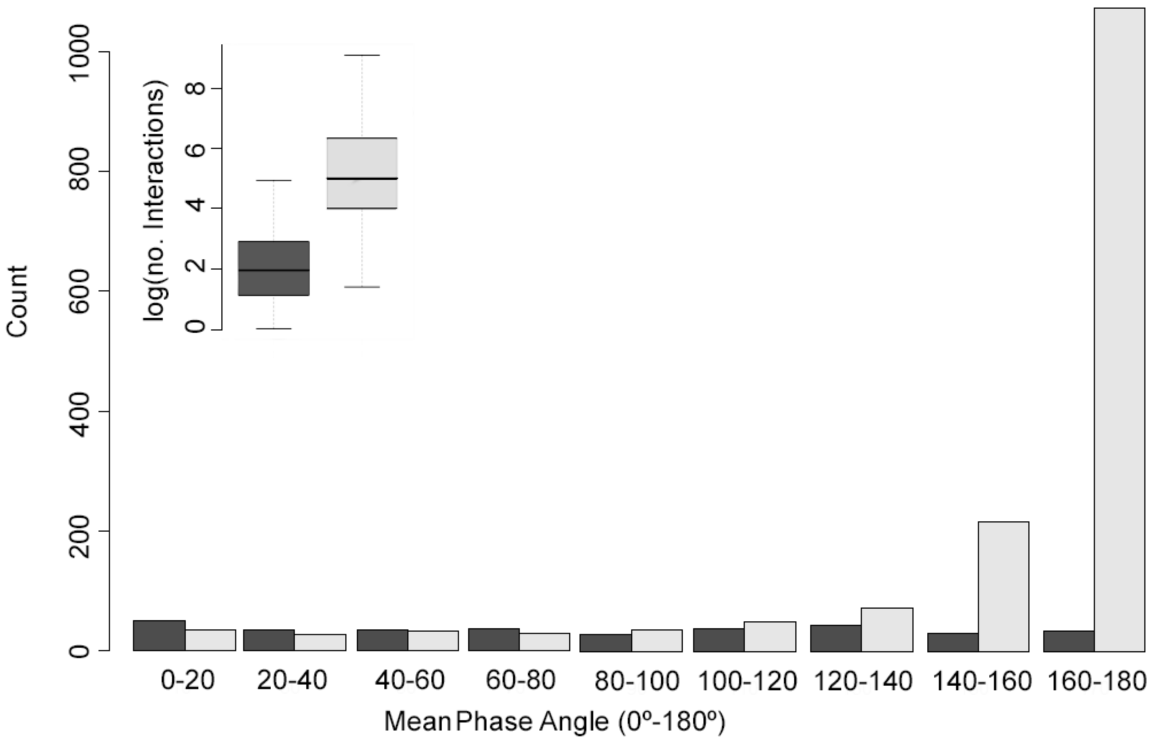

Fig. 10. Histogram representing the means phase angles of the interactions between four groups of 16 individuals. Boxplot representing the logaritm of the number of interaction between each pair. In dark-gray are the interactions cosidered random, and in light-gray the ones considered non-random (Rayleigh test; $p < 0.05$).

Like many other animals, the Lusitanian toadfish appear to have circadian vocalization rhythms with increased activity at night, and peaking at dawn (between 6 and 8 a.m.; Fig. 4). A significant correlation with the daily temperature variations is also observed (Fig. 5). In contrast with these results from the infralitoral recordings, the activity at the intertidal plots, although only considering neap tides, does not appear to follow a clear daily rhythm (Fig. 3). Instead, a marked tidal vocal rhythm was found (Fig.6).

The reason behind the differences between each zone, and the effect in the reproductive success is unknown. We suspect that in the intertidal zone, the effect of the undulations should perturb the fish. Another reason could be some kind of stimulus related to the movements of females into the intertidal zones. In contrast in the infralittoral zone the overall population vocalize continuously, In future works is

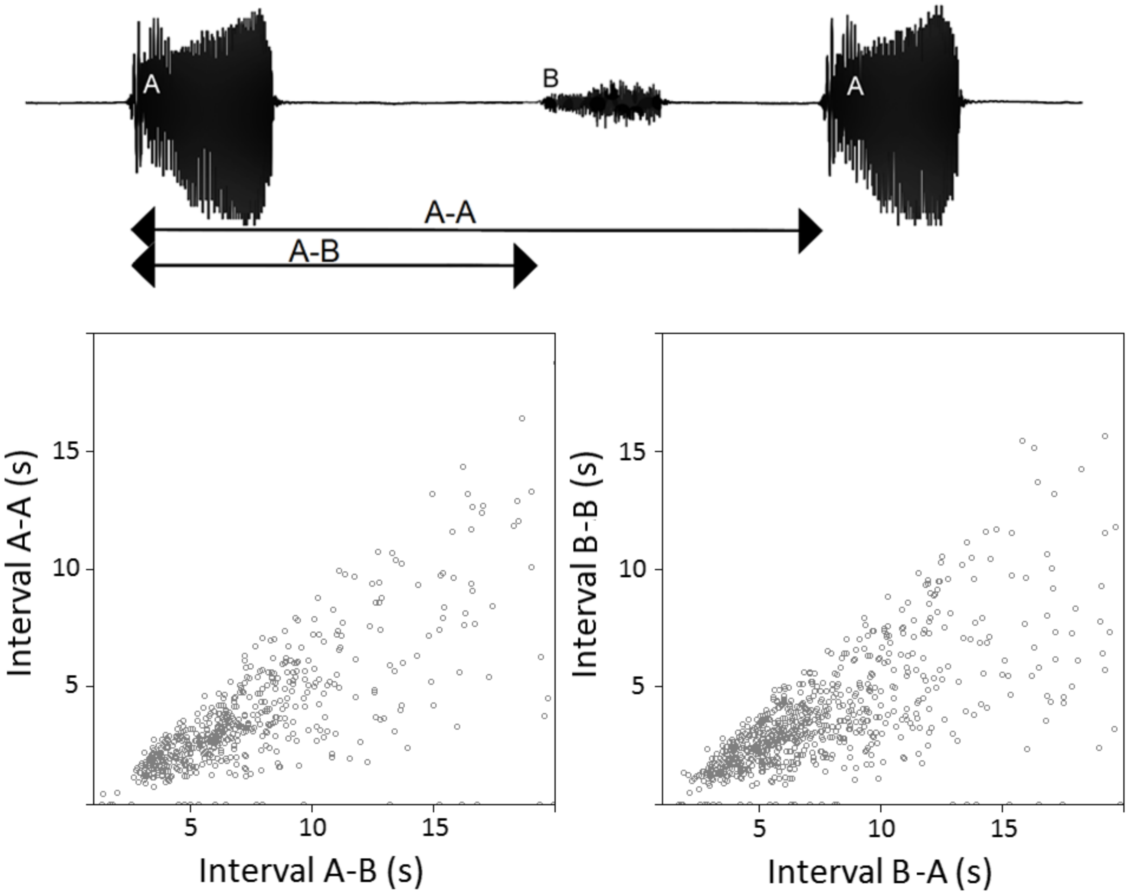

Fig. 11. Timing relationships between one pair of individuals (A and B), showing the call period of one individual (A-A or B-B), against the call delays with the other individual (A-B or B-A) in seconds.

important to analyze the individual activity in the infralittoral zone to characterize each individual acoustic rhythm.

The cause of the differences between night and day acoustic activity, and also the pick at dawn in the infralittoral zone, appear to be related to the temperature variation. Although some other parameters must have some influence. The light intensity is a parameter that could be responsible for the activity at dawn, although in this region there is usually high turbidity and consequently few or no light reach the substrate were this fish usually are. By the other hand the anthropogenic noise could also have some influence having ferry-boats passing by since dawn until midnight.

The breeding males of Lusitanian toadfish must be aware of the vocal activity of the neighbors. They adjust the timing when to sing, having many times a time-locked production of boatwhistles. The most common type of interaction in this fish is the alternation, a chorusing behavior that could facilitate the recognition by the females (Wiley, 1983). There is not followers (Fig. 11), having an apparent adaptation of the calling rate as reported by (Jordão et al., 2012). Apparently there are acoustic interactions even between fish that are positioned at 10 m from each other, suggesting that they perceive the boatwhistlet emitted at such distances. The occasional synchrony situations being usually 3-6 boatwhistles indicate that are the result from an agonistic behaviour. To understand this situations an exhaustive characterization must be done, with the categorization of the boatwhistles as advertising or agonistic calls according to the description of (Raquel O. Vasconcelos et al., 2010).

In conclusion the water level have some influence on the vocal activity in the Lusitanian toadfish, changing the periods of higher activity. When the tidal cycle has less influence the temperature play a important role, that lead to a circadian rhythm. These fish are aware of the other fish, and adjust the timing of calls accordantly having patterns of alternation and synchrony.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology Foundation (projectPTDC/MAR/118767/2010), Portugal.

REFERENCES

Amorim, M. C. P., Simões, J. M., Almada, V. C., and Fonseca, P. J. (2010). "Stereotypy and variation of the mating call in the Lusitanian toadfish, Halobatrachus didactylus," Behavioral Ecology and Sociobiology, 65, 707–716.

Amorim, M. C. P., Simões, J. M., and Fonseca, P. J. (2008). "Acoustic communication in the Lusitanian toadfish, Halobatrachus didactylus: evidence for an unusual large vocal repertoire," Journal of the Marine Biological Association of the UK, 88, 1069–1073.

Bradbury, J. W., and Vehrencamp, S. L. (1998). Principles of Animal Communication, Sinauer.

Davis, R. (1962). "Daily rhythm in the reaction of fish to light," Science, 137, 430–432.

Fonseca, P. J., and Maia Alves, J. (2011). Electret capsule hydrophone: a new underwater sound detector,.

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

Greenfield, M. (1994). “Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions,” American Zoologist, 34, 605–615.

Jordão, J. M., Fonseca, P. J., and Amorim, M. C. P. (2012). “Chorusing Behaviour in the Lusitanian Toadfish: Should I Match My Neighbours’ Calling Rate?,” (V. Janik, Ed.) Ethology, 118, 885–895.

Muller, K. (1998). “The role of conspecifics in habitat settlement in a territorial grasshopper,” Animal Behaviour, 56, 479–485.

Naguib, M. (1999). “Effects of song overlapping and alternating on nocturnally singing nightingales,” Animal Behaviour, 58, 1061–1067.

Ryan, M., Tuttle, M., and Taft, L. (1981). “The costs and benefits of frog chorusing behavior,” Behavioral Ecology and Sociobiology, 8, 273–278.

Vasconcelos, R. O., Carrico, R., Ramos, A., Modesto, T., Fonseca, P. J., and Amorim, M. C. P. (2011). “Vocal behavior predicts reproductive success in a teleost fish,” Behavioral Ecology, 23, 375–383.

Vasconcelos, R. O., Simões, J. M., Almada, V. C., Fonseca, P. J., and Amorim, M. C. P. (2010). “Vocal Behavior During Territorial Intrusions in the Lusitanian Toadfish: Boatwhistles Also Function as Territorial ‘˜Keep-Out’ Signals,” Ethology, 116, 155–165.

Wells, K. (1977). “The social behaviour of anuran amphibians,” Animal Behaviour, 25, 666–693.

Wickler, W., and Seibt, U. (1974). “Rufen und Antworten bei Kassina senegalensis, Bufo regularis und anderen Anuren,” Zeitschrift fuer Tierpsychologie, 34, 524–537.

Wiley, R. (1983). “The evolution of communication: information and manipulation,” Animal behaviour, Vol. 2: Communication, Oxford: Blackwell Publishers, pp. 156–189.

Young, S., Evermann, G., Kershaw, D., and Moore, G. (2002). “The HTK book (for HTK version 32),” Cambridge University …, Retrieved from http://scholar.google.pt/scholar?hl=pt-PT&q=young+htk&btnG=&lr=#1.

. Young, S, Evermann, G., Kershaw, D., & Moore, G. (2002). The HTK book (for HTK version 3.2). Cambridge University Press.

DISCUSSION

Conservation and fisheries management require species-specific, age-based population estimates and the knowledge of their distribution, reproductive patterns and behaviour (Mann et al. 2008). Traditional methods of surveying fish use nets or acoustic surveys (active acoustics) to sample populations but the former is destructive to the target species and habitat, and the latter, although effective, presents limitations in the temporal and spatial scale of data collection (Mann et al. 2008). Sound production in vocal fishes is species-specific and is related to territoriality, courtship and spawning behaviour (Amorim 2006), allowing the use of passive acoustics to locate, for example, spawning aggregations of target species (e.g. Luczkovich et al. 2008; Rowell et al. 2012). With the use of this HMM automatic recognition approach we make possible the study of fish ecology and behaviour using round-the-clock recordings (e.g. assessment of the reproductive success, study rhythms and acoustic interactions, estimating anthropogenic impacts), acquiring the knowledge needed in a cheaper, faster and less invasive way. Is important to refer that the Mel-frequency cepstrum (MFC) and the perceptual linear predictive (PLP) although developed to study human speech recognition, perform well extracting the information needed for an accurate recognition of fish vocalizations.

In the chapter 2 we show an example of what can be done using this recognition system. In this case with the analysis of the overall acoustic activity of several fish, we observed an influence of the day-night cycles and tidal cycles in the vocal activity. Furthermore, we described the interaction between males. The breeding males are aware of the acoustic activity of their neighbors, establishing accurate interactions where males who are close to each other in a chorus either alternated or synchronized their vocalizations. This knowledge is something that might be crucial in the future to understand e.g. the importance of such interactions with the reproductive success.

REFERENCES

Amorim, M. C. P., Simões, J. M., & Fonseca, P. J. (2008). Acoustic communication in the Lusitanian toadfish, Halobatrachus didactylus: evidence for an unusual large vocal repertoire. *Journal of the Marine Biological Association of the UK*, *88*(05), 1069–1073.

Ladich, F. (2004). Sound production and acoustic communication. *The Senses of Fish* (pp. 210–230). Springer Netherlands.

Vasconcelos, R. O., Carrico, R., Ramos, A., Modesto, T., Fonseca, P. J., & Amorim, M. C. P. (2011). Vocal behavior predicts reproductive success in a teleost fish. *Behavioral Ecology*, *23*(2), 375–383.