HAL
open science

Use of recurrence plots for identification and extraction of patterns in humpback whale song recordings

Franck Malige, Divna Djokic, Julie Patris, Renata Sousa-Lima, Hervé Glotin

To cite this version:

Franck Malige, Divna Djokic, Julie Patris, Renata Sousa-Lima, Hervé Glotin. Use of recurrence plots for identification and extraction of patterns in humpback whale song recordings. Bioacoustics, 2020, pp.1 - 16. 10.1080/09524622.2020.1845240 . hal-03008908

HAL Id: hal-03008908

https://hal.science/hal-03008908

Submitted on 3 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH REPORT

Franck MALIGE, Universit d'Aix Marseille, Univ de Toulon, LIS quipe DYNI
Divna DJOKIC, Laboratory of Bioacoustics (UFRN), Natal, Brazil Graduate Program of Psychobiology, Biosciences Center, UFRN, Natal, Brazil
Julie PATRIS, Universit d'Aix Marseille, Univ de Toulon, LIS quipe DYNI
Renata SOUSA LIMA Laboratory of Bioacoustics (UFRN), Natal, Brazil
Herv GLOTIN, Universit d'Aix Marseille, Univ de Toulon, LIS quipe DYNI

November 2020

Use of recurrence plots for identification and extraction of patterns in humpback whale song recordings

A published version of this work in "bioacoustics" is available at
"F. Malige , D. Djokic , J. Patris , R. Sousa-Lima H. Glotin (2020): Use of recurrence plots for identification and extraction of patterns in humpback whale song recordings, Bioacoustics"

Use of recurrence plots for identification and extraction of patterns in humpback whale song recordings

F. Malige[a], D. Djokic [b,c], J. Patris[a], R. Sousa-Lima [b,c] and H. Glotin [a]

[a] Université de Toulon, Aix-Marseille Université, CNRS, LIS, DYNI team, Marseilles, France
[b] Laboratory of Bioacoustics, Department of Physiology and Behavior, Universidade Federal do Rio Grande do Norte (UFRN), Natal, Brazil
[c] Graduate Program of Psychobiology, Biosciences Center, UFRN, Natal, Brazil

ARTICLE HISTORY
Compiled December 3, 2020

ABSTRACT

Humpback whale song is comprised of well-structured distinct levels of organization: combinations of sounds, repetition of combinations, and a sequence of repetitions, which have no clear silent intervals. This continuous sound output can be hard to delimit, rather, it could be interpreted as a long series of states of a system. Recurrence plots are graphical representations of such series of states and have been used to describe animal behavior previously. Here, we aim to apply this tool to visualize and recognize structures traditionally used in inferences about behavior (songs and themes) in the series of units manually extracted from recordings of humpback whales. Data from the Abrolhos bank, Brazil were subjected to these analyses. Our analytical tool has proven efficient in identifying themes and songs from continuous recordings avoiding some of the human perception bias and caveats. Furthermore, our song extraction is robust to errors coming from both manual and automated transcriptions, constructing a level of description largely independent of the first stage of analysis.

HIGHLIGHTS

- The complex hierarchical and cyclical structure of humpback whale song can be visually represented in recurrence plots
- Recurrence plots of humpback whale singing provide insights into the intra-individual variation in song structures of a male
- Our tool can help extracting structures, from long recordings, removing some human perceptual caveats

KEYWORDS
Humpback whale, song, recurrence plot, song structure, sound visualization, sound transcription, sound sequences

1. Introduction

Songs usually describe sequences of sounds showing some structure. Complex strings or sequences of sounds have evolved in many taxa (Kershenbaum et al. 2014) and animals

CONTACT F.Malige. Email: franck.malige@lis-lab.fr

may sing with varying degrees of complexity. Depending on the research question, one may or may not consider the adjacent acoustic context in which such songs, or sound sequences, are delivered, i.e., some are separated by silent intervals, but some are not. These long and continuous complex sequences of sounds impose the added challenge of limiting when one biological meaningful sequence ends, and another begins. The humpback whale male song is an example of such an animal acoustic output structure which is very hard to characterize due to the issues just described.

1.1. Structure of humpback whales songs

Schreiber (1952) was the first to describe sounds recorded in the ocean by the U. S. Navy in 1951 which were later attributed to humpback whales, *Megaptera novaeangliae* (Schevill and Watkins 1962). Nonetheless, the complex structure of the humpback whale acoustic display - the song - was noticed almost a decade later by Katy and Roger Payne in 1969 and formally published by Payne and McVay (1971). Three decades after that, genetic confirmation (Darling and Brub 2001) supported the behavioral and morphological evidence that only males sing (Darling (1983), Glockner (1983)).

According to Payne and McVay's description, singing males emit sound units that are arranged in phrases that are repeated to form a theme. Themes are sung in a fixed order which is a song, and a song session is the continued rendition of the song. Constant changes in the song throughout the singing season, called song evolution characterize the dynamic of singing activity in humpback whales (Payne et al. 1983). This hierarchical song structure that cycles in a fixed order was revised by Cholewiak et al. (2013) to incorporate multi-level variation in song structure and to address some caveats with the original structure and order proposition. Fundamental differences between bird song literature that inspired Payne and McVay (1971), such as the lack of silences between song renditions, hindered the acknowledgement that boxing humpback whale song into static artificial hierarchical levels was potentially misleading inferences (Cholewiak et al. 2013). Specially complicated is to arbitrarily define limits of songs and themes that would vary depending on who was describing it.

Methodology on humpback whale song elements identification (and extraction) is still advancing, as consensus on the best protocols are not yet established. As the song and its elements vary in length and order (Cholewiak et al. 2013), it is challenging to manually identify them (and their limits) without subjectivity. Automating the process of song structure recognition would significantly reduce the frequent human bias. As a mean of going a step closer to this goal of automation, we propose a method based on the "unit" level of the vocalization. Unit is the best defined element of the humpback whale song, and can be described as "the shortest sound entity recognized by the human ear, separated from other sounds by a short period of silence" (Payne and McVay 1971). Even though classifying units can be a tricky process since there is some versatility in the renditions of the same unit, yet as an entity, unit remains the most unambiguous element of the humpback whale song hierarchy.

In order to reliably define patterns of higher hierarchical level - themes and songs, as assemblies of units, we propose a new method to remove some of the human influence in defining the start and end points of songs. This is done by adopting a semi-automated protocol to detect, analyze or extract the aforementioned features from transcribed unit label strings.

1.2. Recurrence plots

Recurrence plots are used to visualize and analyze, at a global level, long series of states of a system (see definition by Eckmann et al. (1987) in the case of a general dynamical system). This tool and its graphical representation have been used in several scientific topics : first in medicine in Zbilut et al. (1990) and then in astronomy, neuroscience, mechanics, geology, climate changes (see review by Marwan et al. (2007)). This tool has recently been proposed to study structures in animal movements or communication in Ravignani and Norton (2017). It was used in acoustics - monitoring of air guns (Miralles et al. 2015), and bioacoustics - shrimps sound production (Hee-Wai et al. 2013). A closer application to our problem of analyzing humpback whale songs has been to visualize structures of a music extract (Foote 1999) or to cut it automatically as in Foote (2000) or Paulus et al. (2010). It has recently been used to study the rhythm of humpback whale sound production in Schneider and Mercado-III (2018), without focusing on the spectral content of the sounds.

The main topic of this paper is to apply this tool to visualize and recognize the main structures (songs and themes) in unit series of humpback whales. We apply this method to data taken in Abrolhos bank, off the northeastern coast of Brazil (see section 2.1). Theses recordings were manually transcribed into a string of units named as letters (see section 2.2). On this input, a matrix of distances, based on the Levenshtein distance is computed as done in recurrence plots (see section 3). A method for the automatic extraction of songs in the series of units is proposed (see section 4) and tested on our data set in the final section 5.

2. Data collection and sound units transcription

2.1. Data collection

Data was collected in the Abrolhos bank, located off the Northeastern coast of Brazil (17^oS and 38^oW) where humpback whales come during the austral winter and which is considered the main calving grounds for the species in the western south Atlantic Ocean (Martins et al. (2001) and Andriolo et al. (2006)). During 2000 and 2001 research cruises, groups of humpback whales were sighted and monitored for acoustic activity using one HTI 90 min hydrophone connected to a portable DAT Sony TCD D-10 (sampling rate 48 kHz). Vocalizing males were then identified and located by monitoring the amplitude decrease of an individual's sounds as it surfaced to breath. Silent approaches to these focal singers were performed using a small zodiac with an improvised sail. Songs were collected from the zodiac at distances that varied from 100-500m to the focal male while its behavioral activity was continuously registered. Several recordings were made but only the best quality ones were used in our analyses. Selected recordings made in September of the year 2000 generated 3 high quality audio files (recordings #1,2,3) lasting respectively 58 minutes, 1 hour 37 minutes and 2 hours and 7 minutes. In September of 2001 the two selected audio files (recordings #4,5) were respectively 26 minutes and 1 hour and 38 minutes long. Another study of song sessions in the same place in 2000 will serve as a comparison for the data we present (Arraut and Vielliard 2004).

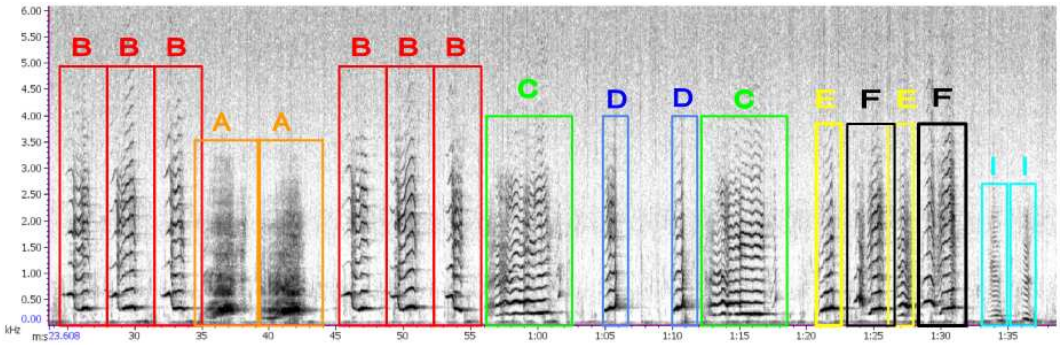

Figure 1. Time (min:sec) / frequency (kHz) representation of an extract of recording #3 (from Abrolhos Archipelago 2000), 1024 point FFTs, Hann window, 43.1 Hz resolution and 50% overlap. Unit transcription into letters is signaled above the boxes

2.2. Sound unit transcription

Classifying units into specific types is a complex problem, since there is a large versatility in the acoustical properties of a single unit (Janik 1999). First, there is a seemingly infinite number of different unit types used by whales, in an ever-changing song. Second, even the units belonging to the same type may vary to a certain level throughout the same song, in the recordings of different singers, and depending on the quality of the dataset. Thus, a different final product of classification can arise from the same dataset due to differences in the methods used to determine the units. The final number of different unit types in different studies varies greatly, from twelve to more than one-hundred (Pines 2018). Challenge to group the units by type (as similar or different), opens the debate on what is the acceptable level of variation within the same unit type.

In order to overcome this difficulty, we used the context i.e. the position of the unit in the song (or more commonly in the phrases within each theme), defined by the arrangement of adjacent units (Green et al. 2011). The context has proven as a good way to help determine the unit type (Cholewiak et al. 2013). For this work, all recordings were inspected by hand. Spectrograms were created using 1024 point FFTs, Hann window, 43.1 Hz resolution and 50% overlap (created by software Raven pro 1.5 - Cornell Lab of Ornithology, Ithaca, NY, Program (2014)), and subjected to aural and visual inspection by two trained analysts. Each analyst separately inspected and labeled the whole dataset. This procedure is a common protocol adopted by humpback whale song researchers (Payne and McVay (1971), Darling et al. (2019)). Our protocol for unit classification also takes context into consideration, i.e. where the unit is placed, so that other hierarchical levels of organization can help define how much variation is allowed in a single unit type. Note that even when studies carry on unit annotation using computational methods, the results are usually validated using manual classification (Pace et al. (2010), Garland et al. (2013), Allen et al. (2017)).

The labeling was done in the following way: every unit type was attributed an alphabet letter based on its distinctiveness from other units (Darling et al. 2019). Analysts considered visually and auditory perceptual characteristics for classification such as unit tonal or pulsed quality, its pitch content and frequency modulation pattern, its duration, and its placement within phrases (context). Each time the specific unit would arise in the recording, it was labeled in Raven, according to its type, minding its context (adjacent units). The final product of every separate recording would be a list

of N consecutive letters, the way the units appeared in that specific humpback whale vocalization (see figure 1 and supplementary material #1). This list of letters (units) will, in the later steps of the method, serve as an input for computing recurrence plots (section 3). Finally, an inexperienced analyst, using another software (Audacity, using a 8192 points FFT, Hanning window, 99% overlap) independently transcribed part of recording #3 into a series of units as a comparison with the first transcription and a check on the robustness of our method (see section 5.2).

3. The Levenshtein distance recurrence plots

3.1. Definition of the Levenshtein distance recurrence plot of order n

To visualize the structures contained in a recording, we compute a distance matrix, using OCTAVE (Eaton et al. 2009) in the following way. Each recording is transcribed as a string of N letters representing sound units (see section 2.2). Then we define $extract_i$ as an n-letter extract beginning with unit number i. The length of the extracts n is taken much smaller than the total length of the string N. The Levenshtein distances $d_{ij} = Levenshtein\ distance\ (extract_i, extract_j)$ between all pairs of n-letters extracts are computed.

The Levenshtein distance between two strings is the minimum number of insertions, deletions and substitutions necessary to transform one string into the other (Levenshtein. 1965). Note that the computation of other types of distance matrix could be achieved choosing other distances between two strings of letters : Jaro-Winkler, Damerau-Levenshtein and Hamming distances for example.

For convenience, we defined a correlation index by :

$$c_{ij} = \frac{n - d_{ij}}{n}$$

Thus the coefficients of the matrix satisfy $c_{ij} \in [0; 1]$. The coefficient 0 means maximum Levenstein distance and thus minimal correlation. The coefficient 1 means zero Levenstein distance and thus maximal correlation.

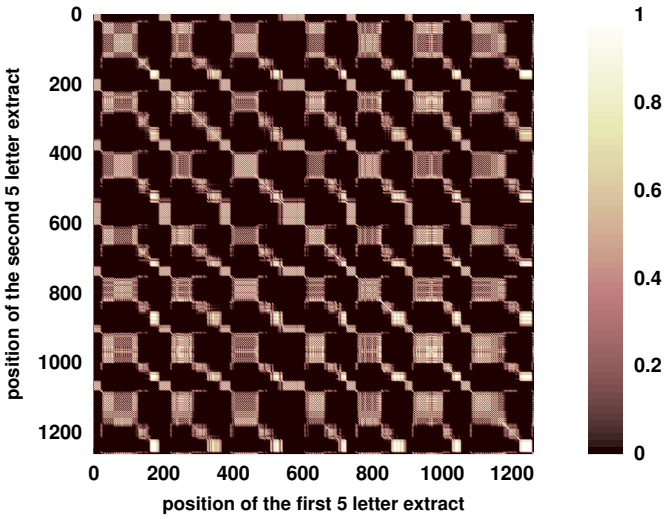

Figure 2. Example of a Levenshtein distance recurrence plot (5 letter extracts) created form recording number #2 from Abrolhos Archipelago (2000). The number of units of this recording is N = 1260. The coefficients of the matrix denote correlation between 5 letter extracts (0 is no correlation, 1 is maximal correlation)

The resulting matrix (c_{ij}) is a square matrix of size $N - n + 1 \simeq N$, symmetrical, and has ones on its diagonal (maximal correlation between one element and itself). As an example, figure 2 shows the graphical representation of this Levenshtein distance matrix of recording #2 (a total of N =1260 units), with 5-letters extracts (n=5).

The result is a **recurrence plot** of a dynamical system (as defined in Eckmann et al. (1987)) in which a state is a vector of n letters. We call this special recurrence

plot the Levenshtein distance recurrence plot of order n ($LDRP_n$).

The value of n is chosen by the analyst. In our data, the visual information is basically the same when n varies (see figure 3). For low values of n, there are few levels of Grey. For high values of n, the resulting figure is more blurred (see figure 3), and the computation of Levenshtein distance is rather time consuming (the time of computation of the distance between two strings of n letters is proportional to n^2 (Wagner and Fischer 1974)). The computation of the $LDRP_n$ for $n = 17$ and $N = 800$ in figure 3, takes around one hour in a domestic computer. The parameter n can be adapted depending on the use of the $LDRP_n$: visual analysis, extraction of structures, etc...

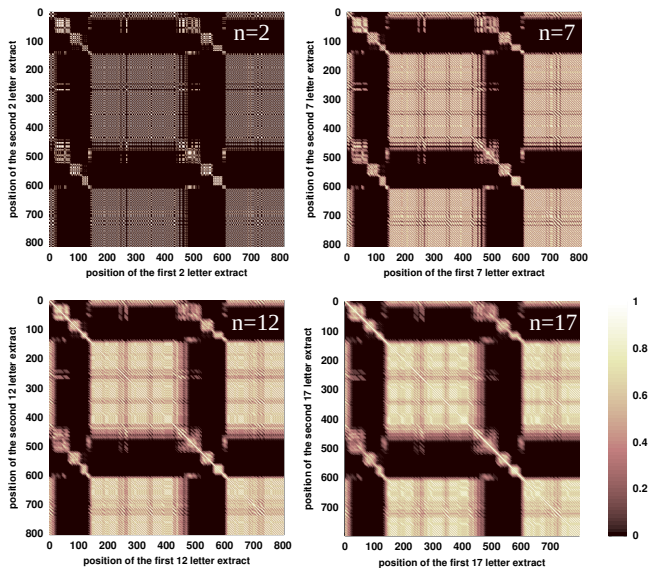

Figure 3. Recording #5 (from Abrolhos Archipelago 2001) : Levenshtein distances recurrence plots for different values of the length n of the extract (2,7,12 and 17)

3.2. Visual identification of main structures in the recurrence plot

Figures 2, 3 and 4 show obvious structures in the recording. On the diagonal of the matrix, squares represent auto-similar structures. Rectangles out of the diagonal show whether theses structures correlate with each other. The high contrast of this representation is due to the sparsity of input representation (letters) added to the efficiency of the distance operator.

In figure 4, we annotated the two different scales of structures found in all recordings of this study. First, a global pattern can be seen that is reproduced periodically (the "period" can be seen in the horizontal or vertical regularly spaced correlations of parts of the recording). We thus define a **song as**

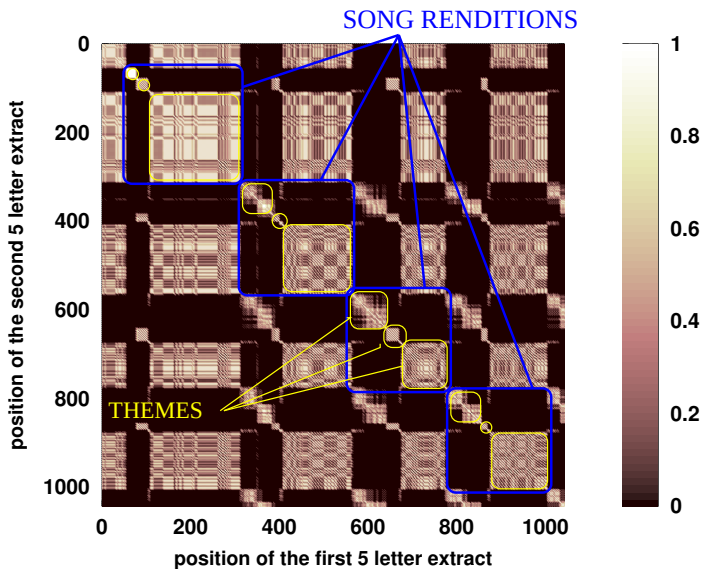

Figure 4. Recording #1 from Abrolhos archipelago, 2000 : Levenshtein distances recurrence plot commented (n=5). The main structures of humpback whale sound production appear : songs in blue, themes in yellow.

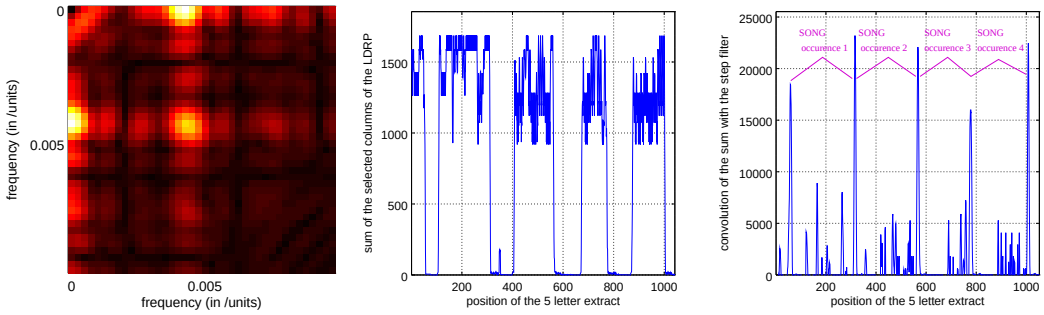

Figure 5. Steps in the extraction of the songs. **Left** : Step 2 of the algorithm, extract of the 2D fast Fourier transform of the Levenshtein distance matrix (for $n = 5$, recording #1). A peak is visible at a frequency of about 0.045 units^{-1} (which means a scale of 200 to 250 units of the recording) **Center** : Step 3 of the algorithm, the horizontal sum of selected columns of the recording 1, n=5. **Right** : Step 4 of the algorithm, each peak shows the transition between two songs.

the largest repeated structure that can be found in a recurrence plot. The song is repeated with a high level of similarity, apparent in the non-zero correlation with other songs. This structure is obvious in our representation, it is also in accordance with the general literature on humpback whale bioacoustics (Payne and McVay (1971), Mercado-III et al. (2003) and Cholewiak et al. (2013)).

A second scale of structure is visible in the $LDRP_n$: each song is itself composed of several auto-similar parts or 'squares' (figure 4). These parts are not repeated within a song (square 1 does not correlate with others during the same song). We define a **<u>theme</u> as the string of units corresponding to a "square", or self-similar part of a song**. With this definition, we are consistent with general literature on humpback whales (Mercado-III et al. 2003) which states that a theme is a repetition of similar phrases composed of similar sound units. Nevertheless, it is difficult to see distinctly visual signature of phrases in the recurrence plot. Anyway, the separation of themes in squares could help a humpback whale song analyst to identify phrases in each series of units corresponding to a square.

4. Automation of the extraction of songs

Based on the computation of the distance recurrence plots, we wrote a routine that separates songs in a fully automatic way. We defined the song as a structure that is repeated with a good amount of similarity, with the largest possible scale in our recording. The approximate size (in number of units) of a song can be estimated by a 2 dimensional FFT of the Levenshtein distance recurrence plot of order n (see figure 5, left). Then, the following algorithm for song extraction is proposed:

(1) Compute the Levenshtein distance recurrence plot for a given value of n
(2) Perform a 2D Fast Fourier transform of the matrix considered as an image. A peak at a frequency inverse to the mean length of the song can be seen (figure 5, left). The abscissa of this peak is then measured to obtain an order of magnitude of the lengths of the songs.
(3) Read the first line of the Levenshtein distance recurrence plot and select the columns with non-zero correlation. Sum all these columns to improve the signal to noise ratio. The resulting column is represented on figure 5, center.

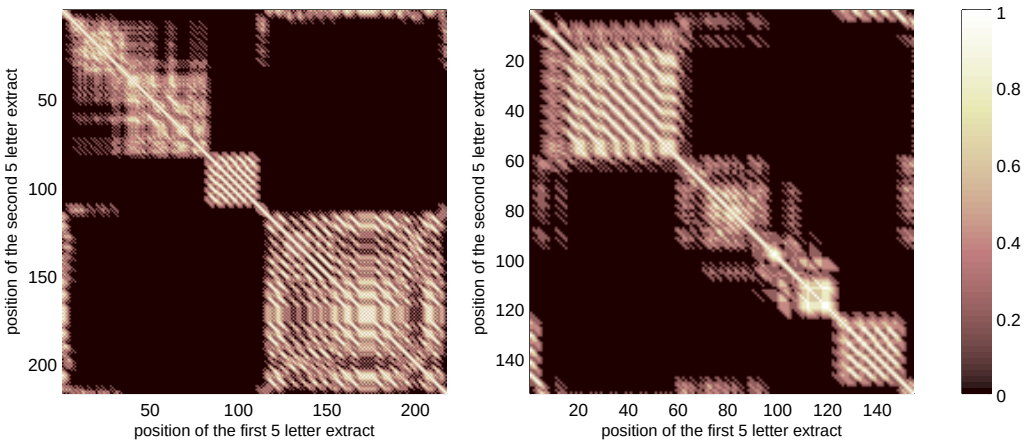

Figure 6. Two songs automatically extracted. **Left** : Recording #1 from Abrolhos archipelago, 2000 : third song automatically extracted from figure 4. In order to see clearly the song, the frame of the figure is chosen a bit larger for representation. **Right** : Recording #2 from Abrolhos archipelago, 2000 : zoom on the third song occurrence in the Levenshtein distance recurrence plot of figure 2. Different types of transition between themes appear.

(4) Find sudden increases or upward steps in the resulting column using convolution with a step filter (figure 5, right). Keep only the steps compatible with the average song size predicted by the 2D Fourier transform. Here, we defined the margin of acceptability at one third of the average size given by FFT analysis.
(5) Extract the units transcribed from the recording between two upward steps: this is a song (figure 6).

5. Results : song extraction

5.1. Extraction of songs in series of units from Abrolhos Archipielago in 2000/2001

We applied the algorithm presented in the previous section to the five recordings of our study. Visually, the same features corresponding to our definition of songs and themes appear in each Levenshtein distance recurrence plot. The recording #4 is too short and do not present a complete song and our routine consequently did not extract any song. In all the other recordings, the extraction of songs was achieved successfully. In recording #5, only one song is present.

We compared the extraction of songs carried out by our routine, based on the Levenshtein distance recurrence plot, and songs delimited by a human expert. Comparison was performed for the recording #3. The analyst followed the instructions described in Cholewiak et al. (2013), the choice for the beginning point of the first song was the first complete theme (composed of the same type of letters). The extraction of the songs (presented in supplementary material #2) is equivalent by both methods. The only difference sits in a few transitional units (less than 5 units per song).

It is interesting to note that recurrence plots can show transitional phrases, which are combinations of units from the previous theme with ones from the next theme the male sings (Frumhoff (1983), Payne et al. (1983), Cholewiak et al. (2013)). Transitional phrases appear in our recurrence plots as blurry areas around self-similar "squares"

that represent the themes. It is particularly clear in figure 6 (right), where the three squares corresponding to themes number 1,2 and 3 inter-penetrate each other. On the contrary, the transition between themes in figure 6 (left) is quite abrupt : the squares are well separated.

Once the songs are extracted, the automatic measurement of the songs parameters is easily done. As an example, we automatically computed songs' length, alphabet (set of different units - or letters - used in the song), and importance of units (% of the number of occurrence of a particular unit in a song occurrence compared to the length of this song). These results are presented in table 1 and figure 7.

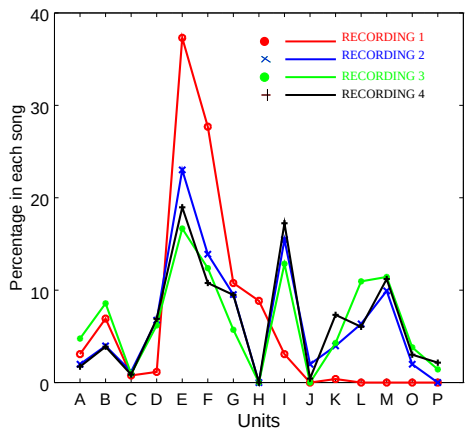

Figure 7. Percentage of each unit in the four song occurrences of recording #1 from Abrolhos archipelago, 2000 (colors are red, blue, green and black for song occurrences 1,2,3,4 respectively). The first occurrence of the song in the recording (see figure 4), in red in this figure is quite different from the other occurrences in term of its units, due to boat noise that prevented an accurate identification of several units.

Figure 7 gives the percentage of each unit in the song occurrence of recording #1. The first occurrence of the song, in red (or circles), is quite different from the other songs. Indeed, during the first song occurrence a series of units are masked by boat noise and were labeled with a special letter (H) in place of a whole theme containing the letters J,K,L,M,O,P. It can be seen in the Levenshtein distance recurrence plot where the first theme of the first rendition of the song does not correlate with any of the themes of other renditions (figure 4). However, the recurrence plot still enabled us to extract the song structure accurately.

The results of table 1, where the mean length of all the songs analyzed is 255, compares well with the results of Arraut and Vielliard (2004), working with the same data but with manual analysis, where the mean song length is around 250-300 units.

Table 1 : Mean and standard deviation of the songs' length for recordings #1,2,3 and 5

Recording	year	nb of songs	mean length (in units)	standard dev. (in units)
#1	2000	4	238	22
#2	2000	6	177	25
#3	2000	8	301	72
#5	2001	1	438	none

5.2. Robustness of the extraction method

In order to test the robustness of our method of songs' extraction, we first checked that the extraction is not dependent on the order n of the Levenshtein distance recurrence plot $LDRP_n$. For each recording (#1 to #3) the number of songs extracted was the same for n going from 4 to 10.

We then checked that the extraction is not heavily dependent on the unit tran-

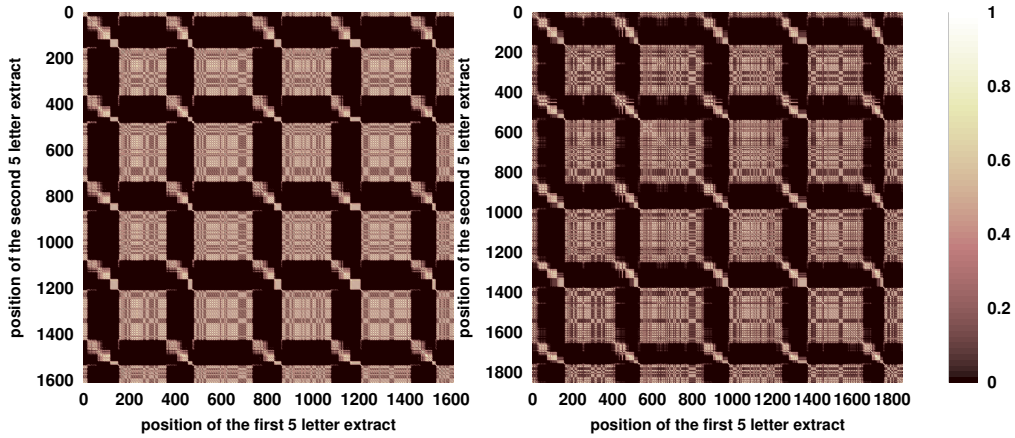

Figure 8. Recurrence plots of part of the recording #3 from Abrolhos archipelago, 2000, with $n = 5$. **Left** : transcription into units by confirmed analysts, **Right** : transcription into units by an inexperienced analyst

scription. For recording #3, an inexperienced analyst transcribed part of it in a string of units. Even though the confirmed analysts transcribed 1614 units and the inexperienced one 1855, the recurrence plots have the same appearance (figure 8). The automatic extraction of the song was then performed on these two strings of units and gave similar results : four songs are extracted in each case. The ratio r_u of the length of the song extracted from the series of units produced by the confirmed analysts and the length of the song extracted from the series of units produced by the inexperienced analyst (measured in unit) is very stable for each song extracted ($r_u = 0.86 \pm 0.006$). We also checked that the determination of the start and end of an automatically extracted song is consistent between the human transcriptions (with a ratio between songs durations (measured in seconds) of $r_t = 1.01 \pm 0.04$ between the two human analysts). Thus, our method of song extraction is remarkably robust to differences in unit labeling due to the subjectivity or lack of training of the analyst.

In addition, we checked the robustness of our method to accidental errors in units transcription, not considering whether this was done automatically or manually. We applied the extraction routine to the strings of units of recording #1 and #2, gradually and randomly changing a percentage of these units (replacing it randomly by a unit from the same recording). For each percentage of error, the test was repeated 20 times and the number of extracted songs was noted. The result is visible in figure 9 where, for each percentage of errors in unit identification, the average number of extracted songs is plotted, along with the error bar corresponding to the variation across the 20 draws. We see that for less than 5% of randomly changed units, the results are very coherent with the original recording. Even as the rate of error grows, the number of extracted songs stays close to the nominal one (1 song of difference for up to 20% of errors).

Last, we applied the extraction routine to strings of units of recording #1 where we gradually and randomly removed a percentage of these units. In this case, until 40% of units were removed, the routine extracts the correct number of songs for more than 90% of the tries. Therefore, our extraction method is robust to missing units during the transcription, for example when the signal to noise ratio of the vocalizations of humpback whale is low. However, this was done for randomly chosen units and do not

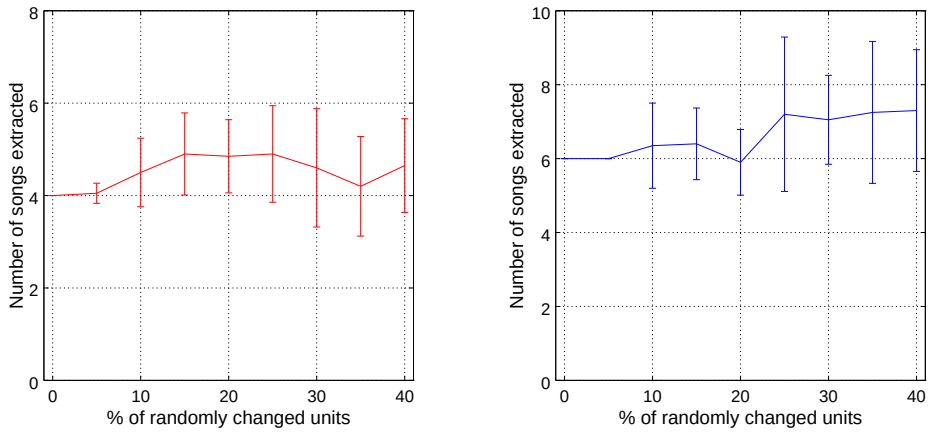

Figure 9. Number of song correctly found in function of the percentage of randomly changed units. The test is done on 20 random tries and the error bar quantifies the distribution of the results over the 20 tries. **Left** : Recording 1 from Abrolhos archipelago, 2000, **Right** : Recording #2 from Abrolhos archipelago, 2000

account for systematic errors (such as would happen if one type of unit only would be lost in the noise, which is not an improbable case).

6. Discussion

6.1. Use of recurrence plots

Recurrence plots are a very visual way of representing the sequence of units, allowing immediate global identification of the structures of the highly organized vocalization emitted by male humpback whales, presented in the precedent sections. It helps with getting a qualitative view of their variability as well as the type of transition (abrupt versus gradual change) between these structures. It could be used as a valuable tool for every highly structured type of sound emission by animals and seems particularly well fitted to study the ever changing structures of humpback whales songs.

Recurrence plots can also be used for automatic unsupervised selection of different structure levels in the recording translated as a series of units. The results of the previous section indicate that our method of extraction is quite robust to errors during manual or automatic transcription of the sound units into a string of letters. The quantification of errors detected for song unit classification systems in studies such as Dunlop et al. (2007), Ou et al. (2013) or Rekdahl et al. (2018) is usually from 10% to 20% which compares well with the percentage of errors that still enables us to automatically extract songs.

The use of recurrence plot to extract songs is still a work in progress but these results are encouraging. Even in the situation of different criteria for unit labeling in different recordings, the song session structure will probably still be visible. This is an additional benefit of the proposed recurrence plot methodology, as it undermines differences in unit labeling among studies. The plot gives a representation of each recording, rather independent on the string it was calculated from. In this way, the recurrence plots provide standardized limits of patterns in the recordings that can be further explored and compared.

This automated song extraction is the first step for many more studies regarding song analysis, and comparison between songs. The automated extraction of other structures (such as themes) could be performed based on this tool but is not as straightforward as songs extraction. Some of the difficulties of theme extraction are : blurred or imprecise transition between themes, unclear definition of a theme, very short or evanescent themes. Thus any method of automated theme extraction would probably be rather ad hoc, which is the reason why we didn't perform it on our limited set of data.

Finally, a matrix comparing two series of units extracted from two different recordings can be computed which shows the similarity between two recordings in term of song rendition. This type of analysis could enable people to compare quickly two song bouts distant in time or space.

6.2. Large scale automatic treatment of humpback whales' recordings

This study is a step towards automatic treatment of large scale recordings. Automatic tools to transcribe sounds into a string of letters such as the one developed in Dunlop et al. (2007), Glotin et al. (2008), Rickwood and Taylor (2008), Pace et al. (2012), Ou et al. (2013), Razik et al. (2015), Bartcus et al. (2015) or Rekdahl et al. (2018) could be used beforehand. Then our method could be applied to extract, classify and compare songs and their features along with other approaches to visualize the structures of humpback whale's songs.

Finally, analyzing humpback whale song structures with recurrence plots would be difficult in the case of a recording where many singers are vocalizing together. In this case, the separation of these singers is a challenge that has to be addressed.

Acknowledgement(s)

Data collection for this work was possible thanks to the support from Fundação O Boticário de Proteção a Natureza / MacArthur Foundation and the Society for Marine Mammalogy (Small-Grants-in-Aid of Research to RSL). We thank the crew of the boat Tomara and Instituto Baleia Jubarte (PBJ) for logistic support. Projeto Baleia Jubarte is sponsored by Petróleo Brasileiro S. A. (Petrobras). We would like to thank Luana Pini for help in data processing and Leonardo Diniz Soares Eugênio who computed the raven tables used to transcribe recordings into series of units . We thank the research program BRILAM STIC AmSud 17-STIC-01, EADM GDRS CNRS MADICS and ANR-18-CE40-0014 SMILES. The collaboration between the authors has been promoted by the Workshop Listening for Aquatic Mammals in Latin America (LAMLA) sponsored by UFRN, CAPES (AUXPE 541/2016) and the U. S. Office of Naval Research (Award number N62909-16-1-2144).

References

J.A. Allen, A. Murray, M.J. Noad, R.A. Dunlop, and E.C. Garland. Using self-organizing maps to classify humpback whale song units and quantify their similarity. *J. Acoust. Soc. Am*, 142(4):1943–1952, 2017.

A. Andriolo, C.C.A. Martins, M.H. Engel, J.L. Pizzorno, S. Mas Rosa, A.C. Freitas, M.E. Morete, and P.G. Kina. The first aerial survey of humpback whales (Megaptera novaean-

gliae) to estimate abundance in the breeding ground off brazil (breeding stock a). *Journal of Cetacean Research and Management*, 8:307–311, 2006.

E. M. Arraut and J.M.E. Vielliard. The song of the brazilian population of humpback whale, Megaptera novaeangliae, in the year 2000 : individual song variations and possible implications. *Anais da Academia Brasileira de Ciencias*, 76(2), 2004.

M. Bartcus, C. Chamroukhi, and H. Glotin. Hierarchical dirichlet process hidden markov model for unsupervised bioacoustic analysis. In *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*, Killarney, Ireland, jul 2015.

D.M. Cholewiak, R.S. Sousa-Lima, and S. Cerchio. Humpback whale song hierarchical structure: Historical context and discussion of current classification issues. *Marine Mammal Science*, 29(3):312–332, 2013.

J.D. Darling. *Migrations, abundance and behavior of Hawaiian humpback whales.* PhD thesis, 1983.

J.D. Darling and M. Brub. Interactions of singing humpback whales with other males. *Marine Mammal Science*, (17):570–584, 2001.

J.D. Darling, J.M.V. Acebes, O. Frey, R.J. Urbn, and M. Yamaguchi. Convergence and divergence of songs suggests ongoing, but annually variable, mixing of humpback whale populations throughout the north pacific. *Scientific reports*, 9, 2019.

R. A. Dunlop, M. J. Noad, D. H. Cato, and D. Stokes. The social vocalization repertoire of east australian migrating humpback whales (Megaptera novaeangliae). *J. Acoust. Soc. Am*, 122(5):2893–2905, 2007.

J.W. Eaton, D. Bateman, and S. Hauberg. *GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations.* CreateSpace Independent Publishing Platform, 2009. ISBN 1441413006.

J.-P. Eckmann, S. Olifsson Kamphorst, and D. Ruelle. Recurrence plots of dynamical systems. *Europhysics letters*, 4(91):973–977, 1987.

J. Foote. Visualizing music and audio using self-similarity. *MULTIMEDIA '99 Proceedings of the seventh ACM international conference on Multimedia (Part 1)*, 1999.

J. Foote. Automatic audio segmentation using a measure of audio novelty. *Conference: Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference onVolume: 1*, 2000.

P. Frumhoff. *Aberrant songs of humpback whales (Megaptera novaeangliae): Clues to the structure of humpback songs.* in Communication and Behavior of Whales, 1983. Westview Press, Boulder.

E.C. Garland, M.J. Noad, A.W. Goldizen, M.S. Lilley, M.L. Rekdahl, C. Garrigue, R. Constantine, N. Daeschler Hauser, M. Poole, and Jooke Robbins. Quantifying humpback whale song sequences to understand the dynamics of song exchange at the ocean basin scale. *J. Acoust. Soc. Am*, 133(1):560–569, 2013.

D.A. Glockner. *Determining the sex of humpback whales (Megaptera novaeangliae) in their natural environment.* in R. Payne, ed. Communication and behavior of whales, 1983. AAAS Selected Symposium 76. Westview Press, Boulder, CO. pages 447-464.

H. Glotin, L. Gauthier, F. Pace, F. Benard, and O. Adam. *New automatic classification for humpback whale songs*, page 93. In P. university, ONR (Eds.), 2008.

S.R. Green, E. Mercado III, A.A. Pack, and L.M. Herman. Recurring patterns in the songs of humpback whales (Megaptera novaeangliae). *Behavioural Processes*, 86:284–294, 2011.

T. Hee-Wai, J.S. Allen, J. Gebbie, and M. Siderius. Nonlinear time series analysis of snapping shrimp sounds. *J. Acoust. Soc. Am.*, 134:41–47, 2013.

V. M. Janik. Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. *Anim. Behav*, 57:133–143, 1999.

A. Kershenbaum, A.E. Bowles, T.M. Freeberg, D.Z. Jin, A.R. Lameira, and K. Bohn. Animal vocal sequences: not the markov chains we thought they were. *Proc. R. Soc. B*, 281, 2014.

V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR*, 163(4):845–848, 1965.

C.C.A Martins, M.E. Morete, M.H. Engel, A.C. Freitas, E.R Secchi, and P.G.Kinas. Aspects

of habitat use patterns of humpback whales in the abrolhos bank, brazil, breeding ground. *Memoirs of the Queensland Museum*, 47:563–570, 2001.

N. Marwan, M. Carmenromano, M. Thiels, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6):237–329, 2007.

E. Mercado-III, L.M. Herman, and A.A. Pack. Stereotypical sound patterns in humpback whale songs : Usage and function. *Aquatic Mammals*, 29(1):37–52, 2003.

R. Miralles, A. Carrin, D. Looney, G.Lara, and D.Mandic. Characterization of the complexity in short oscillating time series: An application to seismic airgun detonations. *The Journal of the Acoustical Society of America*, 138(3):1595–1603, 2015.

H. Ou, W. W. L. Au, L. M. Zurk, and M. O. Lammers. Automated extraction and classification of time-frequency contours in humpback vocalizations. *J. Acoust. Soc. Am.*, 133(1), 2013.

F. Pace, F. Benard, H. Glotin, O. Adam, and P. White. Subunit definition and analysis for humpback whale classification. *Journal of Applied Acoustics*, 7.1, 2010.

F. Pace, P.R. White, and O. Adam. Hidden markov modeling for humpback whale (Megaptera novaeangliae) call classification. In *Societe Francaise d'Acoustique, Proceedings of the Acoustics 2012 Nantes Conference*, pages 2439–2444, 23-27 april 2012.

J. Paulus, M. Muller, and A. Klapuri. Audio-based music structure analysis. *11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010.

K. Payne, Tyack P., and Payne R. *Progressive changes in the songs of humpback whales (Megaptera novaeangliae): A detailed analysis of two seasons in Hawaii*, pages 9–58. in Communication and Behavior of Whales, 1983.

R. S. Payne and S. McVay. Songs of humpback whales. *Science*, 173:585–597, 1971.

H. Pines. Mapping the phonetic structure of humpback whale song units: extraction, classification, and shannon-zipf confirmation of sixty sub-units. *Proc. Mtgs. Acoust.*, 35, 2018.

Bioacoustics Research Program. Raven pro: Interactive sound analysis software (version 1.5). Ithaca, NY: The Cornell Lab of Ornithology, 2014. Available from http://www.birds.cornell.edu/raven.

A. Ravignani and P. Norton. Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, pages 4–19, 2017.

J. Razik, H. Glotin, M. Hoeberechts, Y. Doh, and S. Paris. Sparse coding for efficient bioacoustic data mining: Preliminary application to analysis of whale songs. In *IEEE International Conference on Data Mining (IEEE ICDM) workshop Env. Acoustic Data Mining Forum*, page 7, Atlantic City, NJ, USA, nov 2015.

M.L. Rekdahl, E.C. Garland, G.A Carvajal, C.D. King, T. Collins, and Y. Razafindrakoto. Culturally transmitted song exchange between humpback whales (Megaptera novaeangliae) in the southeast atlantic and southwest indian ocean basins. *R. Soc. open sci*, 5, 2018.

P. Rickwood and A. Taylor. Methods for automatically analyzing humpback song units. *J. Acoust. Soc. Am.*, 123(3), 2008.

W. E. Schevill and W. A. Watkins. *Whale and porpoises voices: A phonograph record.* 1962. Contribution Number 1320, Woods Hole Oceanographic Institution.

J.N. Schneider and E. Mercado-III. Characterizing the rhythm and tempo of sound production by singing whales. *Bioacoustics*, 2018.

O. W. Schreiber. Some sounds from marine life in the hawaiian area. *J. Acoust. Soc. Am.*, 24 (116), 1952.

R.A. Wagner and M.J. Fischer. The string-to-string correction problem. *Journal of the Asociation for Computing Machinery*, 21(I):168–173, January 1974.

J.P. Zbilut, M. Koebbe, H. Loeband, and G. Mayer-Kress. Use of recurrence plots in the analysis of heart beat intervals. *Proceedings Computers in Cardiology*, 1990.