

# Sound production by singing humpback whales

Eduardo Mercado III<sup>a)</sup>

Department of Psychology, University at Buffalo, The State University of New York,  
Buffalo, New York 14260

Jennifer N. Schneider

Evolution Ecology and Behavior, University at Buffalo, The State University of New York,  
Buffalo, New York 14260

Adam A. Pack

Department of Psychology, University of Hawaii at Hilo, 200 West Kawili Street, Hilo,  
Hawaii 96720

Louis M. Herman

The Dolphin Institute, 91-117 Kekai Place, Kapolei, Hawaii 96707

(Received 2 June 2009; revised 8 January 2010; accepted 15 January 2010)

Sounds from humpback whale songs were analyzed to evaluate possible mechanisms of sound production. Song sounds fell along a continuum with trains of discrete pulses at one end and continuous tonal signals at the other. This graded vocal repertoire is comparable to that seen in false killer whales [Murray *et al.* (1998). *J. Acoust. Soc. Am.* **104**, 1679–1688] and human singers, indicating that all three species generate sounds by varying the tension of pneumatically driven, vibrating membranes. Patterns in the spectral content of sounds and in nonlinear sound features show that resonating air chambers may also contribute to humpback whale sound production. Collectively, these findings suggest that categorizing individual units within songs into discrete types may obscure how singers modulate song features and illustrate how production-based characterizations of vocalizations can provide new insights into how humpback whales sing.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3309453]

PACS number(s): 43.80.Ka, 43.80.Ev, 43.60.Uv, 43.25.Ts [WWA]

Pages: 2678–2691

## I. INTRODUCTION

The mechanisms of sound production used by an animal constrain the types of sounds that the animal can produce. Anatomical and physiological constraints on production may in turn limit the ways an animal can use sounds. For example, a smaller animal may not be able to generate the sound amplitudes or frequencies necessary for it to compete reproductively with larger conspecifics (Fitch and Hauser, 1995). Similarly, the amplitudes and frequencies produced by an echolocating animal limit the size and distance of the targets it can detect (Au, 1993). Knowing how a species physically produces sounds can thus provide a useful foundation for understanding the functions that vocalizations may serve, as well as for assessing physical features of the individual producing the sound. The purpose of this paper is to use acoustic data and models of phonation in terrestrial mammals to investigate sound production by singing humpback whales (*Megaptera novaeangliae*).

### A. Anatomy of humpback whale respiratory structures

Most mammals produce sounds by forcing air from the lungs through the glottis, causing membranes (typically vo-

cal folds) to pulsate (Keleman, 1963). Air oscillations or turbulence generated by these membranes are a primary source of sound. As the sound waves travel out of the animal's body, they are modified by the resonating properties of the structures they pass through or by. For example, the vocal tract can act as a resonator, filtering the sound to emphasize certain frequencies while dampening others. As a result, the final form of a produced sound depends on both the initial vibrations produced at the source, and the length, shape, and tissue composition of the vocal tract. Cetaceans produce sounds underwater, necessitating somewhat different mechanisms from those used by terrestrial mammals. Nevertheless, in the few cetacean species that have been studied, the source mechanism still appears to involve pneumatically driven membrane vibrations which travel through internal structures that may impact the acoustic properties of the emitted sound (Au, 1993; Cranford *et al.*, 1996; Cranford and Amundin, 2004). Air-filled chambers are thought to play little role in sound production by toothed whales (however, see Ridgway, *et al.*, 2001).

Humpback whales are the most vocally versatile of the great whales, well known for the long and elaborate songs produced by males during the winter breeding season (Payne and McVay, 1971; Winn *et al.*, 1970). The mechanisms that humpback whales use to internally generate sound are not well understood. Several researchers have postulated that humpbacks produce sounds by pushing an air stream over or through membranes in the larynx as terrestrial mammals do

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: emiii@buffalo.edu

(Purves and Pilleri, 1983; Quayle, 1991; Reidenberg and Laitman, 2007; Schevill, 1964; Sukhovskay and Yablokov, 1979; Reeb and Best, 1999). Anatomical examinations of the respiratory tracts of mysticetes have begun to provide a clearer picture of how humpback whales may produce sounds (Quayle, 1991; Schevill, 1964; Sukhovskay and Yablokov, 1979). Quayle (1991) noted similarities between the human larynx and the larynx of a humpback whale calf suggesting that this structure is capable of sound production. Reidenberg and Laitman (2007) identified a U-fold between the upper and lower respiratory systems as a possible source of vibrations. Unlike typical vocal folds in terrestrial mammals, the U-fold is oriented parallel to airflow. The U-fold borders the entrance of a large laryngeal sac, which may function as a resonance chamber or as a receptacle for air exhaled from the lungs (Reeb and Best, 1999; Reidenberg and Laitman, 2007). Sound may be produced when air flowing from the larynx passes the U-fold, causing vibrations in the air column within the larynx (or laryngeal sac). The vibrating air (or membranes) could then generate pressure fluctuations that are transferred through surrounding tissue into the water.

Singing humpback whales can produce sounds underwater for 15 min or more without any air escaping from either the blowhole or mouth (personal observations, A. A. P. and L. M. H.), suggesting that they recirculate air internally during sound production. How such recirculatory processes are coordinated with sound production is currently unknown.

## B. Models of mammalian sound production

Most hypotheses about sound production by humpback whales are based on anatomical measurements because it is currently not possible to directly observe what humpbacks are doing internally while producing sounds. The dynamics of sound production can be indirectly observed, however, through examination of the acoustic features generated during production (Gaunt, 1983; Taylor and Reby, 2009). For example, techniques developed to describe human speech sounds have greatly increased our understanding of the dynamics of speech production, as well as the production of nonspeech sounds by primates (Owren *et al.*, 1997; Fitch, 1997; Rendall *et al.*, 1998), birds (Beckers *et al.*, 2004; Riede *et al.*, 2004), and deer (Reby *et al.*, 2005; Vannoni and McElligott, 2007). Comparable techniques for characterizing humpback whale sounds might similarly provide insights into the processes underlying their sound production.

One classical approach to modeling mammalian sound production (especially human speech) is a signal processing model called the source-filter model (Fant, 1960). In this model, the “source” simulates features of sound production related to vocal fold vibrations or air turbulence, and the “filter” simulates the transformative effects of the vocal tract. This model can be physically instantiated by combining a silicone vocal fold model with a PVC tube (Riede *et al.*, 2008). From a signal processing perspective, cetacean sound production can be characterized using a source-filter model just as it is for terrestrial mammals—the properties of the source and filter simply need to be adjusted to account for

the differences in cetacean anatomy (e.g., see Altes *et al.*, 1975). Homologies between the sound production mechanisms hypothesized in humpback whales and those present in terrestrial mammals (Reidenberg and Laitman, 2007), increase the likelihood that the source-filter model can adequately characterize humpback whale sound production. For example, U-fold vibrations may be modeled in the same way as vocal fold vibrations, and the effects of the laryngeal sac and/or larynx on these vibrations can be modeled in the same way as vocal tract filtration.

Past analyses of animal sounds based on source-filter models of production have revealed that animals attend to information about vocal filters (Charlton *et al.*, 2007; Fitch and Fritz, 2006; Sommers *et al.*, 1992; Taylor *et al.*, 2008) and that resonant frequencies (called formants) may be used for individual identification and body size (reviewed by Taylor and Reby, 2009). Formant distribution is correlated with individual identity in grunt calls of chacma baboons (Owren *et al.*, 1997; Rendall, 2003) and coos and grunts of rhesus monkeys (Rendall *et al.*, 1998). Both formant features and fundamental frequency may be used for individual identification in fallow deer groans (Vannoni and McElligott, 2007) and African elephant rumbles (Soltis *et al.*, 2005). Formant distribution has been correlated with body size in rhesus macaques (Fitch, 1997), domestic dogs (Riede and Fitch, 1999), red deer (Reby and McComb, 2003), and southern elephant seals (Sanvito *et al.*, 2007). Development of comparable models for humpback whale sound production may similarly provide new ways of assessing individual differences in singing humpbacks.

## C. Singing

Humpback whales produce sounds in a variety of contexts. Researchers have focused most of their attention, however, on sound production during singing. Humpbacks are described as singing when they rhythmically produce sounds (called units) in predictable sequences (Payne and McVay, 1971). The features of humpback whale songs and their constituent units are described in numerous papers (see e.g., Winn and Winn, 1978; Payne and Payne, 1985; Helweg *et al.*, 1998; Mercado *et al.*, 2005). The broad range of frequencies, durations, intensities, and spectrotemporal patterns that humpback whales produce within songs makes songs particularly useful for understanding the dynamics of production mechanisms. Additionally, the fact that songs are often produced by stably positioned, lone individuals makes it easier to obtain recordings of sufficient quality for decomposing units into production-related parameters.

In terrestrial mammals, singing behavior typically involves respiratory patterns and production modes that differ somewhat from nonsinging (Doscher, 1994; Miller, 1986). Nevertheless, the physical mechanisms engaged during singing are typically the same mechanisms used to produce nonsong sounds. For example, human song contains the same basic sounds present in speech. These sounds can be divided into three broad classes based on their periodicity: voiced, unvoiced, and mixed sounds. Voiced sounds (e.g., vowels) are quasiperiodic pulse trains generated when the vocal

folds vibrate in relaxed oscillations. Unvoiced sounds (e.g., most consonants) are aperiodic noise bursts generated by turbulence or an abrupt release of air. Mixed sounds have features of both voiced and unvoiced sounds. If humpback whales sing using vocal mechanisms analogous to those used by humans, then voiced components should be evident in their units.

#### D. Source vibrations

In the source-filter model, two main properties characterize the source: the fundamental frequency or pulse rate and the pulse shape. Voiced sounds are modeled as a periodic series of impulses, the rate of which is quantified using a single value for constant pulse rates, or as a time series for time-varying pulse rates. Pulse rate is determined by the tension of the vocal folds, as well as the pressure below the folds, and typically is perceived as the pitch of the sound. Pulse shape is more difficult to characterize and may be modeled using specific waveforms or spectra that are modified to simulate empirically measured features of individual pulses, some of which may be pulse rate dependent. Pulse shape reflects the degree of vocal fold closure during oscillation and contributes to the perceptual qualities (or timbre) of a voiced sound. In singing humans, pulse shape usually varies with pulse rate (Cleveland, 1994; Gregg, 1992). Acoustically characterizing the “source” of humpback whale sounds within a source-filter model thus entails measuring variability in periodicity, fundamental frequency, and waveform shape within and across units.

When human singers reach certain pitch thresholds, their vocal mechanisms shift into different modes of vibration (Cleveland, 1994; Titze, 2008a). These transitions divide the pitches of singers into four vocal registers: vocal fry, modal, falsetto, and whistle. The vocal fry register includes the slowest pulse rates that singers can produce (5–60 Hz). In this register, individual pulses are perceptually discriminable. The modal register is associated with slightly higher pitches (80–300 Hz). In this register, individual pulses are broadband giving rise to multiple harmonics. The falsetto register is associated with higher pitches (250–900 Hz) and pulses tend to be narrowband, giving rise to fewer harmonics (Miller, 1986). The whistle register includes very high-pitched sounds (850–2300 Hz) that have few harmonics. The full range of song sounds produced by human singers falls along a graded continuum, with sounds containing trains of discrete pulses at one end and sinusoidlike sounds at the other.

Most of the sounds produced by human singers are subjectively very different from those produced by cetaceans. Nevertheless, recent analyses of the acoustic structure of clicks, burst-pulse sounds, and whistles produced by false killer whales (*Pseudorca crassidens*) have revealed that these sounds fall along a similar graded continuum (Murray *et al.*, 1998). At one end of this continuum are trains of clicks that can be modeled as exponentially damped sinusoidal pulses. At the other end are whistles, which are highly similar to sinusoids. False killer whales often gradually modulate from click trains to whistles and can rapidly shift between signal types, for example, shifting immediately from a whistle to a

click train without changing the frequency with peak energy across the transition. Such vocal dexterity is consistent with the proposal that toothed whales, like humans, generate sounds by using air to control the rate at which paired membranes vibrate (Cranford *et al.*, 1996; Brown, 2008).

The current study quantitatively assessed whether humpback whale song units are best described in terms of discrete sounds types, or as points along a graded acoustic continuum. Specifically, the current analysis tested the hypothesis that sounds described as ratchets, moans, and cries (e.g., Winn and Winn, 1978) within humpback whale songs represent points along a pulse rate (i.e., vocal fold tension) continuum. If the sounds of humpback whales fall along a similar continuum to that seen in humans and false killer whales, then this would provide support for Reidenberg and Laitman's (2007) proposal that humpbacks generate sounds by passing air over or through vocal folds.

#### E. Formants

Voiced sounds are categorized based on their spectral features, which are a result of the resonance properties of the vocal tract (i.e., the “filter”). Specifically, they are classified based on the positions and dynamics of regions of frequency enhancement called formants. Generally, about 4–5 formants are evident in sung sounds and in spoken vowels. For males the first formant is typically between 250–700 Hz, and the second formant is between 750–2500 Hz (Sundberg, 1977). The pattern of formants produced by human singers is highly pitch dependent (Rossing *et al.*, 1987; Sundberg, 1977). Lower pitched sounds tend to show more clustering of higher formants than higher pitched sounds. Singers may adjust their formant positions to facilitate the transmission of certain high-pitched sounds by changing the size of their mouth opening and jaw position (Cleveland, 1994; Schutte *et al.*, 1995; Titze, 1994). Some singers are also able to tune their formant regions to specific upper harmonics (called overtone singing) such that they emphasize different frequencies while maintaining a steady pulse rate (Boersma and Kovacic, 2006; Sundberg, 1977).

The prevalence and properties of formants within humpback whale sounds have not previously been analyzed in any detail. Analyses of spectral features of units provide a way to characterize the humpback whale “filter” in terms of formant dynamics and to determine the possible role of air chambers in humpback whale sound production. The current study examined the spectral properties of units to determine whether units contain formants. In terrestrial mammals, formants are the primary acoustic evidence that internal air chambers shape the features of vocalizations. Furthermore, the patterns of formants that distinguish certain classes of sounds (e.g., vowels) correlate with changes in the shape of the vocal tract, and also provide an indication of the length of the vocal tract. If air-filled chambers contribute to sound production by humpback whales, then the properties of these resonators should be evidenced by systematic formant patterns in units.

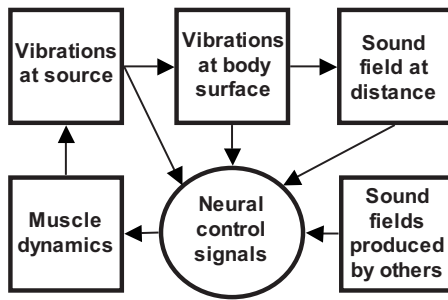


FIG. 1. Multiple internal and external mechanisms contribute to the signals recorded from vocalizing humpback whales. Production-based analyses decompose recorded sounds into acoustic components generated by different mechanisms. Humpbacks neurally control muscle dynamics that affect air-flow as well as the tension of vibrating membranes that are the source of vocalizations. These vibrations are transformed as they travel through the whale's body such that emitted vibrations differ from those generated at the source. Sounds are further distorted as they travel through the ocean. Auditory and somatic sensations provide feedback about produced and received sounds that can potentially influence control mechanisms. By identifying acoustic components resulting from membrane vibrations, internal transformations, and environmental filtering, one can better understand which variations vocalizing humpback whales control and how they control them.

## F. Study goals

The ultimate goal of the current analyses was to better understand how humpback whales sing. A first step in many past analyses of humpback whale song has been to categorize individual sounds into discrete types (often designated using letters or nominal labels) that are then used to describe the structure of songs or the vocal repertoire of singers. Such classifications emphasize the distinctive properties of “prototypical” units heard from a distance without considering what the singer is doing when it sings. In contrast, the production-based analyses reported here attempt to distinguish acoustic features that singers can potentially control from those that they cannot (Fig. 1). Decomposing recorded sounds into components attributable to source vibrations, the filtering effects of tissue transmission, or environmental filtering can help address questions such as the following: do singing whales use similar vocal actions to produce different sound features? If so, could apparent switching between discrete sound types actually reflect more gradual adjustments of vocal control mechanisms? How much control do whales have over the acoustic properties of sound sequences? Which, if any, properties of produced sounds are inadvertent? Knowing more about how singing humpback whales produce sounds can clarify appropriate techniques for describing the properties of songs as well as the vocal acts of singers.

## II. MATERIALS AND METHODS

### A. Study sites and subjects

Singing humpback whales were recorded in waters off the coast of the island of Hawaii between 1981 and 1995, and off the coast of Puerto Rico in 2009. Humpbacks in these two regions, as well as in other regions worldwide, are considered a single species. Recordings were collected during the winter months (February–April) when singing whales can be found in these areas. Singers in both regions are often

separated by many kilometers, and although songs are typically produced at high source levels (Au *et al.*, 2006), in most of the recordings used for these analyses only a single singer was audible. Additionally, ambient noise levels in these waters are lower than in areas where numerous whales can be heard singing simultaneously (e.g., in waters off the island of Maui), and singing humpback whales often maintained a relatively stationary position while singing. Collectively, these features made it possible to collect high quality recordings of individual singing whales.

### B. Recordings and acoustic analysis

All recordings were made from one or more hydrophones suspended approximately 10 m beneath a small boat located in relatively close proximity to the singer being recorded. Various configurations of hydrophones and recording systems were used across years. Recordings collected in Hawaii were made on analog cassette tapes, whereas recordings from Puerto Rico were stored digitally in .wav formatted files. All recordings were imported into the MATLAB programming environment as vectors representing time-domain waveforms. The sampling rate for digitization of Hawaiian songs was 8 kHz, and for Puerto Rican songs was 44 kHz. Signal processing techniques derived from both source-filter modeling, speech analyses, and prior analyses of human song sounds were applied to recordings using a combination of functions from MATLAB's Signal Processing Toolkit as well as customized MATLAB programs. Visual inspections of spectrograms and basic measures of spectrographic features were made using both MATLAB and RAVEN 1.3.

### C. Source characteristics

All measurements of unit source characteristics were collected from a single, 54 min segment of a song session digitally recorded in Puerto Rico in 2009. This recording was selected because it was the highest fidelity recording available and because earlier analyses of lower fidelity recordings showed that the range of pulse rates within songs is relatively similar across individuals and years (Mercado *et al.*, 2005). Use of a single high quality sample controls for variability attributable to individual differences between singers, environmental conditions, time of recording, interference from other sound sources (including singers), and differences in recording systems. This approach is comparable to a case study of an individual human singer during a single performance and is not intended to represent the full vocal repertoire of humpback whales.

Short-term analyses of pulse rate in human speech generally are based on a 10–30 ms waveform sample, which includes three to five periods. Automated techniques developed to extract the pitch of speech can give erroneous results when applied to the full range of rates in units produced by singing whales because too few or too many periods are contained within a 10 ms sample frame; low pulse rate sounds are particularly problematic (Ishi *et al.*, 2008). For medium to high pulse rate sounds, pulse rate can be estimated from spectrograms or spectra by measuring the spacing of spectral bands (Watkins, 1967), but lower pulse rate

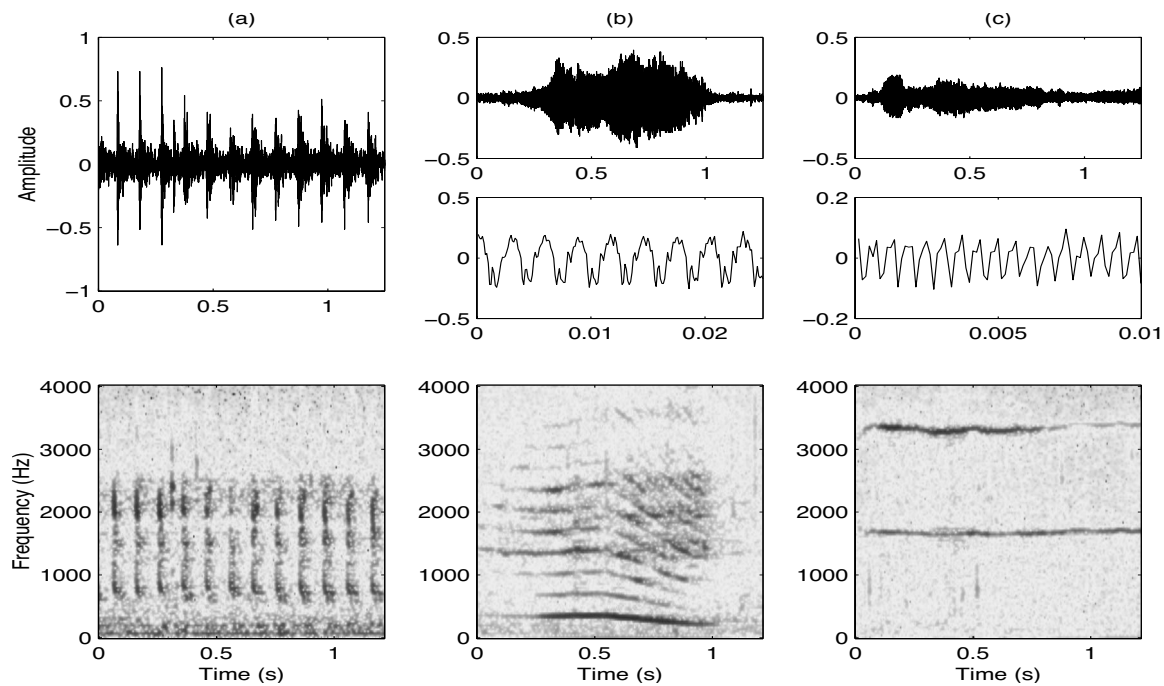


FIG. 2. (a) Low pulse rate units appear as discrete pulses in waveforms (top) and as a series of vertical lines in spectrograms (bottom). (b) Medium pulse rate units typically appear as a coherent increase in amplitude in low-resolution waveforms (top), as quasiperiodic, continuous signals similar to those of human vowels in high-resolution waveforms (middle), and as closely spaced horizontal bands in spectrograms (bottom). (c) High pulse rate units are typically indistinguishable from medium rate units in low-resolution waveforms (top), appear as continuous, sinusoid-, or sawtoothlike signals in high-resolution waveforms (middle), and as widely spaced horizontal bands in spectrograms (bottom). Fast Fourier transform (FFT) size=256 Hz for all spectrograms.

sounds often do not show clear bands without careful selection of the analysis window. In the current analyses, waveforms of individual units were visually inspected and dichotomously classified based on whether discrete repeating pulses were evident [as in Fig. 2(a)] or were not [as in Figs. 2(b) and 2(c)]. For units containing discrete pulses, short-term analyses were applied to nonoverlapping 200 ms sample frames. For all other units, 25 ms frames were used. A MATLAB function (“shrp”) was used to estimate the pulse rate within frames based on the subharmonic-to-harmonic ratio (Sun and Xu, 2002).

In past work characterizing pulsed sounds produced by false killer whales, pulse shape was modeled as an exponentially damped sinusoid based on observations of recorded vocalizations (Murray *et al.*, 1998). The only published report of individual pulses produced by humpback whales shows a representative “megapclick” that is readily modeled as an exponentially damped sinusoid (Stimpert *et al.*, 2007). In the current study, visual inspection of unit waveforms revealed a much broader variety of pulse shapes than those seen in false killer whales (Fig. 2). Although some pulses within units were adequately modeled by damped sinusoids [Fig. 2(a)], others were more comparable to those seen in human speech [Fig. 2(b)]. Pulses within higher pulse rate units often were best modeled as one period of a triangle wave, sawtooth wave, or distorted sinusoid [Fig. 2(c)]. Developing a model of pulse shape that captures the full range of features present in units is beyond the scope of the current paper. Instead, duty cycle was used to characterize variability in waveform shape across units (Murray *et al.*, 1998). Duty cycle refers to the percentage of time a signal is on relative to the duration of the signal. The duty cycle measurement used

here quantifies the similarity of the unit waveform to a sine wave as a value between 1 (a sine wave) and 0 (an impulse train). Duty cycle for units like the one shown in Fig. 2(a) would thus be lower than for the unit in Fig. 2(b), which would in turn be lower than the duty cycle of the unit in Fig. 2(c). Duty cycle was calculated for each sampled frame by rectifying the frame, converting all of its points to a percentage of the maximum value, averaging these percentages, and then normalizing this average relative to that of a sine wave by dividing the average by 0.636. The frequency with peak energy was also calculated for each frame.

#### D. Filter characteristics

Formants can be measured in human singers by having them produce broadband vocalizations such as frequency sweeps, vocal fry, or noisy vocalizations (e.g., see Titze, 2008b). This technique is comparable to using an impulse or sequence of sinusoids to determine the frequency response of a filter. As an initial test for formants in humpback whale sounds, power spectral density functions were calculated [using MATLAB’s “psd” function, with a 128 point fast Fourier transform (FFT)] for units with acoustic features similar to those used to detect formants in human singers. Spectral analyses of these types of vocalizations provide the best opportunity for revealing resonance (and antiresonance) frequency regions, if they exist. Additionally, three standard methods of formant estimation were applied to 60 units sampled from five humpback whale songs recorded in Hawaii: wide-band spectral analysis, linear prediction spectra, and cepstrally smoothed spectra. This small sample of units was chosen simply to test the ease with which formants can

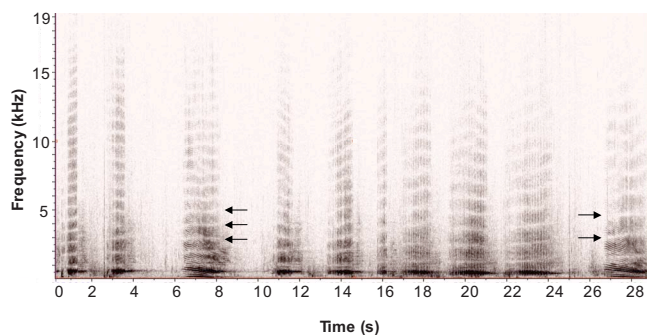


FIG. 3. (Color online) Interference resulting from shallow-water propagation can generate formants in units; the formants appear as dark bands separated by lighter bands (spectral gaps) in this spectrogram. Arrows indicate spectral gaps that are continuously shifting in frequency across units such that two acoustically similar units show different formant positions. The highly regular spacing of gaps across a broad band of frequencies also suggests that most of the formants visible in this spectrogram are environmentally generated. FFT size=1500 Hz.

be measured in units. Units were selected based on recording quality and signal-to-noise ratio, and so that units of different pulse rates were included.

Another technique used to analyze formants in human singers, called long-term average spectrum (LTAS) analysis, involves measuring the spectral content of whole songs. This technique can reveal spectral regions where the greatest energy is being produced, thereby providing not only an indication of prominent formant positions but also information about the spectral bands that singing humpback whales use most extensively. LTAS were created for songs and long segments (>10 min) of song sessions recorded in Hawaii between 1981 and 1995 by averaging the power density spectra (MATLAB's `psd` command,  $FTT=256$ ) of adjacent, nonoverlapping, 625 ms frames across entire recordings. To further explore the origins of patterns observed in LTAS, the spectral content of units was calculated, and spectra from units were plotted in their order of occurrence to create spectral history plots for entire songs. This analysis normalized units in the time domain because spectra were calculated across the entire duration of each unit (i.e., differences in duration across units were converted into differences in spectral energy).

A problem arises in analyzing formants within humpback units that is seldom encountered when analyzing the vocalizations of terrestrial mammals, which is that mechanisms other than internal resonators can generate systematic patterns of spectral degradation and enhancement in song recordings. In the shallow-water environments frequented by singing humpback whales, sounds often reflect from the surface and bottom before reaching a recording sensor, leading to constructive and destructive interference patterns (Mercado and Frazer, 1999; Mercado *et al.* 2007; Thode, 2000). These interference patterns produce formants in recorded units with positions and properties that can be highly similar to those that would be generated by vocal tract filtering. Environmentally generated formants may change gradually over time depending on the source depth, receiver depth, and the distance between the source and receiver (see Fig. 3).

Consequently, any pattern of formants observed in units potentially reflects filtering by both internal structures and the external environment.

To account for this confound, additional spectral measures of units beyond peaks and valleys were made. Specifically, nonlinearities within units called frequency jumps (Fitch *et al.*, 2002; Tyson *et al.*, 2007) were measured to further assess the role of air-filled chambers in humpback whale sound production. In human singers, nonlinear phenomena such as subharmonics and frequency jumps are sometimes observed when the pulse rate is similar to a vocal tract resonance frequency (Titze, 2008b). Titze (2008a) suggested that such instabilities reflect interactions between vocal fold vibrations and the resonance properties of the vocal tract. The standard source-filter model assumes that there are no interactions between the source and the filter (i.e., that they independently contribute to sound generation). However, if air-filled chambers in humpback whales do influence source vibrations, then one would expect that the likelihood of a frequency jump occurring within a unit would be related to the resonance properties of that chamber. In other words, as the pulse rate gets closer to a formant peak, the probability that a frequency jump occurs should increase. Because formant peaks are a function of the dimensions of the air space generating them, the probability of a frequency jump should be as well. Thus, the distribution of frequency jumps within songs can provide indirect evidence of formant positions. Patterns in the probability of frequency jumps should not be impacted by propagation effects, and thus can be combined with more traditional measures of formants to evaluate the extent to which directly measured formant patterns are distorted by environmental factors.

Measures of frequency jumps were collected by visually inspecting and manually measuring spectrograms of units (FFT size=1500) using RAVEN 1.3 to determine whether frequency jumps were more likely to occur near particular frequencies. Measures of frequency jumps were collected from 15 Hawaiian songs and 1 Puerto Rican song (the same one used in analyses of source characteristics). When frequency jumps were identified, the frequency of the lowest visible spectral band before and after the jump was measured (Fig. 4). For units containing multiple frequency jumps, only the three clearest jumps were measured. The number of times jumps occurred within particular frequency ranges was then determined.

## E. Modeling humpback whale sound production

The original source-filter model developed to describe human speech production provides a useful tool for generating and testing specific predictions about physical constraints on humpback whale sound production and for describing recorded sounds in ways that are more closely tied to the processes that whales use to generate sounds. Nevertheless, this simple signal processing approach clearly does not capture all of the factors that determine the acoustic features generated by singing humpback whales. The standard source-filter model does not specify particular sound generating mechanisms, does not account for the possibility that source vibra-

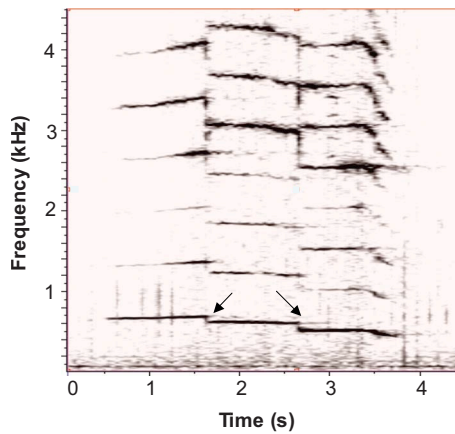


FIG. 4. (Color online) Frequency jumps within a unit appear as discrete vertical shifts in the positions of horizontal bands in spectrograms. Arrows indicate two frequency jumps. FFT size=1500 Hz.

tions might be generated by both the inflow or outflow of air within a closed system, and assumes that the spectral features of produced sounds are largely determined by a single air-filled chamber (the vocal tract) that is weakly coupled to the source. The source-filter model also cannot account for nonlinear phenomena that are commonly produced by singing whales.

Simple geometric models of sound producing mechanisms (which are analogs of physical models) were developed to address some of these limitations. Specifically, patterns of spectral peaks and valleys in unit and LTAS analyses were modeled as the effects of filtering by an air-filled tube. Uniform, closed-end tubes have been used extensively to model the effects of the vocal tract on sound production by humans and other terrestrial mammals (Riede *et al.*, 2008). Such a tube resonates at frequencies that are odd multiples of a fundamental frequency,  $F_o = c/4L$ , where  $L$  is the length of the tube and  $c$  is the sound velocity. The closed-end tube model predicts that antiresonances should occur at even multiples of the fundamental resonating frequency of the tube. In the current study, this basic tube model was modified to account for internal air recirculation by humpback whales, as well as observed patterns in nonlinear frequency jumps.

### III. RESULTS

#### A. Relationships between pulse rate, peak frequency, and duty cycle

Automated unit detection algorithms based on energy levels (see Green *et al.*, 2007) identified 1221 sounds within the 54 min recording of a single whale singing in Puerto Rican waters. Of these, 286 (23%) were subjectively classified as units containing discrete pulses, 733 (60%) were classified as units not containing discrete pulses, and 201 (16%) were not classified as units (e.g., boat noise, ambient noise, and unidentifiable low amplitude sounds). Unit pulse rate varied greatly, ranging from approximately 8 to 6000 Hz within and across units. Figure 5(a) shows the distribution of pulse rates across all 35 734 frames analyzed. The median pulse rates for units containing discrete pulses were 25 Hz (S.D.=17) on average and 230 Hz (S.D.=103) for units not

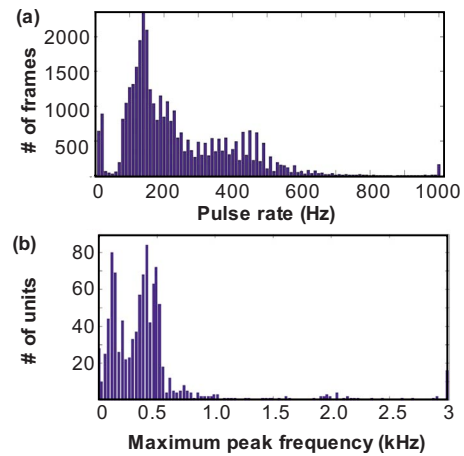


FIG. 5. (Color online) (a) Histogram showing the distribution of pulse rates measured in all unit frames analyzed ( $n=35\,734$ ); (b) Histogram showing the distribution of peak frequencies at the apex of the waveform envelope for all units ( $n=1019$ ).

containing discrete pulses. The acoustic characteristics of some units gradually shifted across repetitions. For example, Fig. 6 illustrates the gradual transformation of high pulse rate, short duration units (often described in the literature as chirps) into low pulse rate, and long duration units (ratchets) within a 2 min song segment.

The precise duration of units was difficult to measure because waveform envelopes typically showed gradual increases and decreases in amplitude at the beginning and end of the unit, respectively, and because ambient noise and water noises occasionally overlapped with the beginning or end of a unit. Automated measures of unit duration ranged from about 200 ms to 4.3 s. The median duration of all units was 1.2 s; the median duration for units containing discrete

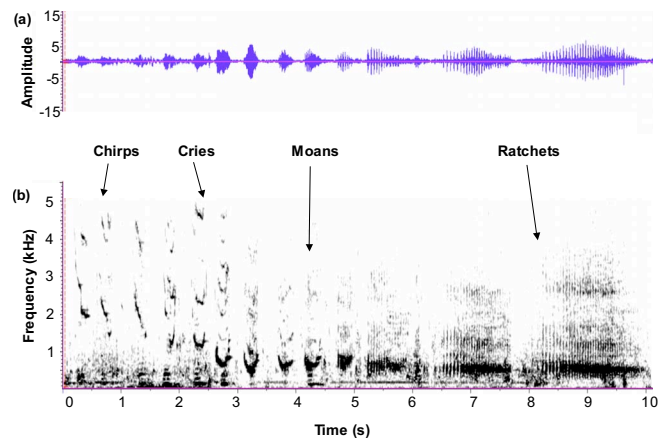


FIG. 6. (Color online) Example of units gradually shifting from being high pulse rate tonal sounds (chirps and cries) to medium rate units (moans), and finally to low pulse rate trains (ratchets). This spectrographic sequence was created by deleting units from a continuous, 2 min segment of song, leaving only a subset of units that were representative of transitions along the pulse rate continuum (i.e., the order of units within the sequence was preserved, but not all units and intervening silences are shown). In both the waveform (top) and spectrogram (bottom), the gradual emergence of discrete pulses is evident. This sequence also illustrates how pulse rate can change while peak frequencies remain stable.

pulses was 1.3 s, and for other units was 1.1 s. The modal unit duration was 260 ms. Unit duration was not correlated with pulse rate ( $r=-0.03$ ).

Peak frequencies ranged from a low of 0–29 Hz (the lowest bin in the spectral analysis) to a high of about 6000 Hz. For units containing discrete pulses, the average peak frequency was 320 Hz (S.D.=220 Hz), whereas for other units, it was 450 Hz (S.D.=480 Hz). The envelope of each unit corresponded to the energy of peak frequencies across consecutive samples. The distribution of peak frequencies at the apex of this envelope (the maximum peak frequency) across units is shown in Fig. 5(b). This apex occurred at 57% (S.D. 24%) of the duration of the unit on average (i.e., near the middle of the unit). Peak frequency at the apex was not correlated with duration ( $r=-0.03$ ) or median pulse rate ( $r=0.09$ ). Peak frequencies could be maintained across or within units despite changes in pulse rate (e.g., see Fig. 6).

Duty cycle measures ranged from a minimum of 0.04 to a maximum of 1.0. The average duty cycle across all frames was 0.52 (S.D.=0.15); the distribution was approximately Gaussian around this mean. For units containing discrete pulses, the average duty cycle was 0.22 (S.D.=0.07), whereas for other units it was 0.52 (S.D.=0.08). Average duty cycle was strongly correlated with median pulse rate ( $r=0.69$ ,  $p<0.001$ ), but not with peak frequency ( $r=0.12$ ) or duration ( $r=0.05$ ).

Collectively, results from source analyses were consistent with the hypothesis that the sound generator in humpback whales involves a pair of vibrating membranes (such as the U-folds) that are dynamically tensed to vary the rate at which they generate pulses. The relationship between pulse rate and waveform shape (quantified in terms of duty cycle) further suggests that techniques for characterizing glottal pulse shape within source-filter models may be appropriate for characterizing pulse features produced by singing humpback whales. Differences in pulse shape between humpback whales and other species may reflect differences in the physical properties of the vibrating membranes in humpbacks (e.g., the parallel orientation of the U-folds) or in how membrane position is controlled. More detailed measures of variability in pulse shape are needed to determine the range of pulses that humpback whales can generate and how pulse shape relates to other acoustic properties of units. Understanding the dynamics of pulse generation can potentially provide new insights into the functional properties of the U-folds.

## B. Spectral markers of resonating structures

Evidence of alternating enhanced and degraded spectral energy regions was found in each of the broadband units examined. Figure 7 shows example spectra for two units modulated across a broad frequency range: a frequency glide and a low pulse rate unit. The frequency glide shows narrow, harmonically related peaks and troughs for frequencies below 2 kHz, with broader peaks centered near 2.5 and 3.25 kHz. The low pulse rate unit shows one narrow peak near 300 Hz, and three or four broad, evenly spaced peaks at higher frequencies. Figure 8 shows the spectra of a very low

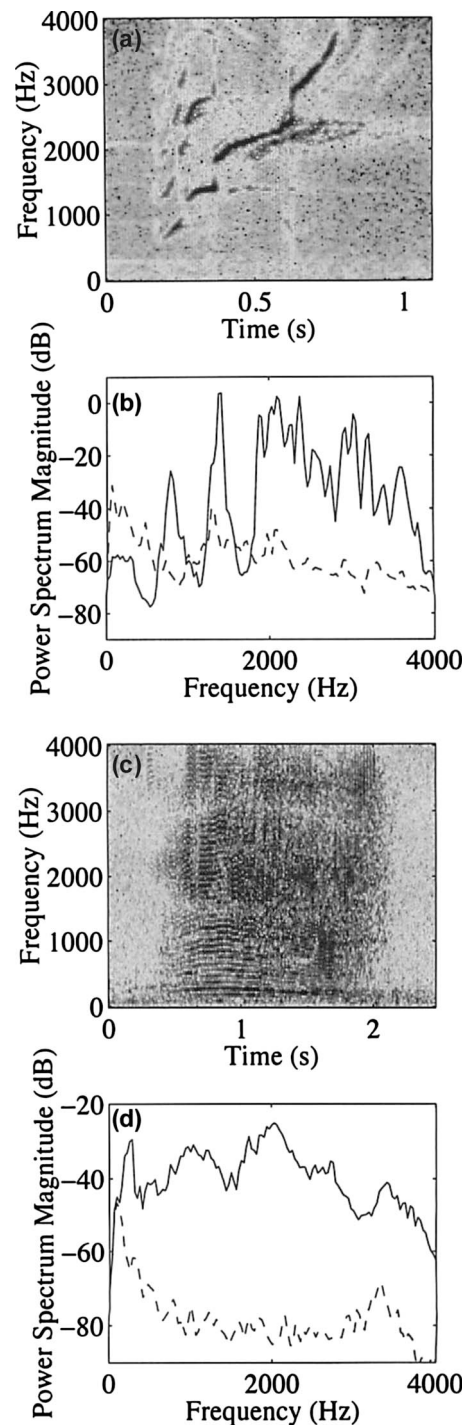


FIG. 7. Example broadband units revealing formants: (a) spectrogram of a frequency glide spanning four octaves; (b) peaks in the power spectral density plot for this unit show formants; (c) spectrogram of a relatively low pulse rate unit shows darker regions indicative of formants; (d) the power spectral density plot shows corresponding formant peaks. Dotted lines show the spectra of ambient noise. FFT size=256 Hz.

pulse rate unit and of an aperiodic unvoiced unit.<sup>1</sup> Both reveal narrow low frequency peaks and broader higher frequency peaks. The number of formants extracted from song units ranged from 3–8. Eight frequency regions where formants were consistently detected across units were identified: 100–500, 600–900, 1000–1500, 1600–1900, 2000–2400, 2500–2900, 3000–3400, and 3500–3800 Hz. These

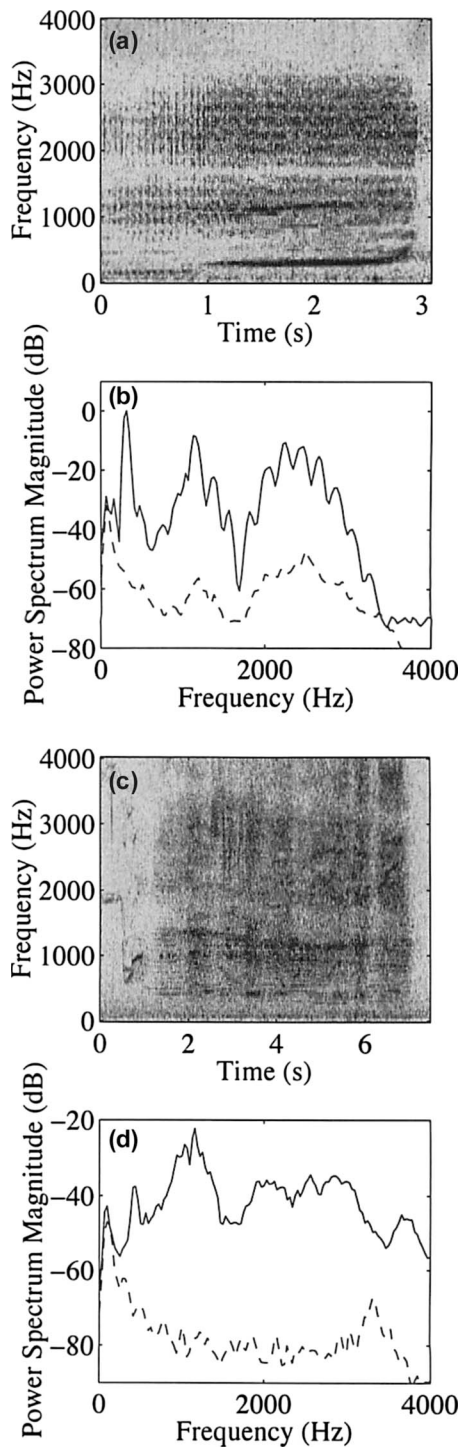


FIG. 8. Example broadband units revealing formants: spectrogram (a) and power spectral density plot (b) of a low pulse rate unit show formants that are narrow at lower frequencies and broader at higher frequencies—note also the regularly spaced peaks within the broad formants above 1 kHz; (c) spectrogram of a noisy, aperiodic unit shows stable darker bands which appear as broad peaks in the power spectral density plot of this unit (d). Dotted lines show the spectra of ambient noise. FFT size=256 Hz.

regions were bounded by narrower spectral regions in which formant peaks seldom occurred (e.g., 500–600 and 900–1000 Hz).

Results from LTAS analyses of Hawaiian songs and spectra of units showed consistent trends in peaks and valleys with peaks occurring between 150–450, 600–900, 1100–

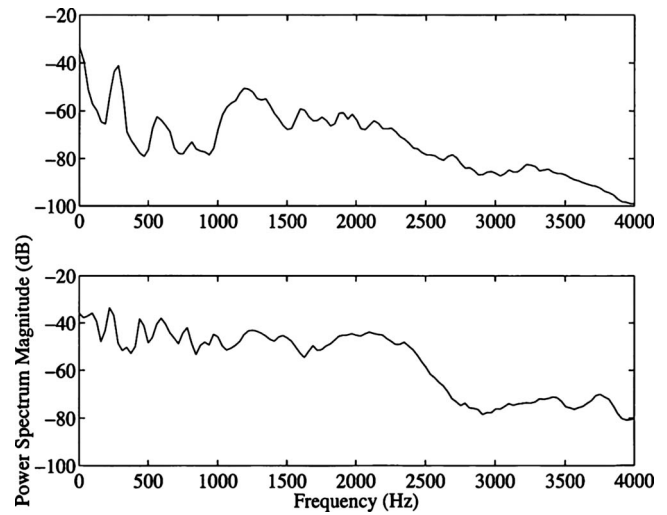


FIG. 9. LTASs for a 41 min segment of song session produced by a single whale, recorded off the coast of Hawaii on 2/16/95 (top), and for a 27 min segment of a song session recorded on 2/19/95 (bottom), show narrow formants at frequencies below 1000 Hz and broader formants above this frequency.

1600, 1700–2200, 2300–2700, and 2800–3100 Hz, and valleys occurring between 500–600, 1000–1100, 1600–1700, and 2100–2200 Hz (Fig. 9). These regions are similar to those identified above from spectral analyses of units.

Spectral history plots of songs revealed that the patterns in LTAS reflected the cumulative spectral content of songs rather than formant patterns that were consistently present across units. For example, in Fig. 10 it is clear that narrow spectral peaks below 500 Hz are occurring in different sec-

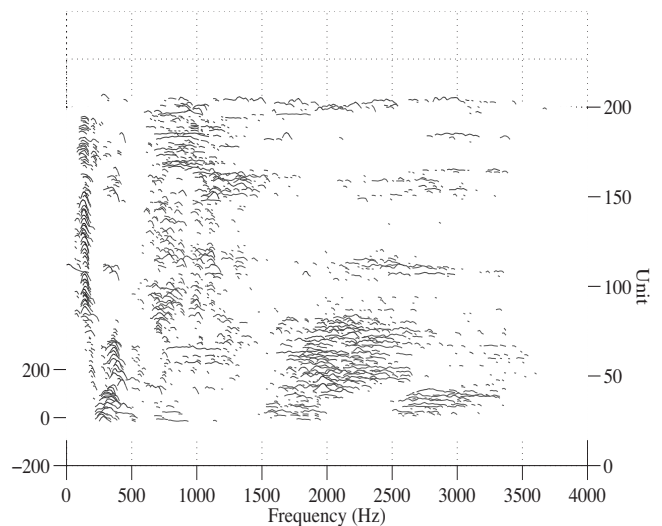


FIG. 10. Three-dimensional spectral history plot for a single humpback whale song recorded off the coast of Hawaii in 1981. Each horizontal trace shows the peaks of the power spectral density plot for a single unit (as in Figs. 7 and 8). The order of traces from the bottom of the plot to the top corresponds to the order of the units (approximately 200) within the song. Viewed as a topographical contour map, the closer peaks (at the bottom of the plot) are from sounds early in the song and the farther peaks (near the top of the plot) are from sounds later in the song. This plot reveals narrow, low frequency peaks and broader, higher frequency peaks that vary progressively throughout the song, but that are relatively stable for long portions of the song. Changes in narrow, low frequency formants are paralleled by changes in broader, higher frequency formants. FFT size=256.

tions of the song, and that broader peaks at higher frequencies are also prevalent at different times during the song. Spectral history plots also revealed that the spectral content of units showed continuity throughout the song. For example, the most prominent lower frequency peak (~200 Hz) evident in Fig. 10 shows a gradual increase in energy that coincides with a decrease in energy at a slightly higher frequency peak (~350 Hz). This lower frequency peak is associated with a broader region of enhanced frequencies (near 1000 Hz) that is evident throughout the song, despite changes in themes. These results are surprising because they suggest that resonant frequencies, if present, are not fixed, and that when singers “switch” to a new theme, the spectral features of units within the new theme are systematically related to those of the preceding theme.

Overall, spectral patterns within frames sampled from units, across whole units, and across songs were consistent with air-filled chambers impacting the features of at least some units. The systematic spectral peaks and valleys observed provide clear evidence that many units have formants. This property of units has not been noted in past studies of humpback whale songs, possibly because most efforts to classify units have focused on features that were clearly audible and easy to trace on spectrograms (e.g., pitch contours). Neither technique is conducive to the identification of formants. Another possibility is that researchers could not consistently distinguish formants produced by humpback whales from those generated by propagation (see also Gaunt, 1983; Watkins, 1967). Propagation-related formant patterns should vary as a function of distance, so that analyses of units recorded at varying distances would tend to reduce such effects, especially in LTAS analyses. Nevertheless, if both the singer and the boat from which recordings were collected were drifting in similar directions, then propagation-generated formants could contribute significantly to formant patterns. Furthermore, environmental formants can themselves be periodic as a function of range (Mercado and Frazer, 1999), so that even if the distance between the singer and hydrophone continuously varied, these formants could potentially overshadow those generated by the whale. In the current analyses of frames, units, and songs, observed spectral peaks varied in bandwidth, with lower frequency peaks being narrower than higher frequency peaks. It is not immediately clear how propagation-induced interference might generate such formant patterns. This trend in bandwidth variability is often observed in vocal tract resonances, however, where it is typically explained as resulting from the clustering of higher frequency formants.

As noted earlier, the closed-end tube model of mammalian sound production predicts that antiresonances should occur at even multiples of the fundamental resonating frequency of the tube. A common multiple within the range of observed spectral valleys in humpback whale units is 500 Hz, corresponding to a tube with a fundamental resonating frequency of approximately 250 Hz (assuming the size of the tract varies between whales, this value would likely vary across individuals). Based on this resonant frequency, the model predicts that a 34 cm long air-filled tube (twice the length of a human vocal tract) or similarly shaped chamber

TABLE I. Frequencies at which frequency jumps were most likely to occur (shown as peaks in the distribution in Fig. 11) and spacing between peaks (all values in Hz). Estimated formant positions were derived by identifying a distribution that was comparable to the measured distribution, and in which the spacing between formants was uniform and equal to the fundamental frequency (i.e., the frequency of the first peak). For formants above 1200 Hz, where regularity was less clear, the estimated formant position is the midpoint between the frequency jumped from and the frequency jumped to, as shown in Fig. 11.

Frequency jumped from	Difference from prior peak	Estimated formant position
163	...	127
263	100	254
363	100	381
488	125	508
613	125	635
738	125	762
888	150	889
1013	125	1016
1188	175	1143
1963	775	2000
2663	700	2800

contributes to humpback whale sound production. Although the tube model is attractive because it is simple and adequately characterizes sound production by many terrestrial mammals, its application to humpback whale sound production is problematic. First, low pulse rate units should show the clearest evidence of formants, yet some of these units do not exhibit them (personal observation, E. Mercado). Second, unlike spoken and sung vowels, the number of formants present in a unit can vary considerably, and not many units show all the formants one might expect to be produced by a resonating tube (e.g., see Figs. 3 and 10). Finally, the distribution of dominant peak frequencies in the one recording from Puerto Rico analyzed here does not show a single peak near 250 Hz, but instead shows two peaks at frequencies surrounding this frequency [Fig. 5(b)]. These factors suggest that the closed-end tube approximation of the vocal tract is insufficient to characterize the spectral variability in units produced by humpback whales, and that if internal air spaces are involved in humpback whale sound production, their role differs somewhat from that seen in most terrestrial mammals.

### C. Possible nonlinear source-filter coupling

A total of 5496 units from 16 recordings were visually examined for evidence of frequency jumps. From these, 827 units (15%) with one or more frequency jumps were identified. The proportion of units within a recording that contained frequency jumps varied considerably across recordings, ranging from 6% to 48%. Overall, the frequencies before and after 997 frequency jumps were measured. These measurements revealed harmonically related frequency regions in which the probability of a frequency jump peaked (Table I). These regions overlapped with those identified in formant analyses, but differed in several respects. In particular, there were a larger number of narrow peaks in the distribution of frequency jumps than was evident in formant analyses (Fig. 11). The spacing between these narrow peaks

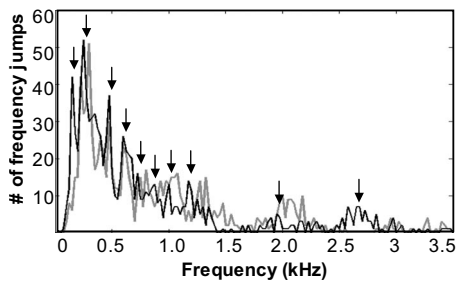


FIG. 11. Histogram showing the distribution of frequencies just before (black line) and after (gray line) a frequency jump occurred within a unit. Arrows show frequencies where jumps were most likely to occur.

was regular across at least six consecutive peaks, making it highly unlikely that the pattern occurred by chance. Spacing between peaks above 1200 Hz was less clear, but also appeared to be nonrandom.

These results further suggest that filtering and resonance of source vibrations in humpback whales differs from what is typically observed in terrestrial mammals. Frequency jumps within sounds do not directly indicate filtration by air-filled resonators. However, recent work with human singers shows that vocal tract resonances can impact the likelihood that such nonlinear phenomena will occur (Titze, 2008b). In the current analysis, the likelihood of a frequency jump within a unit was tightly linked to the frequency content of the unit. The regular spacing of jump-prone regions suggests that modes of source vibration in humpback whales may be more tightly coupled to resonance features of internal structures than is the case in humans. Such coupling is more typical of musical instruments (e.g., brass and woodwinds). If regions where frequency jump probability is high are interpreted as likely formant regions, then the pattern of peak probabilities suggests that the fundamental frequency of the resonator is about 125 Hz rather than 250 Hz (the value indicated above by more traditional formant analyses). Furthermore, the observed pattern of peaks is more consistent with a 1.4 m tube with both ends closed, than it is with a 34 cm closed-end tube. Higher frequency (1200–3500 Hz) jump probabilities showed a different pattern that was more consistent with the closed-end tube model. In fact, the formant frequencies physically measured by Riede *et al.* (2008) from a 20 cm closed-end tube were 380, 1170, 1990, and 2560 Hz. These peaks are startlingly similar to estimated formant positions identified by analyses of unit frequency jumps (381, 1143, 2000, and 2800 Hz; see Table I).

A possible explanation for the discrepancy in formant patterns suggested by the different spectral analyses is that not all units are filtered by the same air-filled cavities. Figure 12 depicts one scenario in which production of different units might engage different resonators. This model of sound production assumes that humpback whales shuttle air bidirectionally while singing, and that they can produce units using air moving in either direction. The model predicts two broad classes of units that have similar source characteristics, but divergent formant features and nonlinear characteristics. Reidenberg and Laitman (2007) described sounds produced by air flowing from the lungs as “egressive vocalizations,” and sounds produced by air flowing into the lungs as “ingres-

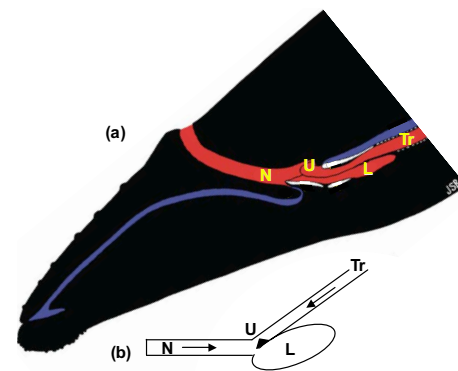


FIG. 12. (Color online) (a) Respiratory tract of a humpback whale showing major components hypothesized to be sound production mechanisms (adapted from Reidenberg and Laitman, 2007); N=nasopharynx, L=laryngeal sac, Tr=trachea; U=U-folds. (b) Egressive units may be produced when air from the lungs induces vibrations in the U-folds, generating airborne vibrations that resonate in the nasopharynx and laryngeal sac. Ingressive units may also be produced when air moves back into the lungs.

sive vocalizations.” In the proposed model, the lungs serve as the pneumatic force for egressive units, the U-folds as the vibratory source, and two chambers act as resonating filters: the laryngeal sac (a closed-end tube) and the nasopharynx (a tube closed at both ends). Air forced from the lungs gradually fills the nasopharynx, creating a resonant air space. Once this space is filled, air begins to inflate the laryngeal sac. This mode of production would account for the gradually increasing and decreasing envelope of units, the evidence suggesting that frequency jumps result from acoustic coupling between the vibrating source and a long air-filled tube, and the evidence of a much shorter closed-end tube generating broad higher frequency formants. During ingressive production, air escapes from the nasopharynx (possibly aided by buoyant forces), and to a lesser extent from the laryngeal sac, back into the lungs. The U-folds would again serve as the vibratory source, but the nasopharynx would no longer resonate because of the air flowing from it. The laryngeal sac might (or might not) continue to resonate during ingressive production. Experiments with doves show that the resonance properties of air sacs in their throats are minimally affected by changes in the volume of the sacs (Riede *et al.*, 2004). Instead, the tissue lining the walls of the dove’s sac appears to determine how it resonates (see also Riede *et al.*, 2008). This geometric model of humpback whale sound production can be instantiated within the source-filter framework by modeling filtration during egressive production as resulting from a uniform, “closed mouth” tube with an inflatable air sac attached as a side branch (Riede *et al.*, 2008) and filtration during ingressive production as resulting from the sac alone. However, more advanced nonlinear models such as those developed to describe sound production by birds (Amador and Mindlin, 2008; Fee *et al.*, 1998) may ultimately prove to be more suitable for modeling sound production by humpback whales.

#### IV. DISCUSSION AND CONCLUSIONS

The analyses of humpback whale songs and units described in this paper have revealed several phenomena rel-

evant to understanding how humpback whales produce sounds. First, many units could be characterized as graded, lying along a continuum with trains of discrete pulses at one end and continuous tonal signals at the other. This finding suggests that traditionally discrete categories of song units such as cries, moans, and ratchets may correspond to points along a continuum from high pulse rates to lower pulse rates (e.g., see Fig. 6), and that humpback whales may gradually modulate specific features of the sounds they produce while singing (e.g., pulse rate) rather than switching between discrete sound types selected from a fixed repertoire. Second, the observed correlation between unit pulse rate and duty cycle suggests that pulse shape may be a function of pulse rate, as is seen in human singers. This suggests that variability in the frequency content of units could reflect variations in the properties of the vibrating membranes that whales use to generate sounds. Third, units contained formants. These formants differed from the ones present in human vocalizations, however, and thus could not resolve the question of whether resonating air-filled chambers contribute to humpback sound production. In contrast, systematic patterns of nonlinear frequency jumps suggested acoustic coupling between a vibrating source and one or more resonant cavities within the respiratory tract, providing more convincing evidence that such cavities do play a role in sound generation.

Past analyses of cetacean vocalizations have often emphasized the differences between sounds produced by the larger baleen whales and those produced by delphinids. The current study suggests, however, that the acoustic structure of humpback whale sounds is similar in many respects to that seen in sounds produced by false killer whales (Murray *et al.*, 1998). Both species show the capacity to continuously modulate their sounds from low rate pulse trains to higher rate tonal sounds. These acoustic similarities indicate that similar analysis techniques should be adequate for characterizing the basic source characteristics of both species (see also Killebrew *et al.*, 2001). The sounds produced by the two species' seem to differ mainly in terms of the shapes and frequency content of individual pulses, and in that false killer whales can simultaneously produce two independent pulse rates (see also Brown, 2008). These differences likely reflect variations in the size and configuration of the membranes that are vibrating (phonic lips in false killer whales versus vocal folds in humpbacks) and the presence of two sets of membranes in false killer whales versus one in humpback whales. Pulse shape appears to be more variable in humpbacks than in false killer whales, suggesting that source membrane vibrations in humpback whales are more complex.

Most units were readily modeled as quasiperiodic pulse trains (see also Mercado and Kuh, 1998), acoustically analogous to the voiced sounds within human speech and song. The range of pulse rates produced by humpback whales was larger than that produced by humans, however. In fact, the range of unit pulse rates within the single recording analyzed covered the full range of registers producible by humans and extended it. Pulse rates were distributed relatively evenly throughout the range of measured rates, consistent with continuous gradations in pulse rate across units. The lowest

pulse rate units identified in this analysis appear to correspond to surface ratchets (Winn *et al.*, 1970). These sounds are analogous to vocal fry in human singers, except that humpbacks use this mode of sound production much more extensively. Interestingly, nonhuman primates also appear to use the vocal fry register more extensively than humans do (Riede and Zuberbuhler, 2003). The overall distribution of unit pulse rates suggests a modal rate between 100 and 200 Hz, which falls within the modal register of human singers. Subjectively, higher pulse rate units have features similar to those of sounds produced by humans singing in the falsetto and whistle registers. These parallels suggest that constraints on vocal fold vibration in singing humans related to pitch range may also apply to the vibrations generated by singing humpback whales.

The graded structure of units within humpback whale songs has important implications for methods designed to sort units into discrete categories or to identify repeating patterns in songs. Consider, for example, the sequence of units shown in Fig. 6. Classification of these units by subjective aural or visual inspection, or by automated sorting of spectrograms, would invariably lead one to conclude that several unit "types" are present (e.g., chirps, cries, and ratchets). The analyses presented here suggest, however, that this sequence corresponds to a gradual decrease in the tension of the singing whale's vibrating membranes over time, analogous to a human singer gradually decreasing the pitch of her notes. In this production-based context, segregating these units into several types is as arbitrary as dividing a chromatic scale into several distinct categories of pitches.<sup>2</sup> It is possible that humpbacks perceive units as falling into discrete categories despite their graded structure, as humans do for vowels. If so, there is currently no way to determine how a whale's categories might map onto the categories that are salient for humans. Consequently, characterizing humpback whale songs in terms of the acoustic properties of their constituent units provides a more objective approach than analyses based on the visual and aural percepts that those sounds induce in human observers, or on similarities between arbitrarily chosen time-frequency representations of units.

Understanding the ways in which humpback whales produce sounds is also relevant to future investigations of humpback whale behavior and physiology. For example, measurements of vibrations at various positions on the head of a humpback whale (see Stimpert *et al.*, 2007) might provide higher quality signals with which to assess variations in pulse shape, which could provide additional information about vocal fold dynamics. Techniques for measuring air movements within the body of a singing whale (e.g., using high frequency active sonar) can potentially be used to examine air recirculation during sound production, which would clarify how physical constraints shape the structure of phrases within songs as well as the contributions of air-filled resonators to sound generation. Knowing how singing whales produce sounds could also provide new insights into the features of their sounds that are most likely to provide information about individual singers, such as their size, age, or identity. For example, more experienced singers might

exhibit more stable control of pulse dynamics, as evidenced by a reduction (or increase) in frequency jumps within their songs.

Finally, understanding how humpback whales sing may clarify why they sing and how their songs function. In the past, extensive variability in individual sounds has been cited as evidence that units are functionally impotent, and that sound patterns (i.e., whole songs, themes, or phrases) are instead more likely to be the functional units (Tyack, 1981). If singing humpback whales are precisely modulating the acoustic features of individual sounds, however, then this would suggest the opposite conclusion, namely, that individual sounds are the core functional units of songs (see, e.g., Frazer and Mercado, 2000). Although the current analyses do not show the extent to which humpback whales control the features of units, they do provide a more coherent model of how singing whales vary the sounds within their songs, which is an important first step toward identifying which of those variations they control.

## ACKNOWLEDGMENTS

This work was supported in part by NSF Grant No. 0718004 and a fellowship from the Center for Advanced Studies in the Behavioral Sciences to E.M., as well as by a grant to L.H. from the Earthwatch Foundation. We thank A. Frankel for his efforts in collecting many of the Hawaiian recordings used in these analyses, as well as the numerous graduate students and volunteers of the Kewalo Basin Marine Mammal Laboratory who assisted with this work. We also thank J. Olsen for his technical assistance with recording equipment, and A. Hollins and P. Banks for their assistance in collecting recordings.

<sup>1</sup>More detailed analyses of “noisy,” unvoiced components of units are warranted, given that these elements may result from nonlinear (e.g., chaotic) source vibrations (Tyson *et al.*, 2007).

<sup>2</sup>It could be argued that differences in time-varying frequency contours justify the categorization of units. For instance, the first units in Fig. 6 are downsweeps, later units are inverted chevrons, the last units are constant frequency, etc. If unit contours were stereotyped and stable over time, then this argument would have force. They are not. The contours of units are known to gradually change over time within songs, across songs, and across years (Payne and Payne, 1985). Even in Fig. 6, it is clear that when a single whale repeats a unit within a short period, frequency contours can change substantially.

Altes, R. A., Evans, W. E., and Johnson, C. S. (1975). “Cetacean echolocation signals and a new model for the human glottal pulse,” *J. Acoust. Soc. Am.* **57**, 1221–1224.

Amador, A., and Mindlin, G. B. (2008). “Beyond harmonic sounds in a simple model for birdsong production,” *Chaos* **18**, 043123.

Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer-Verlag, New York).

Au, W. W. L., Pack, A. A., Lammers, M. O., Herman, L. M., Deakos, M. H., and Andrews, K. (2006). “Acoustic properties of humpback whale songs,” *J. Acoust. Soc. Am.* **120**, 1103–1110.

Beckers, G. J. L., Nelson, B. S., and Suthers, R. A. (2004). “Vocal-tract filtering by lingual articulation in a parrot,” *Curr. Biol.* **14**, 1592–1597.

Boersma, P., and Kovacic, G. (2006). “Spectral characteristics of three styles of Croatian folk singing,” *J. Acoust. Soc. Am.* **119**, 1805–1816.

Brown, J. C. (2008). “Mathematics of pulsed vocalizations with application to killer whale biphonation,” *J. Acoust. Soc. Am.* **123**, 2875–2883.

Charlton, B. D., Reby, D., and McComb, K. (2007). “Female perception of size-related formant shifts in red deer, *Cervus elaphus*,” *Anim. Behav.* **74**, 707–714.

Cleveland, T. F. (1994). “A clearer view of singing voice production: 25

years of progress,” *J. Voice* **8**, 18–23.

Cranford, T. W., and Amundin, M. (2004). “Biosonar pulse production in odontocetes,” in *Echolocation in Bats and Dolphins*, edited by J. A. Thomas, C. F. Moss, and M. Vater, (University of Chicago Press, Chicago), pp. 27–35.

Cranford, T. W., Amundin, M., and Norris, K. S. (1996). “Functional morphology and homology in the odontocete nasal complex: Implications for sound generation,” *J. Morphol.* **228**, 223–285.

Doscher, B. (1994). *The Functional Unity of the Singing Voice* (Scarecrow, Metuchen, NJ).

Fant, G. (1960). *Acoustic Theory of Speech Production: With Calculations Based on X-ray Studies of Russian Articulations* (Mouton and Co., Netherlands).

Fee, M. S., Shraiman, B., Pesaran, B., and Mitra, P. P. (1998). “The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird,” *Nature (London)* **395**, 67–71.

Fitch, W. T. (1997). “Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques,” *J. Acoust. Soc. Am.* **102**, 1213–1222.

Fitch, W. T., and Fritz, J. B. (2006). “Rhesus macaques spontaneously perceive formants in conspecific vocalizations,” *J. Acoust. Soc. Am.* **120**, 2132–2141.

Fitch, W. T., and Hauser, M. D. (1995). “Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on ‘honest’ advertisement,” *Am. J. Primatol.* **37**, 191–219.

Fitch, W. T., Neubauer, J., and Herzel, H. (2002). “Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production,” *Anim. Behav.* **63**, 407–418.

Frazer, L. N., and Mercado, E., III (2000). “A sonar model for humpback whale song,” *IEEE J. Ocean. Eng.* **25**, 160–182.

Gaunt, A. S. (1983). “On sonograms, harmonics, and assumptions,” *Condor* **85**, 259–261.

Green, S. R., Mercado, E., III, Pack, A. A., and Herman, L. M. (2007). “Characterizing patterns within humpback whale (*Megaptera novaeangliae*) songs,” *Aquat. Mamm.* **33**, 202–213.

Gregg, J. W. (1992). “From singing to speech: On glottal configuration,” *J. Sing.* **53**, 47–50.

Helweg, D. A., Cato, D. H., Jenkins, P. F., Garrigue, C., and McCauley, D. (1998). “Geographic variation in South Pacific humpback whale songs,” *Behaviour* **135**, 1–27.

Ishi, C. T., Sakakibara, K.-I., Ishiguro, H., and Hagita, N. (2008). “A method for automatic detection of vocal fry,” *IEEE Trans. Audio, Speech, Lang. Process.* **16**, 47–56.

Keleman, G. (1963). “Comparative anatomy and performance of the vocal organ in vertebrates,” in *Acoustic Behavior of Animals*, edited by R.-G. Busnel, (Elsevier, Amsterdam), pp. 489–521.

Killebrew, D. A., Mercado, E., III, Herman, L. M., and Pack, A. A. (2001). “Sound production of a neonate bottlenose dolphin,” *Aquat. Mamm.* **27**, 34–44.

Mercado, E., III, and Frazer, L. N. (1999). “Environmental constraints on sound transmission by humpback whales,” *J. Acoust. Soc. Am.* **106**, 3004–3016.

Mercado, E., III, Herman, L. M., and Pack, A. A. (2005). “Song copying by humpback whales: Themes and variations,” *Anim. Cogn.* **8**, 93–102.

Mercado, E., III and Kuh, A. (1998). “Classification of humpback whale vocalizations using a self-organizing neural network,” in *Proceedings of the IJCNN '98*, pp. 1584–1589.

Mercado, E., III, Schneider, J. N., Green, S. R., Wang, C., Rubin, R. D., and Banks, P. N. (2007). “Acoustic cues available for ranging by humpback whales,” *J. Acoust. Soc. Am.* **121**, 2499–2502.

Miller, R. (1986). *The Structure of Singing: System and Art in Vocal Technique* (Collier Macmillan, London).

Murray, S. O., Mercado, E., and Roitblat, H. L. (1998). “Characterizing the graded structure of false killer whale (*Pseudorca crassidens*) vocalizations,” *J. Acoust. Soc. Am.* **104**, 1679–1688.

Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1997). “The acoustic features of vowel-like *grunt* calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions,” *J. Acoust. Soc. Am.* **101**, 2951–2963.

Payne, K., and Payne, R. S. (1985). “Large scale change over 19 years in songs of humpback whales in Bermuda,” *Zeit. Tierpsychol.* **68**, 89–114.

Payne, R. S., and McVay, S. (1971). “Songs of humpback whales,” *Science* **173**, 585–597.

- Purves, P. E., and Pilleri, G. E. (1983). *Echolocation of Whales and Dolphins* (Academic, New York).
- Quayle, C. J. (1991). "A dissection of the larynx of a humpback whale calf with a review of its functional morphology," *Mem. Queensl. Mus.* **30**, 351–354.
- Reby, D., and McComb, K. (2003). "Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags," *Anim. Behav.* **65**, 519–530.
- Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., and Clutton-Brock, T. (2005). "Red deer stags use formants as assessment cues during intrasexual agonistic interactions," *Proc. Biol. Sci.* **272**, 941–947.
- Reeb, D., and Best, P. B. (1999). "The laryngeal apparatus of the pygmy right whale, *Caparera marginata* (Gray 1846)," *J. Morphol.* **242**, 67–81.
- Reidenberg, J. S., and Laitman, J. T. (2007). "Discovery of the low frequency sound source in Mysticeti (baleen whales): Anatomical establishment of a vocal fold homolog," *Anat. Rec.* **290**, 745–759.
- Rendall, D. (2003). "Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons," *J. Acoust. Soc. Am.* **113**, 3390–3402.
- Rendall, D., Owren, M. J., and Rodman, P. S. (1998). "The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations," *J. Acoust. Soc. Am.* **103**, 602–614.
- Ridgway, S. H., Carder, D. A., Kamolnick, T., Smith, R. R., Schlundt, C. E., and Elsberry, W. R. (2001). "Hearing and whistling in the deep sea: Depth influences whistle spectra but does not attenuate hearing by white whales (*Delphinapterus leucas*) (Odontoceti, Cetacea)," *J. Exp. Biol.* **204**, 3829–3841.
- Riede, T., Beckers, G. J. L., Blevins, W., and Suthers, R. A. (2004). "Inflation of the esophagus and vocal tract filtering in ring doves," *J. Exp. Biol.* **207**, 4025–4036.
- Riede, T., and Fitch, W. T. (1999). "Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*)," *J. Exp. Biol.* **202**, 2859–2867.
- Riede, T., Tokuda, I. T., Munger, J. B., and Thomson, S. L. (2004). "Mammalian laryngeal air sacs add variability to the vocal tract impedance: Physical and computational modeling," *J. Acoust. Soc. Am.* **124**, 634–647.
- Riede, T., and Zuberbühler, K. (2003). "Pulse register phonation in Diana monkey alarm calls," *J. Acoust. Soc. Am.* **113**, 2919–2926.
- Rossing, T. D., Sundberg, J., and Ternstrom, S. (1987). "Acoustic comparison of soprano solo and choir singing," *J. Acoust. Soc. Am.* **82**, 830–836.
- Sanvito, S., Galimbert, F., and Miller, E. H. (2007). "Vocal signaling of male southern elephant seals is honest but imprecise," *Anim. Behav.* **73**, 287–299.
- Schevill, W. E. (1964). "Underwater sounds of cetaceans," in *Marine Bioacoustics*, edited by W. N. Tavolga, (Pergamon, New York), pp. 307–316.
- Schutte, H. K., Miller, D. G., and Svec, J. G. (1995). "Measurement of formant frequencies and bandwidths in singing," *J. Voice* **9**, 290–296.
- Soltis, J., Leong, K., and Savage, A. (2005). "African elephant vocal communication II: Rumble variation reflects the individual identity and emotional state of callers," *Anim. Behav.* **70**, 589–599.
- Sommers, M. S., Moody, D. B., Prosen, C. A., and Stebbins, W. C. (1992). "Formant frequency discrimination by Japanese macaques (*Macaca fuscata*)," *J. Acoust. Soc. Am.* **91**, 3499–3510.
- Stimpert, A. K., Wiley, D. N., Au, W. W. L., Johnson, M. P., and Arsenault, R. (2007). "'Megapclicks': Acoustic click trains and buzzes produced during night-time foraging of humpback whales (*Megaptera novaeangliae*)," *Biol. Lett.* **3**, 467–470.
- Sukhovskay, L. I., and Yablokov, A. V. (1979). "Morpho-functional characteristics of the larynx in *Balaenopteridae*," in *Investigations on Cetacea*, edited by G. Pilleri (Brain Anatomy, Berne), pp. 205–214.
- Sun, X., and Xu, Y. (2002). "Perceived pitch of synthesized voice with alternate cycles," *J. Voice* **16**, 443–459.
- Sundberg, J. (1977). "The acoustics of the singing voice," *Sci. Am.* **236**(3), 82–91.
- Taylor, A. M., and Reby, D. (2009). "The contribution of source-filter theory to mammal vocal communication research," *J. Zool.* **280**, 221–236.
- Taylor, A. M., Reby, D., and McComb, K. (2008). "Human listeners attend to size information in domestic dog growls," *J. Acoust. Soc. Am.* **123**, 2903–2909.
- Thode, A. (2000). "Source ranging with minimal environmental information using a virtual receiver and waveguide invariant theory," *J. Acoust. Soc. Am.* **108**, 1582–1594.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).
- Titze, I. R. (2008a). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**, 2733–2749.
- Titze, I. R. (2008b). "Nonlinear source-filter coupling in phonation: Vocal exercises," *J. Acoust. Soc. Am.* **123**, 1902–1915.
- Tyack, P. (1981). "Interactions between singing humpback whales and conspecifics nearby," *Behav. Ecol. Sociobiol.* **8**, 105–116.
- Tyson, R. B., Nowacek, D. P., and Miller, P. J. O. (2007). "Nonlinear phenomena in the vocalizations of North Atlantic right whales (*Eubalaena glacialis*) and killer whales (*Orcinus orca*)," *J. Acoust. Soc. Am.* **122**, 1365–1373.
- Vannoni, E., and McElligott, A. G. (2007). "Individual acoustic variation in fallow deer (*Dama dama*) common and harsh groans: A source-filter theory perspective," *Ethology* **113**, 223–234.
- Watkins, W. A. (1967). "The harmonic interval: Fact or artifact in spectral analysis of pulse trains," in *Marine Bioacoustics 2*, edited by W. N. Tavolga, (Pergamon, New York), pp. 15–43.
- Winn, H. E., Perkins, P. J., and Poulter, T. C. (1970). "Sounds of the humpback whale," in *Proceedings of the 7th Annual Conference on Biological Sonar and Diving Mammals*, pp. 39–52.
- Winn, H. E., and Winn, L. K. (1978). "The song of the humpback whale *Megaptera novaeangliae* in the West Indies," *Mar. Biol. (Berlin)* **47**, 97–114.