HAL
open science

# A case study of biological pulsed sound : the blue whale's southeast Pacific song unit, models and properties

Julie Patris, Franck Malige, Hervé Glotin, Mark Asch, Susannah J Buchan

## To cite this version:

Julie Patris, Franck Malige, Hervé Glotin, Mark Asch, Susannah J Buchan. A case study of biological pulsed sound : the blue whale's southeast Pacific song unit, models and properties. 2019. hal-02074437

**HAL Id: hal-02074437**

**https://hal.science/hal-02074437v1**

Preprint submitted on 20 Mar 2019

**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, University of Toulon, DYNI team, LIS Laboratory, CNRS**
preprint submitted to JASA (Journal of the Acoustical Society of America) in january 2019

# A case study of biological pulsed sound : the blue whale's southeast Pacific song unit, models and properties

Julie Patris (1,6), Franck Malige (1,6), Hervé Glotin (1,6), Mark Asch (2) and Susannah J. Buchan (3,4,5,6)

*1 : AMU, Université de Toulon, CNRS, LIS, Marseille, DYNI team, France*
*2 : Université de Picardie Jules Verne, CNRS, LAMFA, Amiens, France*
*3 : COPAS Sur-Austral, University of Concepción, Chile*
*4 : Centro de Estudios Avanzados en Zonas Áridas, Chile*
*5 : Woods Hole Oceanographic Institution, USA*
*6 : BRILAAM, STICAmSud*

***Abstract***

*Pulsed sounds are interesting as an example of complex biological sound. We propose a classification into two groups for the pulsed sounds: tonal or non-tonal. Two mathematical models allow to see the properties of the sound in both cases. This classification is useful for developing new measurements that can be more accurate and can distinguish between two possibilities of sound production. We apply our method to blue whale vocalization and find that the pulse rate corresponds to the fundamental frequency (not expressed in the spectrum) of the song. Thus, we reinforce the hypothesis that the sound is produced by only one organ and then filtered by the body of the giant.*

*March 20, 2019*

# 1 Introduction

## 1.1 Pulsed sounds in biological context

*Among the numerous possibilities of animal sounds, pulsed sounds are particularly complex and interesting and are common in marine mammals [4] [27]. Usually a pulsed sound is the repetition of similar 'pulses' or short signals, with a constant pulse rate, often described in odontocete's vocalizations [20]. Aurally, these sounds are often perceived by listeners as amplitude modulated sounds. The aspect of pulsed sound in the frequency domain is characterized by a series of equally-spaced frequency peaks [30].*

*Two biological pulsed sounds from marine mammals are represented in the time domain in figure 1.*

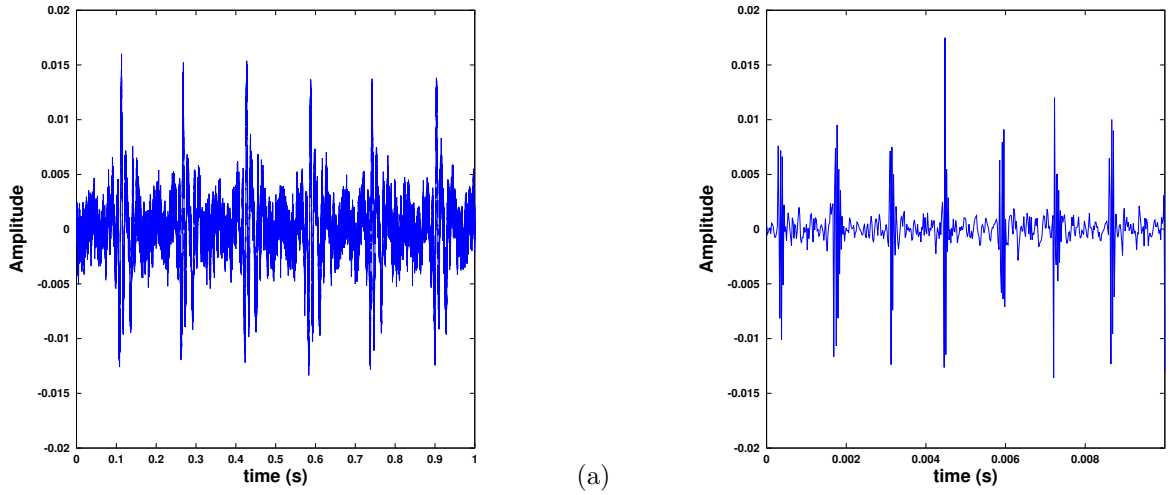


Figure 1: Waveform of two biological pulsed sounds, both recorded off Chañaral de Aceituno Island, Chile, in 2017 with an autonomous recorder at $f_s = 48\,000$ Hz. (a) Top : unit $B$ of the southeast Pacific song type 2 of a blue whale (*balaenoptera musculus*). (b) Bottom : buzzed sound of a bottlenose dolphin (*tursiops truncatus*) (color online)

*Though pulsed sounds are frequent in animal vocalization [30], no general index has been proposed to measure and compare its properties. Thus, in literature references we may read about 'modulation rate' [8], 'pulse rate' [19], 'inter pulse interval' (when the pulses are clearly separated as in fin whales, Balaenoptera Physalus) [31], 'inter note interval' [21], 'time separation pitch' [4] or 'harmonic intervals' [30], [8], all referring to the same parameter.*

*A pulsed sound is also often characterized by its peak frequency, along with classical tonal sounds [6]. However, this definition implies a comparison of energy in the frequency bands that is dependent on propagation, thus it is usually not a stable measurement (in pulsed sounds as in tonal sounds [3] for the typical 'B' call of north Pacific blue whales). Peak frequency will also depend on the representation of the pulsed sound and size of the Fourier transform window used for measurements.*

*The fundamental frequency is also used [8] but is not always visible, nor relevant for a pulsed sound.*

*A precise analysis of a pulsed sound is necessary to discriminate between all these tools and to compare new measurements with historical data. For instance, it allows to compute the long-term frequency decline that has been found in most of blue whales' songs [17] or when a joint decrease of pulse rate and peak frequency is shown as in [19] for blue whales and [31] for fin whales.*

## 1.2 Sound production

*Sound production in large marine mammals is a difficult subject since live animals cannot be examined. However, physical analysis and anatomical research have been proposed, mainly in Aroyan's large chapter [3]. Sound production in mysticetes is still poorly known compared to odontocetes [4] and is an active area of research, [26]. Some studies try to reproduce sound production using theoretical models or sound production machines [1]. According to these studies and anatomical analyses,*

*sound is produced by vibrating U-shaped vocal folds and a complex system of resonators (lungs, laryngeal sac, trachea and other tissues) that modify the sound as a passive filter. Interestingly, another organ has been proposed for the production of pulses in the mysticetes' sounds [26]. In this case the sound produced is called two-voiced sound (or biphonation). Different frequencies are produced by the different active organs e.g. the vocalizations of male North American Wapiti, Cervus canadensis [25]. The non-pulsed part of north Pacific blue whales has recently been studied in search of a cue for sound production [9], showing the growing interest in finding clues in signal processing and signal modelling for sound production.*

### 1.3 Classification of pulsed sound for analysis

*The goal of this paper is to propose a simple classification of pulsed sound, along with two mathematical models. We show the interest of such methods to have a better description of the sound, to achieve more accurate measurements of marine mammals vocalizations, but also as a cue to sound production mechanisms. While most of the efforts of describing pulsed sounds have been done on odontocetes [5], we propose analyzing pulsed sounds emitted by blue whales (Balaenoptera musculus) in Chilean coastal seas [6] as a example of the application of our method.*

## 2 Pulsed sounds

### 2.1 Description of the spectrum of a pulsed sound

*The Fourier transform of a pulsed sound shows peaks (or lines) of frequencies, with an approximately constant separation between frequencies (see figure 2).*

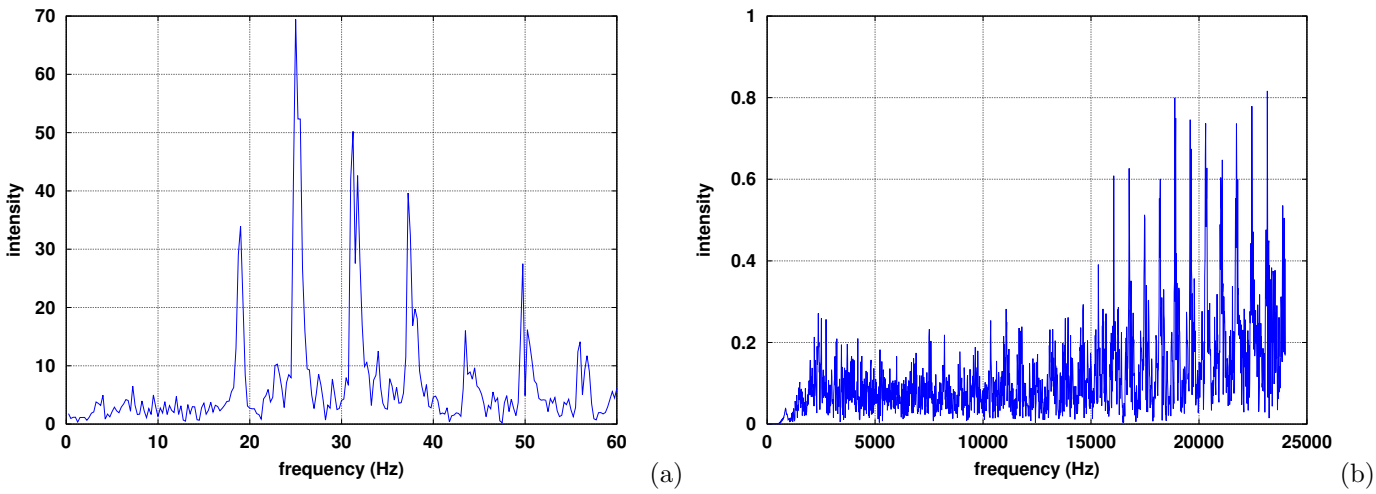


Figure 2: Spectra (by mean of a Fast Fourier Transform or FFT) of two biological pulsed sounds, both recorded off Chañaral de Aceituno Island, Chile in 2017 with an autonomous recorder at $f_s = 48\ 000$ Hz. (a) Top : unit $B$ of the southeast Pacific songtype 2 of a blue whale (*balaenoptera musculus*), signal length : 4 seconds. (b) Bottom : buzzed sound of a bottlenose dolphin (*tursiops truncatus*), signal length : 0.07 s. (color online)

*The intensity of each frequency is not a good marker of the sound because it can be influenced strongly by propagation [15]. On the contrary, the abscissa of the peaks is only weakly affected by propagation and thus is a very good marker of the sound. The various frequencies corresponding to the peaks will be called the $\{f_i\}$, and the constant band interval is called $\Delta f$. As can be verified in our models below, and is also shown in Watkins' original paper [30], this band interval corresponds, in time domain, to the repetition rate of pulses, or pulsed rate called $f_{\text{pulse}}$ in our study. Table 1 gives the peak frequencies $\{f_i\}$ and the averaged band interval or pulse rate $\Delta f = f_{\text{pulse}}$ for the two examples of pulsed sound presented in figure 2, simply measured in the spectrum. These measures have rather high uncertainty and one aim of this study is to list better methods to compute the frequencies $f_i$ and $\Delta f = f_{\text{pulse}}$.*

Table 1: Peak frequencies and average pulsed rate of the two examples shown in figure 2 as measured from the spectra.

| Spectrum | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $\Delta f$ |
|---|---|---|---|---|---|
| Blue whale | 19.0 | 25.2 | 31.3 | 37.6 | 6.26 |
| Bottlenose dolphin | 18900 | 19610 | 20330 | 21040 | 712 |

*We propose the following criterion for a classification of pulsed sound in two groups: if* $\exists(k_i) \in \mathbb{N} \ / \ \forall i, \ \ f_i = k_i.\Delta f$ *that is if the peak frequencies* $\{f_i\}$ *are all integer multiples of the pulse rate* $\Delta f$*, then the sound will be called tonal (although the fundamental frequency is not visible or expressed in the spectrum). In this case the signal is periodic of period* $T_{pulse} = 1/f_{\text{pulse}}$*. Else, the pulsed sound will be called non-tonal.*

*If we examine again our two examples of the blue whale song and the dolphin buzz, we see that the first one can be classified as a tonal signal, whereas the second one cannot (see table 2).*

Table 2: Ratio between frequencies $f_i$ of table 1 and pulsed rate of the two examples shown in figure 2

.

| Frequencies ratio | $\frac{f_1}{\Delta f}$ | $\frac{f_2}{\Delta f}$ | $\frac{f_3}{\Delta f}$ | $\frac{f_4}{\Delta f}$ |
|---|---|---|---|---|
| Blue whale | 3.04 | 4.03 | 5.00 | 6.01 |
| Bottlenose dolphin | 26.55 | 27.54 | 28.55 | 29.55 |

*Before examining the consequences of this classification in section 2.3, we propose two mathematical descriptions of pulsed sounds. These mathematical descriptions will help in understanding the benefit that comes from classifying biological pulsed sounds as well as in choosing the best way to measure* $\{f_i\}$ *and* $\Delta f = f_{\text{pulse}}$*.*

## 2.2 Mathematical models for the interpretation of different pulsed sounds

*In this section we present two mathematical models of pulsed sounds. The first one, model A, can only apply to a signal where the peak frequencies* $\{f_i\}$ *are integer multiples of the band interval or pulse rate* $\Delta f$ *(tonal signals). The second one, model B, is more general and can be applied to any of the two possibilities described in the preceeding paragraph.*

*In both models, we consider the pulsed sound as infinite in time, which means we are not addressing the effects of the global duration of the sound. Eventually, if needed, a window* $w$ *corresponding to the global duration of the signal will be used in the computation of theoretical formulas.*

*For each model we present and compute the Fourier transform and autocorrelation function of the signal. These two operators are often used in signal processing to analyze the signals and to measure parameters such as peak frequency and pulse rate.*

**Model A** *Let us first consider the pulsed sound as the repetition of distinct and similar pulses (actually like the blood pulse), separated by a duration* $T_{\text{pulse}}$*. This is the point of view developed in [9] to model the northeast Pacific blue whale song type. Then the easiest way to mathematically represent such a function of time is the convolution of a specific wavelet* $p$ *(the pulse) by a Dirac comb* $\text{Ш}_{T_{pulse}}$*, characterized by the time* $T_{\text{pulse}}$ *between each impulse. We note* $f_{\text{pulse}} = 1/T_{\text{pulse}}$*.*

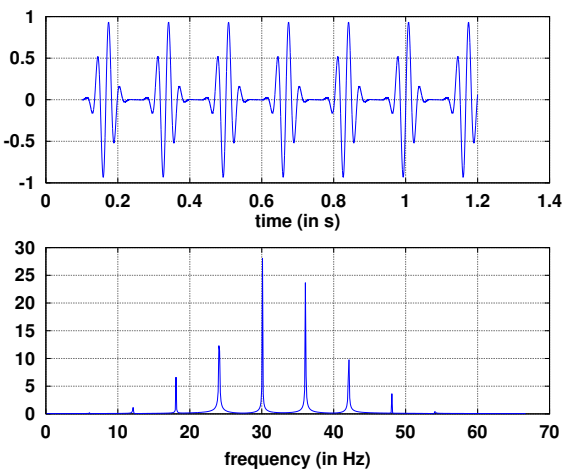


Figure 3: Model A in waveform (top) and its FFT (bottom). In model A we choose the pulse as the multiplication of a sine function of frequency $f_0 = 31.7$ Hz by a Gaussian with standard deviation ($\sigma = 0.02$s). We also choose $f_{pulse} = 6$ Hz. In the FFT we can see that the peaks are multiples of $f_{\text{pulse}} = 6$ Hz (color online)

*The signal of the figure 3 can then be written as :*

$$s_A(t) = \sum_{n \in \mathbb{Z}} p(t - nT_{\text{pulse}}) = [p * \text{Ш}_{T_{\text{pulse}}}](t),$$

*where $*$ is the convolution symbol, $[g*h](t) = \int_{-\infty}^{+\infty} g(u)h(t-u)du$, $\text{Ш}_{T_{pulse}}(t) = \sum_{n\in\mathbb{Z}} \delta_{nT_{\text{pulse}}}(t)$ is the Dirac comb distribution of period $T_{\text{pulse}}$, and $\delta_{t_0}(t)$ is the Dirac distribution centered at $t_0$.*

*A good reference on these techniques is [2]. We note that $s_A$ is then a periodic function, of period $T_{\text{pulse}}$. There is no phase difference between the pulses (see figure 3). A more complete model would include an additive noise term, $\nu(t)$, usually assumed to be of zero mean and known (estimated) variance.*

*The Fourier transform of the function $s_A(t)$ is defined by $S_A(f) = \int_{-\infty}^{+\infty} s_A(t)e^{-2i\pi ft}dt$ and gives*

$$S_A(f) = P(f) \times \text{Ш}_{f_{\text{pulse}}}(f),$$

*where $P(f)$ is the Fourier transform of the wavelet. In figure 3, $p(t)$ is a Gaussian multiplied by a sine of frequency $f_0 = 31.7$ Hz and its Fourier transform $P(f)$ is therefore a Gaussian centered on $f_0$. We observe that the Fourier transform of the signal is the spectrum of the pulse $P$ multiplied by a Dirac comb. Thus the spectrum of $s_A$ is a set of frequency bands at integer multiples of $f_{\text{pulse}}$ (figure 3, bottom). The frequency band with higher energy (30 Hz) does not correspond to $f_0 = 31.7$ Hz (even if the wavelet maximum of energy was at this frequency) because of the multiplication by the Dirac comb.*

*It is important to underline that in practice, the signals analyzed are finite, of duration $T_{\text{signal}}$. In this case, we can write $s_{A,\text{finite}}(t) = s_A(t) \times w(t)$ where $w$ is a window of duration $T_{\text{signal}}$. A classic window is the rectangular window, build on an indicator function : $w(t) =_{[-T_{\text{signal}}/2;T_{\text{signal}}/2]}(t)$ but any kind of window can be used. In this case, the Fourier transform is*

$$S_{A,\text{finite}}(f) = [(P \times \text{Ш}_{f_{\text{pulse}}}) * W](f),$$

*where $W$ is the Fourier transform of $w$. In the case of a rectangular window, the Fourier transform is $W(f) = T_{\text{signal}} \times \text{sinc}(\pi T_{\text{signal}} f)$, a cardinal sine giving a width to the peaks in figure 3, linked to the value of $T_{\text{signal}}$.*

*The autocorrelation function of a signal $s$ is $C_s(\tau) = lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} s(t)s^*(t+\tau)dt$ where $s^*$ is the complex conjugate of $s$. In the case of the finite signal $s_{A,\text{finite}}$ and a rectangular window $w(t) =_{[-T_{\text{signal}}/2;T_{\text{signal}}/2]}(t)$, the autocorrelation function is*

$$C_{s_{A,\text{finite}}}(\tau) \simeq \Lambda(\frac{\tau}{T_{\text{signal}}}) \times (\sum_{n\in\mathbb{Z}} |P(nf_{pulse})|^2 e^{2i\pi nf_{pulse}\tau}),$$

*where $\Lambda(t)$ is the triangular function ($\Lambda(t) = 1+t$ on [-1;0] and $\Lambda(t) = 1-t$ on $[0;1]$ and zero outside of $[-1;1]$). The proof is very similar to the proof given in the appendix for the model B and will not be detailed. An important remark is that the first maximum of the modulus of the autocorrelation function (other than $\tau = 0$) is obtained for $\tau = T_{\text{pulse}}$, period of the signal. Thus, for this model of pulsed sound, the autocorrelation or the summed autocorrelation [33] is a good, unbiased tool to measure the pulse rate.*

**Model B** *We will now examine the case when the pulsed sound can be described as a tonal sound modulated in amplitude by a periodic function. This kind of pulsed sound has been described by [30] or [5]. A straightforward way to represent this signal is to multiply a tonal function $g_{T_0}$ (characterized by a fundamental frequency $f_0$) by a signal that could be an envelope $e$ convoluted by a Dirac comb (of period $T_{\text{pulse}}$). We assume that $f_{\text{pulse}} \ll f_0$ so that the tonal function $g_{T_0}$ is modulated in amplitude by a function with a much smaller frequency.*

*We write*

$$s_B(t) = g_{T_0}(t) \times \sum_{n\in\mathbb{Z}} e(t - nT_{pulse}) = g_{T_0}(t) \times \left[e * \text{Ш}_{T_{pulse}}\right](t).$$

*In this case, the signal is not a periodic function. If we examine each of the 'pulses', they do not have the same phase (see figure 4). This is due to the multiplication of two tonal functions with different periods.*

*Let us compute the Fourier transform of such a signal. We obtain*

$$S_B(f) = \sum_{n\in\mathbb{Z}} E(nf_{pulse})G_{f_0}(f - nf_{pulse}),$$

*where $E$ and $G_{f_0}$ are the Fourier transforms of $e$ and $g_{T_0}$. The proof is given in the appendix. In this formula, as $g_{T_0}$ is a tonal sound, only multiples of $f_0$ are found in its spectrum and thus $S_B(f)$ is different from zero only if $f - nf_{\text{pulse}}$ is a multiple of $f_0$, that is only if $f = nf_{pulse} + mf_0$, which is usually not a multiple of $f_{\text{pulse}}$. In addition, if the tonal signal $g_{T_0}(t)$ had its energy concentrated at the frequency $f_0$, and if the Fourier transform of the envelope $E(f)$ is sufficiently regular, we see that the pulsed signal $s_B$ has a maximum of energy also at $f_0$ (see figure 4).*

*As $g_{T_0}$ is a tonal signal with fundamental $f_0$, we can write it as $g_{T_0}(t) = \sum_{n\in\mathbb{Z}} a_n e^{2i\pi nf_0 t}$. Then the autocorrelation function of the finite signal associated to model B is*

$$C_{s_{B,\text{finite}}}(\tau) \simeq \Lambda(\tau/T_{\text{signal}})(\sum_{n\in\mathbb{Z}} |a_n|^2 e^{2i\pi nf_0\tau}) \times \left(\sum_{m\in\mathbb{Z}} |E(mf_{\text{pulse}})|^2 e^{2i\pi mf_{\text{pulse}}\tau}\right)$$

*The proof is given in the appendix. Contrary to the model A, the non-zero maximum of this function is not obtained for $\tau = T_{\text{pulse}}$ (see remark in the appendix). Thus the maximum of the autocorrelation function is a biased estimator of the pulse rate in this case.*

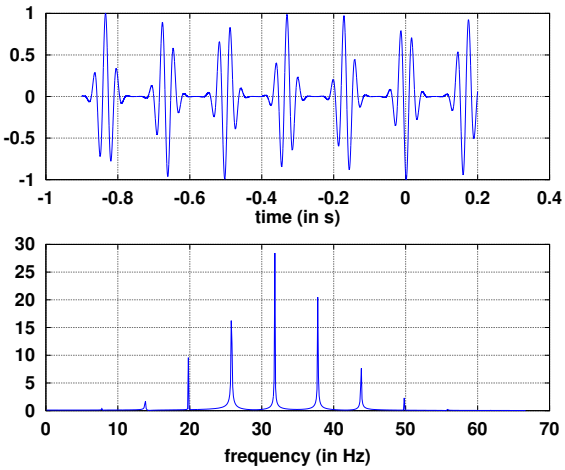


Figure 4: Model B in waveform (top) and its FFT (bottom). In this model, we choose the tonal function $g_{T_0}$ as a pure sine function of period $T_0$ and the envelope $e$ as a Gaussian with standard deviation $\sigma = 0.02$s. As in figure 3, we choose $f_0 = 31.7$ Hz and $f_{\text{pulse}} = 6$ Hz. In the FFT we can see that the frequency peaks are not centered at multiples of $f_{\text{pulse}}$.(color online)

### 2.3 Consequences of the classification of pulsed sounds

**Consequences on measurement methods.** *As can be seen in the previous section, if the pulsed sound is tonal, the autocorrelation function is a good tool to measure the pulse rate $f_{\text{pulse}}$. This can be of importance, because the autocorrelation function method can be much more precise than other methods, as is shown in an example in part 3.5.*

*Thus, our recommendation for a precise measurement of a pulsed sound characteristics would be to:*

- *compute the Fast Fourier Transform (FFT) of the whole pulsed signal (the FFT resolution in frequency is $1/T_{\text{signal}}$, so it is important to have as long a signal as possible);*
- *measure the frequencies $\{f_i\}$, corresponding to peaks in the FFT;*
- *and compute their interval $\Delta f = f_{\text{pulse}}$ or find an approximation of $\Delta f = f_{\text{pulse}}$ by getting the envelope of the signal (the envelope can be obtained by squaring the signal and low-pass filtering it) and then compute the maximum of the autocorrelation function of the envelope;*
- *check whether the signal is tonal or not by examining the quotients $f_i/\Delta f$;*
- *if the signal is tonal, get a better approximation of $\Delta f = f_{\text{pulse}}$ by finding the first maximum of the modulus of the summed autocorrelation function.*

**Consequences on sound production** *If the sound is tonal, then it is compatible with only one active source of sound production altered by a passive filter (source-filter theory). This is the same case as formants [11] in human voice production [14], but the sound will appear 'pulsed' when the first harmonics are not visible in the spectrum. It has been shown to be the case in some musical instruments (e.g. timpani or trombone), or some birds e.g. oscine birds Parus atricapillus [22].*

*If the sound is not tonal (as in model B for instance), then it is the combination of sounds with two different frequencies ($f_0$ and $f_{\text{pulse}}$) that are not linked. In this case, it is not explainable by only one source of energy. Thus, we can infer that two independent organs are used to produce the sound. One produces a signal, and the other acts as an amplitude modulation of the first signal. As shown in part 2.1, the dolphin's buzz is not a tonal signal. One mechanism proposed to explain this rapid train of clicks involves the concomitant action of two generators [4].*

## 3 Application to blue whale songs

### 3.1 The southeast Pacific blue whale song type

*As most baleen whales, blue whales produce high energy, low frequency and long duration vocalizations [8] highly structured in time, with endless repetition of phrases remarkably self similar. Since only males have been reported to produce these sounds [23], they are thought to play a role in reproduction, as happens in birds songs [7]. Interestingly, several songs have been registered for blue whales worldwide and each is characteristic of a population [18]. Several of these song types include pulsed units.*

*In this paper we are interested in a southeast Pacific blue whale song called SEP2, first recorded in 1996 [28] and first described in detail in 2014 [6]. A representation of the repeated phrase is given in figure 5. This phrase, composed of several units, is usually repeated every two minutes, in a sequence lasting from some minutes to a few hours.*

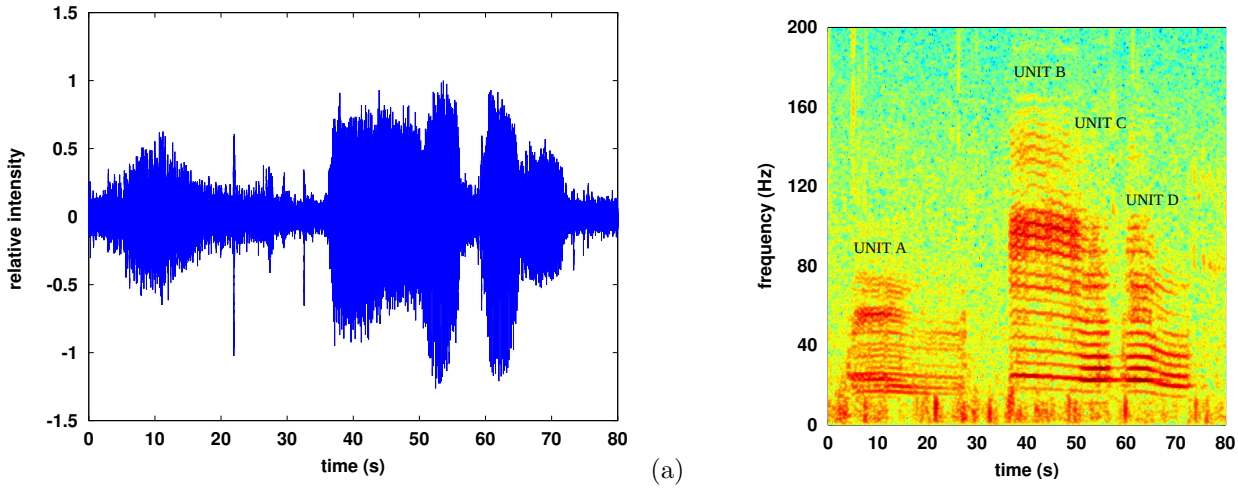


Figure 5: Phrase of the southeast Pacific blue whale song SEP2, recorded off Isla Chañaral, Chile, February 2nd, 2017, sample frequency $f_s$=48 kHz. (a) Left : waveform of relative intensity. (b) Right: time-frequency representation: FFT $2^{12}$ points, overlap of 90%, Hanning window. Low frequency bars are background noise. (color online)

*The pulsed nature of this kind of source is visible in the spectrum because of the various harmonics without the fundamental frequency [30]. This aspect is not due to propagation effect, since all recording of this song show the same aspect, independently of the place and technology of the recording device [16]. Alternatively, it is also visible if we zoom into the waveform, as in figure 1, top. However, the amplitude modulation visible on the waveform is not rectangular as in separated pulses, but rather like a sinusoidal modulation.*

## 3.2 Data collection

*Data were collected close to the Isla Chañaral marine reserve in northen Chile, between the Isla Chañaral and the mainland, at* $29°00'44''$ *south and* $71°31'26''$ *west during the austral summer of 2016/2017, between the 16th of January 2017 and the 27th of February 2017. The hydrophone and recording package 'BOMBYX II' was deployed at 15/20 meters below the surface on a mooring where water column depth was 70 meters. Data were collected during three periods of two weeks in January and February [24]. The hydrophone package 'BOMBYX II' was mounted by the University of Toulon and includes a Cetacean Research C57 hydrophone (very high sensibility, flat response down to 20 Hz, omnidirectional at low frequencies and listening in a plane orthogonal to its axis in high frequencies), alimented by 9 V through a high-pass filter (C=47*$\mu$*F, frequency cut 0.15 Hz) and a commercial SONY PCM-M10 recording device (gain 6, Rin = 22 kOhm) equipped with a 256 GB memory card, set up in a specialized tube made by Osean able to resist high pressure. Recording was done at a sample rate of 48 kHz so as to record a vast diversity of cetaceans, ranging from large whales to dolphins (namely bottlenose dolphins, tursiops truncatus), and at 16-bits, allowing for high sensibility without saturating the memory.*

*A systematic analysis showed that blue whales song were present almost all days of recording (Naysa Balcazar and Giselle Alosilla, private communication). Long series of up to 70 phrases of high signal to noise ratio were recorded, especially on Feb, 2nd 2017.*

## 3.3 Analysis

*On these high signal to noise ratio (SNR) song phrases, we decided to apply our criterion to characterize the nature of these blue whales' 'pulsed' sound. To this end, we measured the peak frequency set* $\{f_i\}$ *and pulse rate* $\Delta f = f_{\text{pulse}}$ *for 100 phrases, extracted on six different days of our recording. For the selected high SNR signals, we analyzed the four units A, B, C and D of the signals (see figure 5) that have different frequency characteristics but are all pulsed. These units are described in detail in [6] or [16].*

**Peak frequency** *For all selected units, we performed a FFT on the first 4s of the unit by a routine in OCTAVE [10]. We measured one of the peak frequencies, which is one that in average shows the higher SNR. This* $f_i$ *is measured as the frequency corresponding to the maximum value (in modulus) of the FFT between 23 and 25 Hz for unit A and between 22 and 26 Hz for units B, C and D. As we did a FFT on* $T_{\text{signal}} = 4s$ *of the signal, there is a quantification of the measure of the frequency which is equal to* $1/T_{\text{signal}} = 0.25$ *Hz and thus the uncertainty on this measure is of the order of 1% [16]. Due to the fact that the precision in frequency is inverse to the duration of the signal, it is important to use as long a signal as possible.*

**Pulse frequency** *The estimation of $f_{\text{pulse}}$ by a difference of two frequencies $\{f_i\}$ obtained by the FFT, would lead to a poor precision, of the order of 8%. Thus, to measure the pulse rate $f_{\text{pulse}}$ of the signal with a better precision, we first performed an envelope detection. To this effect we squared the signal and then low pass filtered it using a fifth order Butterworth filter with frequency cut-off at 10 Hz. Other methods of reconstructing the envelope of the signal can be used [13] giving the same kind of results. Then a summed autocorrelation [33] on the first 4s of the signal was performed to measure the pulse rate [16]. The relative uncertainty on this measure is around 1.5% (see section 3.5).*

## 3.4 Results

*The results of the measures of the ratio between $f_i$ and $f_{\text{pulse}}$ for the four units of the SEP2 phrase are shown in the figure 6.*

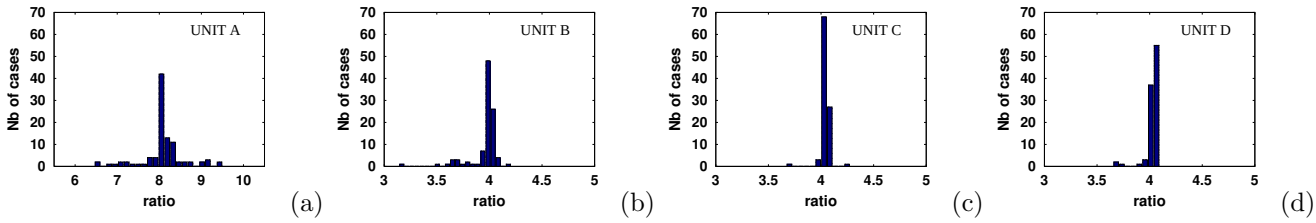


Figure 6: For 100 high SNR SEP2 phrases in 2017, histograms of the ratio between the peak frequency, measured by an FFT, and the pulse rate, measured by envelope detection and summed autocorrelation for units $A$, $B$, $C$ and $D$ (figures (a),(b),(c) and(d) respectively)(color online)

*As we can see, the dispersion of the ratio value $f_i/f_{pulse}$ around a fixed integer number is small, especially for units $C$ and $D$, which usually have a better signal to noise ratio. This dispersion can be explained by errors in measurements (see precedent section), presence of additional low-frequency noise (see figure 5) or variability in the frequency (especially for unit $B$). For unit $A$, the ratio is near 8, and for the other units near 4. Thus our measures are compatible with the hypothesis of a tonal signal for the four units of the SEP2 song phrases. The values of the very low fundamental frequencies $f_0$ (which coincides with the pulse rates $f_{\text{pulse}}$) are given in the table 3 for 2017. This fundamental frequency is very stable between two phrase occurrences of the same year but undergoes a yearly decrease [16].*

Table 3: Mean fundamental frequencies $f_0$ (shown to be the same as the pulse frequency) with standard deviation of the four units of the SEP2 song for the 100 phrases recorded in 2017

| Unit | $A$ | $B$ | $C$ | $D$ |
|---|---|---|---|---|
| $f_{\text{pulse}}$ (Hz) | $2.98 \pm 0.19$ | $6.52 \pm 0.17$ | $5.88 \pm 0.08$ | $5.89 \pm 0.11$ |

## 3.5 Discussion

*As seen in part 2.3, in the case of a tonal pulsed sound, the measure of $f_{\text{pulse}}$ can be done without bias by at least three different methods: FFT of the signal and measure of the gap between two frequency peaks, summed autocorrelation of the envelope of the signal, summed autocorrelation of the signal.*

*In figure 7, we present three histograms of the values of $f_{\text{pulse}}$ (for unit $C$) measured by these three methods on our set of 100 signals.*

*In the case of the measure of the difference between two peaks of the FFT, the result has a quantification value of 0.25 Hz. This value is clearly seen in the figure 7 and this method is an inefficient method to measure $f_{\text{pulse}}$ in this configuration (short duration of the signal compared to the pulse period $T_{\text{pulse}}$). The mean value and standard deviation of the measure is in this case $f_{\text{pulse}} = 5.9 \pm 0.2$ Hz. However, the statistical distribution of the values is far from being a normal distribution (see figure 7, left), so the standard deviation is clearly not a tool that is adapted to this result.*

*In the case of the autocorrelation of the envelope of the signal, we obtain $f_{pulse} = 5.88 \pm 0.08$ Hz. In the case of the summed autocorrelation, we obtain $f_{\text{pulse}} = 5.88 \pm 0.02$ Hz. The best precision is thus obtained by summed autocorrelation of the signal. This justifies the need to be able to discriminate between a tonal and a non-tonal signal by the method presented in the section 2.3.*

*On the other hand, since we have shown that the signal is tonal, it means that, for all the four parts, it is probably produced by only one organ, and filtered by a passive filter that can be the head of the animal or any other part, as in human production of vocals.*

*Interestingly, units of blue whale songs type worldwide are not always tonal sounds as are SEP2 units. For example, in the first unit of the pygmy blue whale song type from south and west Australia (SE Indian song type) [12] and [29] show that, during a song occurrence, the peak frequency increases (up-sweep) while the pulse rate decreases (the gap between frequency bands narrows) [30]. Thus these two frequencies are not linked in a simple way, and are not explained by a source-filter mechanism.*

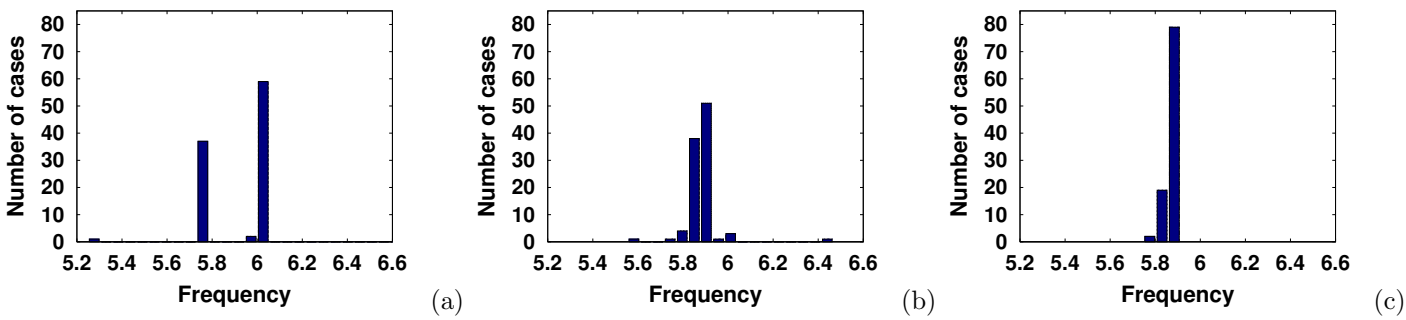


Figure 7: Histograms of the estimation of pulse rate $f_{\text{pulse}}$ for unit $C$ of 100 high SNR blue whale phrases. Three different methods of estimation were employed. (a) Top : estimation by difference of frequency peaks $f_i$ in the FFT. (b) Center : estimation by summed autocorrelation of the envelope of the signal. (c) Bottom : estimation by summed autocorrelation of the signal. See text for comments. (color online)

# 4 Conclusion

*In this paper we presented a tool to better characterize and thus understand a pulsed sound, and, though we applied it to only one example, the SEP2 blue whale phrase, it could be applied to numerous other biological pulsed sounds. The four units of the SEP2 song type are found tonal in spite of their apparent pulsed nature. Thus, the fundamental frequency (or pulse rate) is the better tool to characterize it. Some studies, like the study of frequency decline in baleen whales [17], compare several song types worldwide. It would be useful to have one common criterion to characterize a sound, and the fundamental frequency is probably the best one.*

*The authors thank very much the help of Cesar Villaroel and all the divers of Explorasub diving center (Chile), Agrupación turística Chañaral de Aceituno (Chile) and the research program BRILAM STIC AmSud 17-STIC-01. We are grateful to colleagues at DCLDE 2018 and SOLAMAC 2018 conferences for useful comments on the preliminary version of this work. In this work we used only free and open-source software: Latex, Audacity and OCTAVE.*

*

# A Computation of theoretical formulas and proofs

*In the appendix we presente proofs of the results stated in section 2.2.*

## A.1 Fourier transform of the model B

<u>*Statement*</u> *: If* $s_B(t) = g_{T_0}(t) \times \left[e * \text{Ш}_{T_{\text{pulse}}}\right](t)$*, then* $S_B(f) = \sum_{n\in\mathbb{Z}} E(nf_{\text{pulse}})G_{f_0}(f - nf_{\text{pulse}})$ *is its Fourier transform.*

<u>*Proof*</u> *:* $S_B(f) = \left[G_{f_0} * (E \times \text{Ш}_{f_{\text{pulse}}})\right](f)$
$= \left[\sum_{n\in\mathbb{Z}} E(nf_{\text{pulse}})\delta(f - nf_{\text{pulse}}) * G_{f_0}\right](f)$
$= \sum_{n\in\mathbb{Z}} E(nf_{\text{pulse}})G_{f_0}(f - nf_{\text{pulse}})$

## A.2 Autocorrelation function of the model B

*The signal* $s_{B,\text{finite}} = g_{T_0}(t) \times \left[e * \text{Ш}_{T_{\text{pulse}}}\right](t) \times w(t)$ *is of the form*

$$s_{B,\text{finite}} = \sum_{n\in\mathbb{Z}} a_n e^{2i\pi n f_0 t} \times \Big(\sum_{n\in\mathbb{Z}} e(t - nT_{\text{pulse}})\Big) \times w(t)$$

*considering that* $g_{T_0}$ *is a tonal sound with fundamental equal to* $f_0$ *and thus can be expressed as* $\sum_{n\in\mathbb{Z}} a_n e^{2i\pi n f_0 t}$*.*
<u>*Statement*</u>*: Let a finite pulsed sound*

$$s_{B,\text{finite}}(t) = \Big(\sum_{n\in\mathbb{Z}} e(t - nT_{pulse})\Big) \times \sum_{n\in\mathbb{Z}} a_n e^{2i\pi n f_0 t} \times w(t)$$

*where* $w(t) = 1_{[-\frac{T_{\text{signal}}}{2};\frac{T_{\text{signal}}}{2}]}(t)$ *which satisfies the two hypotheses*

- *the duration of the signal* $T_{\text{signal}}$ *is high compared to* $T_{\text{pulse}}$*;*
- *the bandwidth of* $e$ *is within the interval* $[-f_0/2; f_0/2]$*;*

*then its autocorrelation function is approximately* $C_{s_{B,\text{finite}}}(\tau) \simeq \Lambda(\frac{\tau}{T_{\text{signal}}})(\sum_{n\in\mathbb{Z}} |a_n|^2 e^{2i\pi n f_0 \tau})$
$\times \Big(\sum_{m\in\mathbb{Z}} |E(mf_{pulse})|^2 e^{2i\pi m f_{pulse}\tau}\Big)$, *where* $\Lambda(t)$ *is the triangular function* ($\Lambda(t) = 1+t$ *on [-1;0] and* $\Lambda(t) = 1-t$ *on [0;1] and zero outside of [-1;1]).*
*Proof : The Fourier transform of* $s_{B,\text{finite}}$ *is (see former paragraph)* $S_{B,\text{finite}}(f) = \sum_{m\in\mathbb{Z}} E(mf_{\text{pulse}}) G_{f_0}(f - mf_{\text{pulse}}) * W(f)$
$= T_{\text{signal}}[\sum_{m\in\mathbb{Z}} E(mf_{\text{pulse}}) \times \sum_{n\in\mathbb{Z}} a_n \delta(f - mf_{pulse} - nf_0) * sin_c(\pi T_{\text{signal}} f)](f)$
$= T_{\text{signal}} \sum_{n,m\in\mathbb{Z}} a_n\ E(mf_{pulse}) \times \text{sinc}(\pi T_{\text{signal}}(f - mf_{pulse} - nf_0))$

*The Wiener-Khinchin theorem [32] states that the autocorrelation function* $C_{S_{B,\text{finite}}}(\tau)$ *is the inverse Fourier transform of the spectral density* $|TF(s_{B,\text{finite}})|^2(f)$ *of the signal. Thus*

$$C_{S_{B,\text{finite}}}(\tau) = FT^{-1}\Big(|\sum_{n,m\in\mathbb{Z}} a_n\ E(mf_{pulse}) \times T_{\text{signal}}\ \text{sinc}(\pi T_{\text{signal}}(f - mf_{pulse} - nf_0))|^2\Big)$$

*The two facts that the duration of the signal* $T_{\text{signal}}$ *is high compared to* $T_{pulse}$ *and that the bandwidth of e is within the interval* $[-f_0/2; f_0/2]$ *imply that for a particular t all but one term of this sum are very close to zero. Thus, we can say that*
$C_{S_{B,\text{finite}}}(\tau) \simeq FT^{-1}\Big(\sum_{n,m\in\mathbb{Z}} |a_n|^2\ |E(mf_{\text{pulse}})|^2 \times |T_{\text{signal}}\ \text{sinc}(\pi T_{\text{signal}}(f - mf_{\text{pulse}} - nf_0))|^2\Big)$
$\simeq \sum_{n,m\in\mathbb{Z}} |a_n|^2\ |E(mf_{\text{pulse}})|^2 \times FT^{-1}(|T_{\text{signal}}\ \text{sinc}(\pi T_{\text{signal}}(f - mf_{\text{pulse}} - nf_0))|^2)$
$\simeq \sum_{n,m\in\mathbb{Z}} |a_n|^2\ |E(mf_{\text{pulse}})|^2 \times e^{2i\pi(nf_0 + mf_{\text{pulse}})\tau} \Lambda(\tau/T_{\text{signal}})$
$\simeq \Lambda(\tau/T_{\text{signal}})(\sum_{m\in\mathbb{Z}} |a_n|^2 e^{2i\pi n f_0 \tau}) \times (\sum_{m\in\mathbb{Z}} |E(mf_{\text{pulse}})|^2 e^{2i\pi m f_{\text{pulse}}\tau})$
*Remark The maximum of* $\sum_{m\in\mathbb{Z}} |a_n|^2 e^{2i\pi n f_0 \tau}$ *is obtained when* $\tau$ *is an integer multiple of* $T_0$ *and the maximum of* $\sum_{m\in\mathbb{Z}} |E(mf_{\text{pulse}})|^2 e^{2i\pi m f_{\text{pulse}}\tau}$ *is obtained when* $\tau$ *is an integer multiple of* $T_{\text{pulse}}$. *Thus, in the case of a tonal signal where* $T_{\text{pulse}} = k\ T_0$, *we will have a maximum of* $C_{S_{A,\text{finite}}}$ *for* $\tau = T_{\text{pulse}}$. *In the case of a non-tonal signal* ($f_0/f_{\text{pulse}}$ *is not an integer), we will have a maximum of* $C_{S_{B,\text{finite}}}$ *at the multiple of* $T_0$ *which is the nearest value to* $T_{\text{pulse}}$. *In this case the determination of* $T_{\text{pulse}}$ *by autocorrelation has a bias.*

# References


[1] *O. Adam, D. Cazau, N. Gandilhon, B. Fabre, J.T. Laitman, and J.S. Reidenberg. New acoustic model for humpback whale sound production.* Applied Acoustics, *74:1182–1190, 2013.*

[2] *W. Appel.* Mathématiques pour la physique et les physiciens. *2008.*

[3] *James L. Aroyan, Mark A. McDonald, Spain C. Webb, John A. Hildebrand, David Clark, Jeffrey T. Laitman, and Joy S. Reidenberg.* Acoustic Models of Sound Production and Propagation, *chapter 10, pages 409–469. Springer-Verlag, 2000.*

[4] *W.W.L. Au, A.N. Popper, and R.R. Fay.* Hearing by whales and dolphins. *Springer, 2000.*

[5] *J. C. Brown. Mathematics of pulsed vocalizations with application to killer whale biphonation.* J. Acoust. Soc. Am., *123(5):2875–2883, May 2008.*

[6] *Susannah Buchan, Rodrigo Hucke-Gaete, Luke Rendell, and Kathleen Stafford. A new song recorded from blue whales in the corcovado gulf, southern chile, and an acoustic link to the eastern tropical pacific.* Endangered Species Research, *23:241–252, 2014.*

[7] *K.C. Catchpole and J. B. Slater.* Bird Song: Biological Themes and Variations. *01 1995.*

[8] *W.C. Cummings and P.O. Thompson. Underwater sounds from the blue whale, balaenoptera musculus.* Journal of the Acoustical Society of America, *(50):1193–1198, 1971.*

[9] *Bob Dziak, J Haxel, T.-K Lau, Sara Heimlich, Jacqueline Caplan-Auerbach, David Mellinger, Haru Matsumoto, and B Mate. A pulsed-air model of blue whale b call vocalizations. 7, 12 2017.*

[10] *John W. Eaton, David Bateman, and Soren Hauberg.* GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations. *CreateSpace Independent Publishing Platform, 2009. ISBN 1441413006.*

[11] *J.L. Flanagan.* Speech Analysis Synthesis and Perception. *Springer, 1965.*

[12] *A.N. Gavrilov, R.D. McCauley, C. Salgado-Kent, J. Tripovitch, and C. Burton Wester. Vocal characteristics of pygmy blue whales and their change over time.* . Acoust. Soc. Am. 130 (6), December 2011, *130(6):3651–3660, December 2011.*

[13] *Hervé Glotin. Dominant speaker detection based on voicing for adaptive audio-visual asr robust to speech noise. In* ISCA Tutorial and Research Workshop (ITRW) on Adaptation Methods for Speech Recognition, *2001.*

[14] *D. M. Howard and J. A. S. Angus.* Acoustics and Psychoacoustics. *Elsevier, 2006.*

[15] *Finn B. Jensen, William A. Kuperman, Michael B. Porter, and Henrik Schmid.* Computational Ocean Acoustics. *Springer, 2 edition, 2011.*

[16] *F. Malige, J. Patris, S.J. Buchan, K.M. Stafford, F.W. Shabangu, K.P. Findlay, R. Hucke-Gaete, S. Neira, C.W. Clark, and Herve Glotin. Annual decrease in pulse rate and peak frequency of southeast pacific blue whale song types since 1970. submitted to JASA, 2019.*

[17] *M.A. McDonald, J.A. Hildebrand, and S.Mesnick. Worldwide decline in tonal frequencies of blue whale songs.* Endangered species research, *9:13–21, 2009.*

[18] *M.A. McDonald, S.L. Mesnik, and J.A. Hildebrand. Biogeographic characterisation of blue whale song worldwide: using song to identify populations.* J. Cetacean Res. Manage., *2006.*

[19] *B. S. Miller, K. Collins, J. Barlow, S. Calderan, R. Leaper, M. McDonald, P. Ensor, P.A. Olson, C. Olavarria, and M.C. Double. Blue whale vocalizations recorded around new zealand : 1964-2013.* J. Acoust. Soc. Am., *135(3):1616–1623, March 2014.*

[20] *Ramón Miralles, Guillermo Lara, Estaban Antonio, and Alberto Rodriguez. The pulsed to tonal strength parameter and its importance in characterizing and classifying beluga whale sounds.* J. Acoust. Soc. Am., *131(3):2173–2179, 2012.*

[21] *Janelle L. Morano, Daniel P. Salisbury, Aaron N. Rice, Karah L. Conklin, Keri L. Falk, and Christopher W. Clark. Seasonal and geographical patterns of fin whale song in the western north atlantic ocean.* J. Acoust. Soc. Am., *132(2):1207–1212, 2012.*

[22] *S. Nowicki and R. R. Capranica. Bilateral syringeal interaction in the production of an oscine bird sound.* Science, *(231):1297–1299, 1986.*

[23] *Erin M. Oleson, John Calambokidis, William C. Burgess, Mark A. McDonald, Carrie A. LeDuc, and John A. Hildebrand. Behavioral context of call production by eastern north pacific blue whales.* Mar Ecol Prog Ser, *330:269–284, 2007.*

[24] *Julie Patris, Franck Malige, and Hervé Glotin. Construction et mise en place d'un système fixe d'enregistrement à large bande pour les cétacés "bombyx 2" isla de chañaral, été austral 2017. Technical Report 2017-03, LSIS CNRS, march 2017.*

[25] *D. Reby, M. T. Wyman, R. Frey, D. Passilongo, J. Gilbert, Y. Locatelli, and B. D. Charlton. Evidence of biphonation and source–filter interactions in the bugles of male north american wapiti (cervus canadensis).* Journal of Experimental Biology, *(219):1224–1236, 2016. doi:10.1242/jeb.131219.*

[26] *J.S. Reidenberg. Terrestrial, semiaquatic, and fully aquatic mammal sound production mechanisms.* Acoustics Today, *13(2), 2017.*

[27] *W. John Richardson, Jr. Charles R. Greene, Charles I. Malme, Denis H. Thomson, Sue E. Moore, and Bernd Wiirsig.* Marine Mammals and Noise. *Academic Press, 1995.*

[28] *Kathleen M. Stafford, Sharon L. Nieukirk, and Christopher G. Fox. Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical pacific.* J. Acoust. Soc. Am., *106(6):3687–3698, 1999.*

[29] *K.M. Stafford, E. Chapp, and D.W.R. Bohnenstiel. Seasonal detection of three types of "pygmy" blue whale calls in the indian ocean.* Marine mammal science, *27(4):828–840, 2011. DOI: 10.1111/j.1748-7692.2010.00437.x.*

[30] *W.A. Watkins, editor.* The Harmonic interval fact or artifact in spectral analysis of pulse train, *volume 2, American Museum of Natural History, New York, April 13-15 1968. Pergamon Press-Oxford New-York.*

[31] *M. J. Weirathmueller, K. M. Stafford, W. S. D. Wilcock, R. S. Hilmo, R. P. Dziak, and A. M. Trehu. Spatial and temporal trends in fin whale vocalizations recorded in the ne pacific ocean between 2003-2013.* PLOS ONE, *12(10), 2017. https://doi.org/10.1371/journal.pone.018612.*

[32] *N. Wiener. Generalized harmonic analysis.* Acta mathematica, *55, 1930.*

[33] *J.D. Wise, J.R. Caprio, and T.W. Parks. Maximum likelihood pitch estimation. In IEEE, editor,* Transactions on acoustics, speech and signal processing, *volume ASSP-24, 1976.*