Applied Acoustics 74 (2013) 1182–1190

Contents lists available at SciVerse ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

New acoustic model for humpback whale sound production

Olivier Adam [a,b,*], Dorian Cazau [a], Nadege Gandilhon [c], Benoît Fabre [a], Jeffrey T. Laitman [d], Joy S. Reidenberg [d]

[a] Institut Jean Le Rond d'Alembert, University UPMC Paris 6, CNRS UMR 7190, Lutheries Acoustique Musicale (LAM) Team, Paris, France
[b] Centre de Neurosciences Paris Sud, University Paris Sud, CNRS UMR 8195, Bioacoustics Team, Orsay, France
[c] Laboratoire de Biologie Marine, Université des Antilles et de la Guyane, Dynecar Team, France
[d] Center for Anatomy and Functional Morphology, Mount Sinai School of Medicine, NY, USA

ARTICLE INFO

Article history:
Received 1 December 2012
Received in revised form 10 April 2013
Accepted 15 April 2013
Available online 22 May 2013

Keywords:
Bioacoustics
Humpback whale
Baleen
Sound generator

ABSTRACT

The mechanism by which baleen whales (Mysticeti) produce sounds has remained largely unknown, due in part to our limited knowledge of the relationship between the sound-producing anatomy and the vocal characteristics of calls. Recent studies on mysticete anatomy indicate that the laryngeal vocal folds are the sound source, and the surrounding air spaces may play an important role in airflow, and sound modification or transduction. This current study offers a theoretical model to describe the mysticete vocal production system, which is much more complex than that of typical terrestrial mammal species. Metric data from laryngeal structures and air space volumes are combined with frequency and duration ranges defined by recordings of humpback whales off the coast of Madagascar. The resulting model delivers a prediction of sound unit durations and frequency formants that are constrained by the measurements of the trachea, laryngeal sac, and nasal cavities. Results predicted by the model are comparable to those obtained from real recordings. Errors between the frequencies of real vocalizations and the frequencies estimated using our theoretical model are less than 60 Hz for the low frequency band. Then, this new model should hopefully advance our growing understanding of sound generation in humpback whales, and mysticetes in general.

1. Introduction

Physiological mechanisms of the sound generation in baleen whales remain a current subject not well understood. Noticeably, they have long been thought to have no vocal folds, a misconception frequently held even through the scientific literature [8,63,20,34,47,48,44,28]; see also review in [51]. Actually, marine mammal anatomy has been studied over the last century [4,5,34,61,31,49], but without investigations on the potential functions of the larynx in the sound generation. These functions were reported and discussed by more recent works, in particular through the descriptions of the lungs and the larynx [21,47,38]. However, some analyses were limited by the apparent absence of glottal vibrators and others wrestled with the question of radiating ends in the sealed vocal system of humpback whales during vocalizations. Their conclusions are largely confined to details of the structure of a potential vocal generator located in the larynx(e.g., [50,51,52]. Furthermore, analogy-based deductions with an extensively studied related species, the odontocetes, are not possible as mysticetes do not present similar anatomical details, such as nasal air sacs and fatty tissues Mead, 1975; [2,37,54]. As a direct consequence, no production-related descriptions of mysticete sound characteristics have been feasible, although such sounds referring to one or different specific parts of the sound-producing anatomy are commonly used for many different animal species. Examples are the nasal complex of toothed whales for echolocation clicks [14], the supra-laryngeal system for human and terrestrial animal formants, (e.g. Fant [22] for humans, Vannoni and McElligott [68] for male deers, Fitch and Hauser [24] for non-human primates, Riede and Fitch [55] for dogs) However, many analysis were done on songs that some individual male humpback whales, *Megaptera novaeangliae,* emit during the breeding seasons, using the definition of sound units (SUs) proposed by Payne and McVay [42]. The SU, sometimes just called *units* in the scientific literature, are continuous sounds between 2 silences. Classical characteristics of SU are, in the time domain, the duration, the attack phase, the decay phase, the sustain phase, the release phase of the sounds, and, in the frequency domain, the fundamental frequency, the presence of harmonics. Silences could also be considered in the analysis. Successive SU compose phrases and some successive phrases are repeated to form the theme. A combination of themes gives the leitmotiv of the song.

The SU contents change with geographical areas or from one year to another, and even during the season [27,9]. However the

* Corresponding author at: Institut Jean Le Rond d'Alembert, University UPMC Paris 6, CNRS UMR 7190, Lutheries Acoustique Musicale (LAM) Team, Paris, France. Tel.: +33 620 744 607.
E-mail address: olivier.adam@upmc.fr (O. Adam).

0003-682X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apacoust.2013.04.007

songs can also be transferred from one population to another [42]. Potential social roles of these songs have also been investigated [71,66,25,32,1,16] and used to classify SU with associated behaviors. Dunlop et al. [19] describes the different social sounds into 6 main groups of vocalizations: (i) low-frequency sounds, (ii) mid-frequency sounds, (iii) high-frequency sounds, (iv) amplitude modulated sounds, (v) repetitive sounds, and (vi) broadband, "noisy" and complex sounds. From acoustical observations during the 2002 winter season in Hawaiian waters, Au et al. [2] suggested 9 different types of SU including: vibrating upsweep, double upsweep, frequency sweeping cry to short low-frequency upsweep and mid-frequency tonal wail. SU classification is not trivial [43]. For example, the number of SU classes is still not well established. To go further, we propose to use the recent knowledge from anatomy to inform about the humpback whale sound production system. As all vocalizations are channeled through the production system, its biomechanical characteristics determine what sounds are physically producible. Then, a better understanding of the complete system should help to identify the relevant types of SU and then should help to define the different classes for these SU. Until now, two studies have proposed a model of sound generation based on the anatomical and physiological mechanisms of humpback whales [1,53]. Reidenberg and Laitman [53] discovered the presence of vocal folds in mysticetes (Fig. 1). Reidenberg and Laitman [54] also revealed the presence of the laryngeal sac that would receive air (Although the term of "air" will be mostly used through this text, a reference is actually implicitly made to any kind of internal gases flowing into the respiratory tractus of the whale) from the lungs through the trachea. Then, the authors propose that sounds might be made bidirectionally, i.e., with airflow both into (ingressive) and out of (egressive) the lungs.

In this paper, we introduce a new theoretical acoustic model for the humpback whale sound production system based upon past anatomical investigations. Explanations are given about the vibrator's mechanics and the directions of airflow in this model. Spectrum of sounds provided by this theoretical model will be compared to spectrum provided from SU detected in real signals recorded in Madagascar over three successive years.

Objectives of this paper are: (1) to propose a theoretical model corresponding to real recordings of MN sounds, (2) to explain sound characteristics based on the anatomy of the vocal tract, (3) to justify the specific quasi-vertical body position of MN singers, and (4) to suggest a new theory, called the "4Ls" theory, from the specific characteristics of these songs.

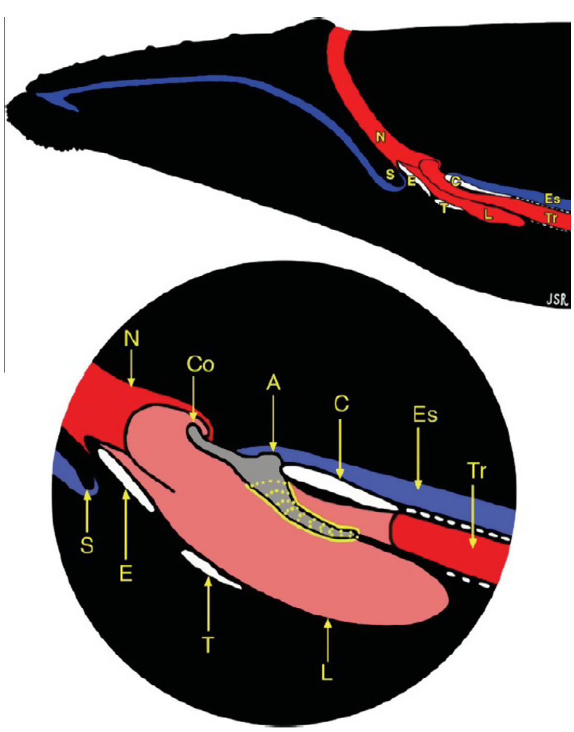

Fig. 1. laryngeal anatomy for baleen whales (from [53] (A : arytenoid cartilage, C : cricoid cartilage, Co: corniculate cartilage, E : epiglottic cartilage, Es: esophagus, L : laryngeal sac, N : nasal passageway/nasopharynx, S : soft palate, T : thyroid cartilage, Tr: trachea).

2. Method and materials

2.1. Humpback whale recordings

Humpback whales were recorded in the Ste Marie Channel, Madagascar, during August 2007, August 2008 and August 2009. Visual and acoustic observations were done close to the singer from the motor-boat (motor off). Special care was taken so that nomarine traffic was around the studied area, and that the World Metrological Organization (WMO) sea states were always below than 3. Our acoustic dataset is based on 44 h of MN songs (cf. Table 1) recorded with the ColmarItalia GP0280 hydrophone (omnidirectional, [5–90 kHz], 170 dB re 1 V/μPa), analog lowpass band filter [0–20 kHz] and digitalized by the recorder Tascam HD-P2 (16 bits, Fs = 44.1 kHz). The maximum duration for the continuous recordings is around 1 hr. The hydrophone was at 15–18 m depth (the water depth was between 30 and 40 m) and the distance from the singer approximately 100 m to avoid the frequency attenuation due to the acoustic propagation. Sounds from more than one singer are audible in the background. Although it is almost impossible with separate an individual from others using recordings from a single hydrophones, we selected the SUs with the highest signal-to-noise ratio, assuming that these sounds were emitted by the closest MN singer.

Initially, parts of these songs were manually segmented from the spectrogram analysis to extract the beginning and the end of SUs, with the objectives (1) of the estimation of the maximum duration of the longest SUs , (2) the duration of the longest silences between 2 successive SUs, and (3) the largest frequency ranges (including harmonics). This manual segmentation was used to set the parameters of the automatic segmentation algorithm. It is based on the detection of at least one frequency using a sliding 100 ms-length Hanning window (overlap 50%) [46]. To minimize the false alarm rate, we made the choice to define: (1) a minimal duration for SU (empirically fixed at 100 ms), (2) a threshold on the time-energy signal (the threshold based on the mean acoustic intensity was fixed for each recording), and (3) that the most energetic frequency should be above 20 Hz [46].

2.2. The theoretical sound producing model

Our model of the sound generator is based on the general mysticete anatomy (Fig. 1), combining the trachea tube (T), the laryngeal sac (L) and the paired nasal cavities (N) (Fig. 2). All these air-filled pipes are considered as uniform tubes. The T tube is assumed to be closed at one side (vocal folds) and opened at the

Table 1
Recordings during 3 years in the Sainte Marie Channel (Madagascar).

Season	Period (days)	Recordings (h)	*N* songs
August 2007	16	15	6
August 2008	15	12	7
August 2009	16	17	9