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

5 authors, including:

Franck Malige
Université de Toulon
10 PUBLICATIONS **58** CITATIONS

SEE PROFILE

Julie Patris
Aix-Marseille Université
17 PUBLICATIONS **33** CITATIONS

SEE PROFILE

Divna Djokić
Universidade Federal do Rio Grande do Norte
5 PUBLICATIONS **15** CITATIONS

SEE PROFILE

Renata Santoro Sousa-Lima
Universidade Federal do Rio Grande do Norte
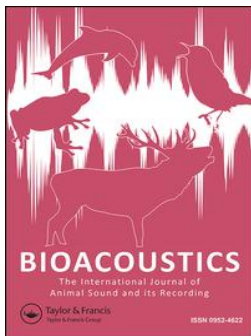**94** PUBLICATIONS **1,131** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project AQUO Project View project

Project sabiod View project

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

F. Malige , D. Djokic , J. Patris , R. Sousa-Lima & H. Glotin

View supplementary material ⧉

Published online: 16 Nov 2020.

Submit your article to this journal ⧉

View related articles ⧉

View Crossmark data ⧉

Taylor & Francis
Taylor & Francis Group

Check for updates

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]

aCNRS, LIS, DYNI Team, Université De Toulon, Aix-Marseille Université, Marseilles, France; bLaboratory of Bioacoustics, Department of Physiology and Behavior, Universidade Federal do Rio Grande do Norte (UFRN), Natal, Brazil; cGraduate Program of Psychobiology, Biosciences Center, UFRN, Natal, Brazil

ABSTRACT

Humpback whale song is comprised of well-structured distinct levels of organisation: 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 behaviour previously. Here, we aim to apply this tool to visualise and recognise structures traditionally used in inferences about behaviour (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.

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 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 characterise 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 behavioural 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 characterise 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 vocalisation. Unit is the best-defined element of the humpback whale song, and can be described as 'the shortest sound entity recognised 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, the 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, analyse or extract the aforementioned features from transcribed unit label strings.

1.2. Recurrence plots

Recurrence plots are used to visualise and analyse, 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 analysing humpback whale songs has been to visualise 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 (2018), without focusing on the spectral content of the sounds.

The main topic of this paper is to apply this tool to visualise and recognise 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). These 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 were collected in the Abrolhos bank, located off the Northeastern coast of Brazil (17ºS and 38ºW) 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). Vocalising 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 to 500 m to the focal male while its behavioural 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 h 37 min and 2 h and 7 minutes. In September 2001, the two selected audio files (recordings #4,5) were, respectively, 26 min and 1 h and 38 min 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).

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 12 to more than 100 (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, (2014)), and subjected to aural and visual inspection by two trained analysts. Each analyst separately inspected and labelled 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 organisation 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 labelling was done in the following way: every unit type was attributed to 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 labelled 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 vocalisation (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
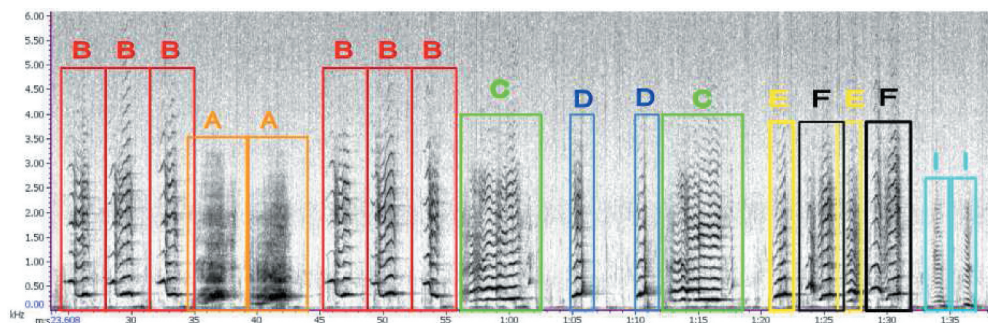

Figure 1. Time (min:sec)/frequency (kHz) representation of an extract of recording #3 (from Abrolhos Archipelago 2000), 1024 point FFTs, Hann window, 43.1 Hz resolution and 50% overlap. Unit transcription into letters is signalled above the boxes.

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 visualise 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.

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 1 h in a domestic computer. The parameter n can be adapted depending on the use of the $LDRP_n$: visual analysis, extraction of structures, etc.
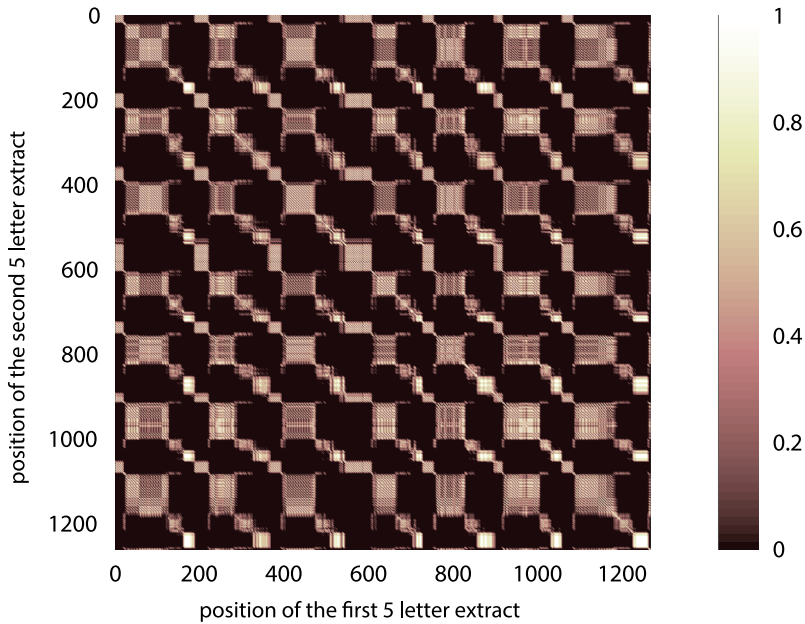
Figure 2. Example of a Levenshtein distance recurrence plot (5 letter extracts) created form recording number #2 from Abrolhos Archipelago (2000). The number of units of this recording is $N = 1260$. The coefficients of the matrix denote correlation between 5 letter extracts (0 is no correlation, 1 is maximal correlation).
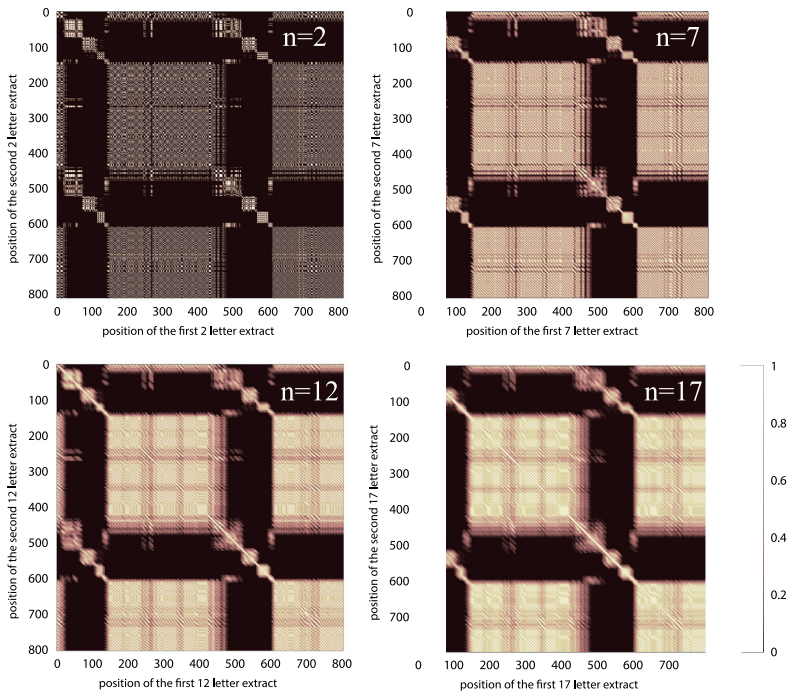

Figure 3. Recording #5 (from Abrolhos Archipelago 2001): Levenshtein distances recurrence plots for different values of the length n of the extract (2,7,12 and 17).
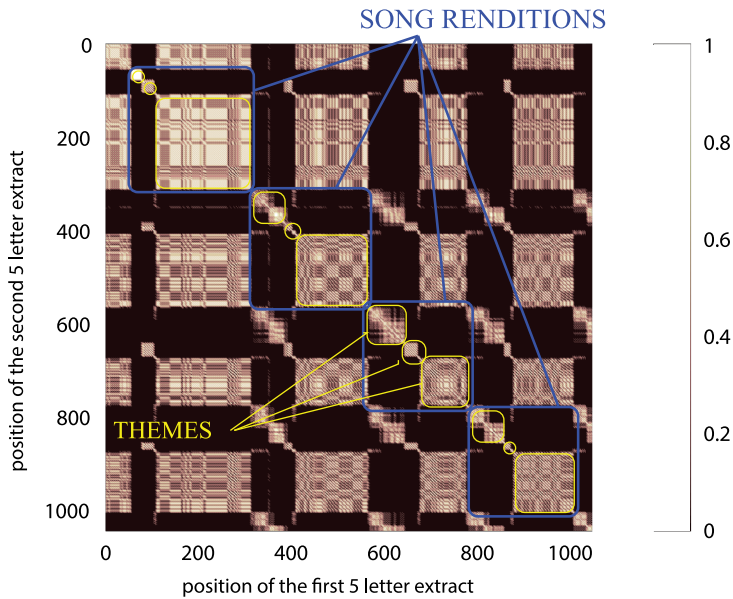
Figure 4. Recording #1 from Abrolhos archipelago, 2000: Levenshtein distances recurrence plot commented (*n*=5). The main structures of humpback whale sound production appear: songs in blue, themes in yellow.

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 these 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 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; Cholewiak et al. 2013).

A second scale of the 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 **theme 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 a 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 two-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 in Figure 5, centre.
(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
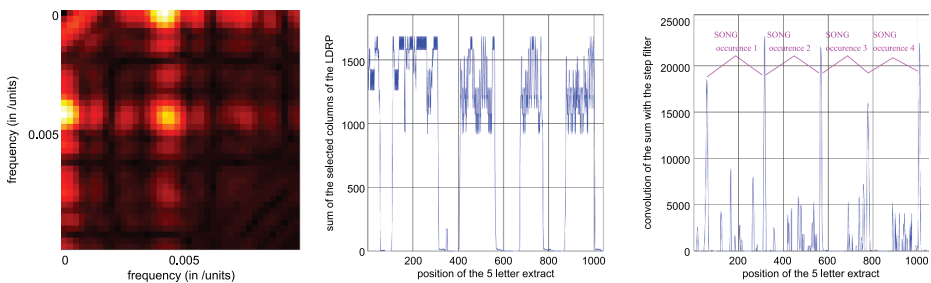

Figure 5. Steps in the extraction of the songs. **Left**: Step 2 of the algorithm, extract of the 2D fast Fourier transform of the Levenshtein distance matrix (for $n = 5$, recording #1). A peak is visible at a frequency of about 0.045 units^{-1} (which means a scale of 200 to 250 units of the recording) **centre**: 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.
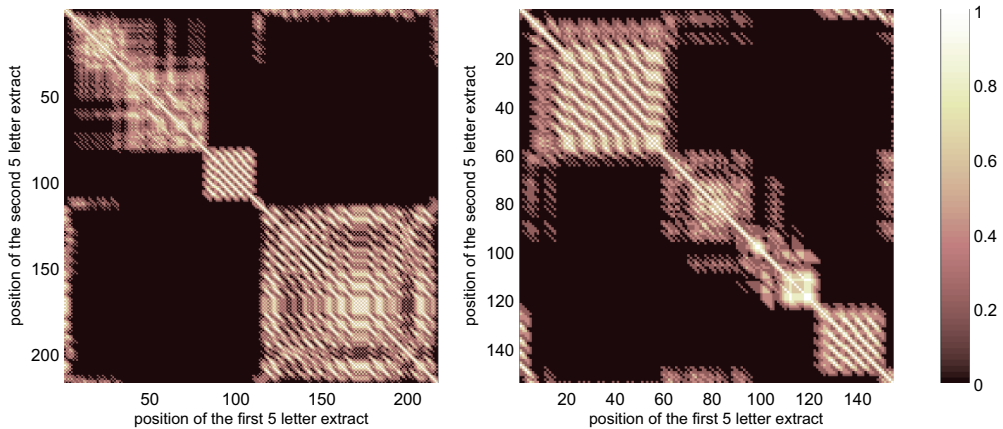
Figure 6. Two songs automatically extracted. **Left**: Recording #1 from Abrolhos archipelago, 2000: third song automatically extracted from Figure 4. In order to see clearly the song, the frame of the figure is chosen a bit larger for representation. **Right**: recording #2 from Abrolhos archipelago, 2000: zoom on the third song occurrence in the Levenshtein distance recurrence plot of Figure 2. Different types of transition between themes appear.

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 numbers 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 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 labelled 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

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

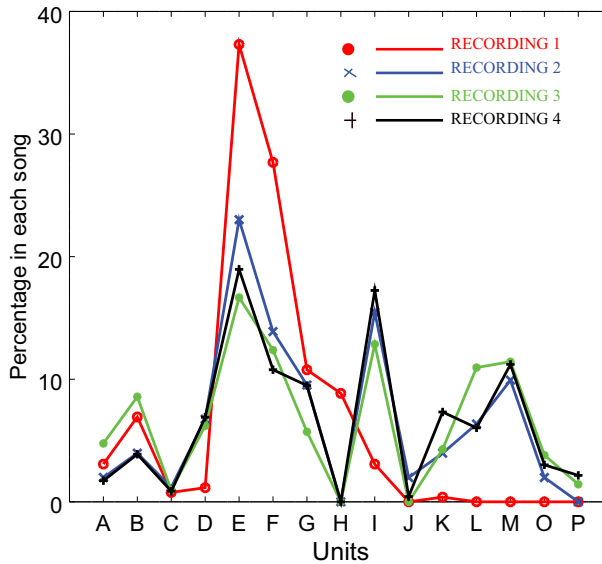| Recording	Year	nb of songs	Mean length (in units)	Standard dev. (in units)
#1	2000	4	238	22
#2	2000	6	177	25
#3	2000	8	301	72
#5	2001	1	438	None

Figure 7. Percentage of each unit in the four song occurrences of recording #1 from Abrolhos archipelago, 2000 (colours 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.

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 analysed is 255, compare 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.

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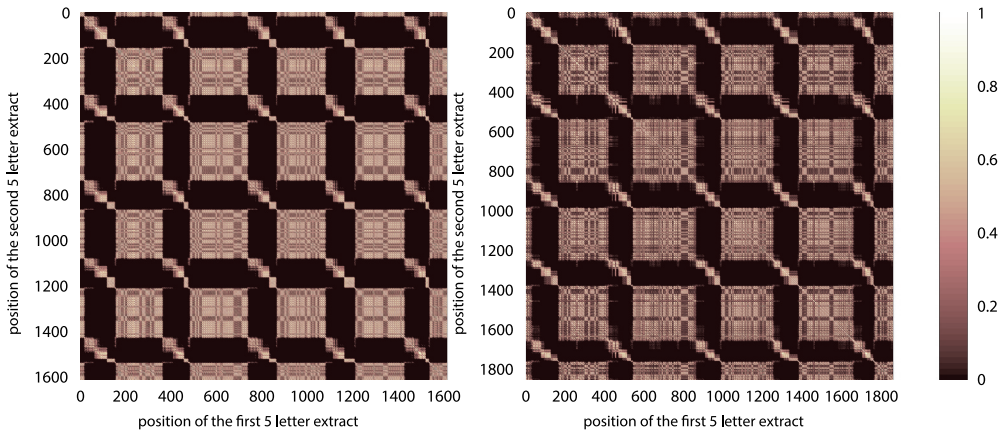
Figure 8. Recurrence plots of part of the recording #3 from Abrolhos archipelago, 2000, with $n = 5$. **Left**: transcription into units by confirmed analysts, **Right**: transcription into units by an inexperienced analyst.

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 labelling 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 vocalisations of humpback whale is low. However, this was done for randomly chosen units and do not account
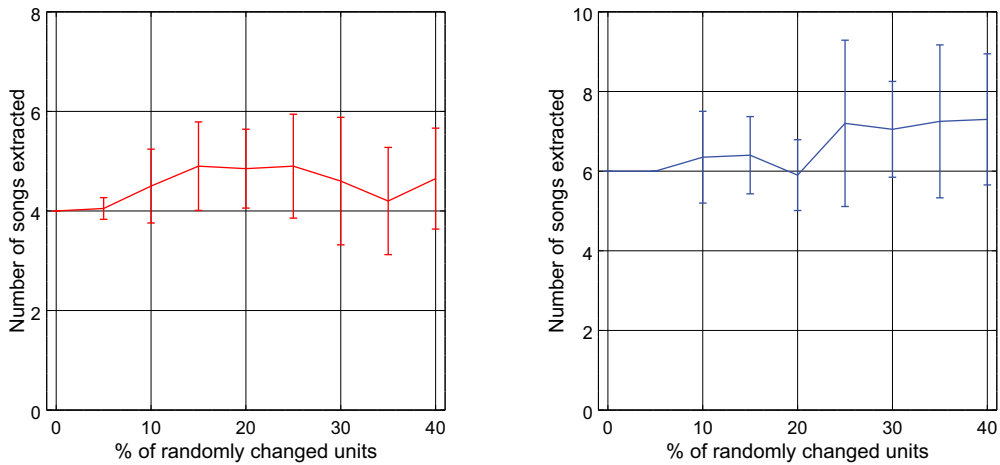
Figure 9. Number of song correctly found in function of the percentage of randomly changed units. The test is done on 20 random tries and the error bar quantifies the distribution of the results over the 20 tries. **Left**: Recording 1 from Abrolhos archipelago, 2000, **right**: recording #2 from Abrolhos archipelago, 2000.

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 organised vocalisation emitted by male humpback whales, presented in the preceding 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 whale 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 labelling 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 labelling among studies. The plot gives a representation of each recording, rather independent of the string it

was calculated from. In this way, the recurrence plots provide standardised 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 did not 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 the 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. (2009), 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 visualise the structures of humpback whale's songs.

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

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.

Disclosure statement

Funding

ORCID

F. Malige http://orcid.org/0000-0003-4567-8483
J. Patris http://orcid.org/0000-0001-6281-7380

References

Allen JA, Murray A, Noad MJ, Dunlop RA, Garland EC. 2017. Using self-organizing maps to classify humpback whale song units and quantify their similarity. J Acoust Soc Am. 142 (4):1943–1952. doi:10.1121/1.4982040.

Andriolo A, Martins CCA, Engel MH, Pizzorno JL, Mas Rosa S, Freitas AC, Morete ME, Kina PG. 2006. The first aerial survey of humpback whales (Megaptera novaeangliae) to estimate abundance in the breeding ground off brazil (breeding stock a). J Cetacean Res Manage. 8:307–311.

Arraut EM, Vielliard JME. 2004. The song of the brazilian population of humpback whale, Megaptera novaeangliae, in the year 2000: individual song variations and possible implications. An Acad Bras Cienc. 76:2. doi:10.1590/S0001-37652004000200028.

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

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

Cholewiak DM, Sousa-Lima RS, Cerchio S. 2013. Humpback whale song hierarchical structure: historical context and discussion of current classification issues. Mar Mammal Sci. 29 (3):312–332. doi:10.1111/mms.12005.

Darling JD. 1983. Migrations, abundance and behavior of Hawaiian humpback whales [PhD thesis].

Darling JD, Acebes JMV, Frey O, Urbn RJ, Yamaguchi M. 2019. Convergence and divergence of songs suggests ongoing, but annually variable, mixing of humpback whale populations throughout the north pacific. Sci Rep. 9. https://www.nature.com/articles/s41598-019-42233-7#citeas

Darling JD, Brub M. 2001. Interactions of singing humpback whales with other males. Mar Mammal Sci. 17:570–584. doi:10.1111/j.1748-7692.2001.tb01005.x.

Dunlop RA, Noad MJ, Cato DH, Stokes D. 2007. The social vocalization repertoire of east australian migrating humpback whales (Megaptera novaeangliae). J Acoust Soc Am. 122 (5):2893–2905. doi:10.1121/1.2783115.

Eaton JW, Bateman D, Hauberg S. 2009. GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations. CreateSpace Independ Publ Platform; Orlando, FL. ISBN 1441413006.

Eckmann J-P, Olifsson Kamphorst S, Ruelle D. 1987. Recurrence plots of dynamical systems. Europhys Lett. 4(91):973–977. doi:10.1209/0295-5075/4/9/004.

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

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

Frumhoff P. 1983. Aberrant songs of humpback whales (Megaptera novaeangliae): clues to the structure of humpback songs. In: Payne RS, editor. Communication and behavior of whales. Westview Press: Boulder; p. 81-127.

Garland EC, Noad MJ, Goldizen AW, Lilley MS, Rekdahl ML, Garrigue C, Constantine R, Daeschler Hauser N, Poole M, Robbins J. 2013. 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. doi:10.1121/1.4770232.

Glockner DA. 1983. Determining the sex of humpback whales (Megaptera novaeangliae) in their natural environment. In: Payne R, editor. Communication and behavior of whales. AAAS selected symposium 76. (Boulder, CO): Westview Press; p. 447–464.

Glotin H, Gauthier L, Pace F, Benard F, Adam O. 2009. 4th International Workshop on Detection, Classification and Localization of Marine Mammals Using Passive Acoustics. Pavia: Office of Naval Research; p. 93.

Green SR, Mercado EIII, Pack AA, and Herman LM. 2011. Recurring patterns in the songs of humpback whales (Megaptera novaeangliae). Behav Processes. 86:284–294. doi:10.1016/j. beproc.2010.12.014.

Hee-Wai T, Allen JS, Gebbie J, Siderius M. 2013. Nonlinear time series analysis of snapping shrimp sounds. J Acoust Soc Am. 134:41–47. doi:10.1121/1.4831199.

Janik VM. 1999. Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. Anim Behav. 57:133–143. doi:10.1006/anbe.1998.0923.

Kershenbaum A, Bowles AE, Freeberg TM, Jin DZ, Lameira AR, Bohn K. 2014. Animal vocal sequences: not the markov chains we thought they were. Proc R Soc B. 281:20141370. doi:10.1098/rspb.2014.1370.

Levenshtein VI. 1965. Binary codes capable of correcting deletions, insertions, and reversals. Dokl Akad Nauk SSSR. 163(4):845–848.

Martins CCA, Morete ME, Engel MH, Freitas AC, Secchi ER, Kinas PG. 2001. Aspects of habitat use patterns of humpback whales in the abrolhos bank, brazil, breeding ground. Mem Queensland Mus. 47:563–570.

Marwan N, Carmenromano M, Thiels M, Kurths J. 2007. Recurrence plots for the analysis of complex systems. Phys Rep. 438(5–6):237–329. doi:10.1016/j.physrep.2006.11.001.

Mercado-III E, Herman LM, Pack AA. 2003. Stereotypical sound patterns in humpback whale songs: usage and function. Aquat Mammal. 29(1):37–52. doi:10.1578/016754203101024068.

Miralles R, Carrin A, Looney D, Lara G, Mandic D. 2015. Characterization of the complexity in short oscillating time series: an application to seismic airgun detonations. J Acoust Soc Am. 138 (3):1595–1603. doi:10.1121/1.4929694.

Ou H, Au WWL, Zurk LM, Lammers MO. 2013. Automated extraction and classification of time-frequency contours in humpback vocalizations. J Acoust Soc Am. 133:1.

Pace F, Benard F, Glotin H, Adam O, White P. 2010. Subunit definition and analysis for humpback whale classification. J Appl Acoust. 7:1.

Pace F, White PR, Adam O. 2012. Hidden markov modeling for humpback whale (Megaptera novaeangliae) call classification. In Societe Francaise d'Acoustique, Proceedings of the Acoustics 2012 Nantes Conference; April 23–27; Nantes, France; p. 2439–2444

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

Payne K, Tyack P, Payne R. 1983. Progressive changes in the songs of humpback whales (Megaptera novaeangliae): A detailed analysis of two seasons in Hawaii. In: Payne R, editor. Communication and behavior of whales. Colorado: Westview Press; p. 9–58.

Payne RS, McVay S. 1971. Songs of humpback whales. Science. 173:585–597. doi:10.1126/ science.173.3997.585.

Pines H. 2018. Mapping the phonetic structure of humpback whale song units: extraction, classification, and shannon-zipf confirmation of sixty sub-units. Proc Mtgs Acoust. 35:010003. doi: 10.1121/2.0000957.

Ravignani A, Norton P. 2017. Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. J Lang Evol. 2:4–19. doi:10.1093/jole/lzx002.

Razik J, Glotin H, Hoeberechts M, Doh Y, Paris S. 2015. 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; Nov; Atlantic City, NJ, USA. p. 7.

Rekdahl ML, Garland EC, Carvajal GA, King CD, Collins T, Razafindrakoto Y. 2018. Culturally transmitted song exchange between humpback whales (Megaptera novaeangliae) in the southeast Atlantic and southwest Indian Ocean basins. R Soc Open Sci. 5:172305. doi:10.1098/rsos.172305.

Rickwood P, Taylor A. 2008. Methods for automatically analyzing humpback song units. J Acoust Soc Am. 123:3. doi:10.1121/1.2836748.

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

Schneider JN, Mercado E III. 2018. Characterizing the rhythm and tempo of sound production by singing whales. Bioacoustics. doi:10.1080/09524622.2018.1428827

Schreiber OW. 1952. Some sounds from marine life in the hawaiian area. J Acoust Soc Am. 24:116. doi:10.1121/1.1917427.

Wagner RA, Fischer MJ. 1974. The string-to-string correction problem. J Asoc Comput Mach. 21 (I):168–173. January . doi:10.1145/321796.321811.

Zbilut JP, Koebbe M, Loeband H, Mayer-Kress G. 1990. Use of recurrence plots in the analysis of heart beat intervals. Proceedings of the IEEE Conference onComputers in Cardiology; Chicago: IEEE Computer Society Press; p. 263–266.