Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification

Anne Bøttcher, Shane Gero, Kristian Beedholm, Hal Whitehead, and Peter T. Madsen

Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification

Anne Bøttcher,[a] Shane Gero, and Kristian Beedholm
Department of Bioscience, Zoophysiology, Aarhus University, Denmark

Hal Whitehead
Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada

Peter T. Madsen
Aarhus Institute of Advanced Studies, Høegh-Guldbergs Gade 6B, DK-8000 Aarhus C, Denmark

Sperm whales generate multi-pulsed clicks for echolocation and communication with an inter-pulse interval (IPI) determined by the size of their hypertrophied sound producing nose. The IPI has therefore been used to estimate body size and distinguish between individuals, and it has been hypothesized that conspecifics may use IPIs to recognize each other. However, the degree to which IPIs vary within individuals has not explicitly been tested, and therefore the inherent precision of this measure and its applicability for size estimation for researchers and sperm whales alike remain unknown. Here, the variability in IPI from both animal-borne Dtags and far-field recordings from echolocating and communicating sperm whales is quantified. Three different automatic methods (envelope, cepstrum, and cross-correlation) are tested and it is found that the envelope approach results in the least dispersion. Furthermore, it is shown that neither growth, depth, nor recording aspect fully explains the observed variability among clicks recorded from the same individual. It is proposed that dynamics in the soft structures of the nose are affecting IPIs, resulting in a variation of approximately 0.2 ms. Therefore, it is recommended that this variation be considered in IPI studies and that IPIs may have limited functionality as an identity cue among large groups of conspecifics. © 2018 Acoustical Society of America. https://doi.org/10.1121/1.5047657

I. INTRODUCTION

The sperm whale has a hypertrophied nasal complex that serves as sound generator; the largest in the animal kingdom (Norris and Harvey, 1972; Møhl *et al.*, 2003). The nasal complex consists of the dorsal spermaceti organ and the ventral junk complex, which are both filled with spermaceti oil and terminated anteriorly and posteriorly by the distal and frontal air sacs (Clarke, 1978a; Cranford, 1999; Madsen, 2002a). Clicks are produced at the front of the nose by a single pair of phonic lips connected to the spermaceti organ and the right nasal passage (Cranford *et al.*, 1996; Madsen *et al.*, 2003). A small proportion of the energy generated during a click is transmitted directly into the water to form a so-called p0 pulse (Adler-Fenchel, 1980; Møhl, 2001), whereas most of the energy travels back through the spermaceti organ, reflects off the frontal air sac and travels through the ventral junk complex and is projected into the water (Zimmer *et al.*, 2005b) to form the main signal, p1. However, a small proportion of the forward propagating energy in the junk bounces off the distal air sac and is once again reflected back and forth between the air sacs before being emitted. These reflections give rise to a multi-pulsed click, where the time between two pulses (inter-pulse interval, IPI) therefore should reveal the size of the nose and hence body size if the allometry of the nose-body relationship is known (Norris and Harvey, 1972; Møhl *et al.*, 1981; Gordon, 1991).

Such an acoustic measure of the sizes of the clicking sperm whales was recognized early on as a potential tool for whale researchers (Norris and Harvey, 1972; Møhl *et al.*, 1981; Gordon, 1991). Accordingly, studies estimating size distribution (Adler-Fenchel, 1980; Caruso *et al.*, 2015) and growth rates (Miller *et al.*, 2013) of sperm whales have taken advantage of the IPI for these purposes. Moreover, the difference in IPI between individuals has been used to tell individuals apart during exchanges of codas, the primary communication click pattern (Rendell and Whitehead, 2004; Antunes *et al.*, 2011; Schulz *et al.*, 2011; Gero *et al.*, 2016). This has enabled detailed information on the differences and similarities in coda usage between individuals. Similarly, it may be hypothesized that the multi-pulsed structure is used by clicking sperm whales to radiate and decode size and therefore perhaps identity (Cranford, 1999; Gero *et al.*, 2016). Based on a pronounced sexual dimorphism in both absolute and relative nose size between male and female sperm whales, Cranford (1999) proposed that evolutionary pressures have selected for a large IPI, and therefore nose, in an acoustically mediated mating scheme. In contrast, Møhl *et al.* (2003) have advocated that the IPI is a mere byproduct of a large nose that evolved to provide high directionality and high sound pressure of the produced echolocation clicks.

[a] Electronic mail: annebottcher@hotmail.com

However, this does not exclude that individuals may use the information on size imbedded in the IPI to distinguish between individuals, or classes of individuals.

To decode size acoustically from sperm whale multi-pulsed clicks, it is critical for both listening sperm whales and sound recording researchers alike to know both the relationship between body size and IPI and the accuracy of the IPI estimates. Several studies have convincingly shown that there is a close relationship between body size and IPI if enough clicks are used, giving rise to a series of increasingly fine-tuned equations to estimate sperm whale size from their clicks (Møhl *et al.*, 1981; Gordon, 1991; Rhinelander and Dawson, 2004; Growcott *et al.*, 2011).

However, there is variation in IPI among series of clicks recorded from a single sperm whale (Gordon, 1991; Growcott *et al.*, 2011; Schulz *et al.*, 2011) which suggests limitations on the functional utility of the acoustic size estimation and individual identification for both researchers and among sperm whales themselves. Furthermore, the aspect to the clicking whale has been shown to dramatically change the received waveform (Zimmer *et al.*, 2005a; Teloni *et al.*, 2007; Schulz *et al.*, 2009), and thus the resulting variation in IPI estimate; but even when the recording aspect is stable, IPI estimations may show substantial variations for the same animal (Growcott *et al.*, 2011; Schulz *et al.*, 2011; Mathias *et al.*, 2009) that are currently not understood.

Such uncertainty calls for a study that tests the precision of IPI estimates and identifies the sources of potential variation in IPI estimation. Several factors may result in variation in IPI estimates of the same click, including recording aspect (Zimmer *et al.*, 2005a; Teloni *et al.*, 2007), recording system, signal-to-noise ratio, different analytical method used to estimate IPI, and finally actual biological variation in IPIs of the produced clicks.

To quantify the influence of these potential sources of IPI variation we use recordings of nine sperm whales from a well-studied community in the eastern Caribbean (Gero *et al.*, 2014). Specifically, we test for differences caused by four potential sources of variation: (1) Three different analytical methods; envelope, cepstrum (Antunes *et al.*, 2010), and cross-correlation (Goold, 1996); (2) two recording systems; we use a combination of commonly used recording technologies: array and SoundTrap (Ocean Instruments, NZ) recordings and animal-borne, sound recording Dtags (Johnson and Tyack, 2003); (3) time; we use recordings of the same whale on the same day and over longer time periods; and (4) the effect of depth. To do so, we calculated estimates of IPI for both echolocation and coda clicks. We use these findings to discuss the potential for using IPIs as reliable size estimation and individual identification of sperm whales. Furthermore, we discuss whether IPIs are sound production byproducts or selected for in acoustically mediated communication.

II. MATERIALS AND METHODS

A. Field methods

Data were collected in an area of approximately $2000 \, \text{km}^2$ along the leeward, western coast of the island nation of Dominica. Sperm whale social units were followed by visual and acoustic tracking, and recordings were made using either a towed far field hydrophone or an animal-borne Dtag between January and May in 2005, 2008, 2010, 2014, 2015, and 2016. Photographs of the flukes of individuals that were either tagged or recorded were taken when the animals were diving, allowing for photo-identification using distinct markings on the trailing edge of the fluke. See methods about click detection for details on how to ensure that the photo-identified animal is the one recorded.

Dtags were deployed in three different years from 2014 to 2016. The Dtags were attached on the back of the sperm whales with suction cups using a 9-m-long pole from an 11-m rigid-hulled inflatable boat (RHIB). The Dtags provided stereo sound recordings (120 kHz sampling rate, 16 bit) with a flat frequency response (± 2 dB) from 0.5–50 kHz. Additionally, the Dtag collected ambient pressure measurements sampled at 50 Hz, 16 bit.

In 2005, 2008, 2010, and 2015 a custom-built hydrophone [Benthos AQ-4 elements with a flat frequency response (± 2 dB) between 0.1 and 30 kHz] was towed from a 12-m sailboat. The total length of the towing cable was 100 m with the depth of the hydrophone varying depending on the speed of the boat. Recordings were either made with a sampling rate of 44.1, 48, or 96 kHz, 16 bit, using an amplifier and filter box with high pass filters up to 1 kHz. In 2016 a single-hydrophone soundtrap [SoundTrap 300 with a flat frequency response (± 3 dB) between 20 Hz and 60 kHz, 16 bit] with a sampling rate of 96 kHz was deployed from an 11 m RHIB. For both acquisition techniques, recordings of the onset of echolocation clicks were made within the calmed water called the flukeprint of the photo-identified diving animal, so that the hydrophone was approximately on the body axis of the diving whale and 180° off the acoustic axis.

B. Click detection

1. Dtag recordings

Detection of echolocation clicks was conducted with a custom written click detector written in MATLAB R2015b (The Mathworks, Inc., MA, USA). The click detector automatically detected all clicks with a peak amplitude above 0.03 of the normalized clip level corresponding to a received level above 154 dB re 1 μPa (peak). Afterwards all clicks were manually assessed to ensure that clicks from other whales or noise were not marked as well as to minimize the risk that no focal clicks were missed. Focal clicks are defined as the clicks emitted by the tagged individual and were recognized by their constant high amplitude. If another animal was clicking in close proximity, making it difficult to distinguish which clicks belonged to the tagged animal, no clicks were marked in that period.

Focal coda clicks were marked using a custom written MATLAB script (The Mathworks, Inc., MA, USA) and a LABVIEW program (National Instruments, TX, USA). In Dtag recordings, to ensure that only focal coda clicks were marked, the angle from which echolocation and coda clicks arrived at the Dtag (angle of arrival, AoA) was used. The

AoA of each click was calculated as $90 - \arccos(Tc/D)$, where T is the maximum time delay between the stereo hydrophones, D is the distance between the hydrophones, and c the speed of sound in water (approximated as 1500 m/s). This AoA approximation relies on an assumption of the distance between the hydrophones and the sound producing organ being much greater than the distance between the two hydrophones. An AoA of 0° indicates that the sound is coming from directly in front or behind the Dtag. Coda clicks were only accepted as focal if the median AoA of the coda was within ± three standard deviations of the mean AoA of 250 echolocation clicks in immediate proximity of the coda clicks. This ensured that only clicks emitted from the correct direction relative to the Dtag were marked, greatly reducing the risk of marking non-focal codas. In cases where the AoA of the echolocation clicks changed with more than one standard deviation from the end of one dive to the beginning of the next (due to a change in Dtag position) codas in between were excluded. Furthermore, codas with a median amplitude below 0.05 of the normalized clip level, corresponding to a received level of 158 dB re 1 μPa (peak), were not considered. As 90% of all accepted focal echolocation clicks were above 161 dB re 1 μPa (peak), the probability that codas falling below 158 dB re 1 μPa (peak) were non-focal was considered too high. A conservative approach was taken, meaning that coda clicks were not marked if they were overlapped in time by other sounds, including neighboring whales' codas.

2. Towed hydrophone and soundtrap recordings

Recordings were only used when the animal of interest was in a singleton cluster or in a few cases where a small, dependent calf was present. The calf, however, was unlikely to echolocate and demonstrate an inter-pulse interval (IPI) within the area of interest (2–5 ms, approximate IPI range for adult female sperm whales (Rendell and Whitehead, 2004; Schulz et al., 2011). Any clicks produced by the accompanying calf would therefore be excluded. All accepted recordings were audited, and all recordings with a clear onset of echolocation clicks were selected. The onset of echolocation clicks is recognized by the large inter-click intervals (referred to as "first clicks") before slowly decreasing to normal echolocation rates of approximately two clicks per second (Whitehead and Weilgart, 1990). The selected recordings used for further analysis lasted for up to three minutes after onset of echolocation clicks.

A custom written MATLAB R2015b script (The Mathworks, Inc., MA, USA) and LABVIEW program (National Instruments, TX, USA) enabled detection of clicks. The restriction of only using recordings of a singleton cluster and with first clicks present ensured that only clicks from the photo-identified whale were marked.

C. Estimation of the inter-pulse interval

A time-window from 15 ms before to 15 ms after the marking of the peak of each click was extracted and each click was filtered with a high pass filter at 3 kHz using a second order Butterworth filter to avoid low-frequency noise. Subsequently, all clicks had to meet two criteria to be included for further analysis: (1) A signal-to-noise ratio of more than 20 dB, defined as the ratio between the root-mean-square (rms) level of the click (0.2 ms before the peak to 0.8 ms after the peak) and the rms level of the noise in a 3 ms window starting 5 ms before the peak of the click. (2) An amplitude below 1 dB from the full amplitude range of the recording to avoid effects of clipping. Furthermore, in order to exclude buzz clicks, all echolocation clicks had to meet a third criteria: (3) recorded in a sequence where inter-click-intervals were greater than 0.2 s (Teloni et al., 2007).

Previous studies have suggested several methods for estimating IPI. Here, three of the most widely used methods (envelope, cepstrum, and cross correlation) were chosen for evaluation (Goold, 1996; Teloni et al., 2007; Schulz et al., 2008; Schulz et al., 2009; Schulz et al., 2011; Antunes et al., 2010; Gero et al., 2016). Only automatic techniques were used in this study, as manual assessment is very time consuming and adds the risk of introducing observer bias. Furthermore, for automated techniques, clear criteria can be stated to render the study reproducible. The IPI was only calculated from p0 to p1 for all methods, due to the high decay rate of echolocation clicks (Madsen et al., 2002a). As the true IPI is unknown, it is not possible to say anything about the degree of accuracy for the different techniques. Thus, the only way of determining the relative utility of each technique, is by looking at the precision of the methods by assessing their dispersion. Hence, the method with the least dispersion for most whales was considered the best method, as this allows for a better estimation of a usable IPI when fewer clicks are available. As a way of quantifying the precision of the three methods, the interquartile range (IQR) was calculated, which is an outlier resistant quantifier of the dispersion. IQR values were calculated for each whale for each of the three IPI methods for both echolocation and coda clicks. The method with the least dispersion for most whales within either echolocation or coda clicks was considered the best method when calculating IPI for echolocation and coda clicks, respectively.

1. Method 1: Envelope

IPIs were measured from peaks within the envelope of the wave function, which eliminates potential problems arising from phase differences between the pulses. The envelope was calculated as the absolute value of the analytical signal (called Hilbert transformation in MATLAB). The IPI was then taken to be the delay from the largest positive peak of the envelope in the immediate proximity of where the click was detected (p0) (±3 ms from the marking) to the largest positive peak 2–5 ms after p0 (p1). The lower boundary of 2 ms was chosen as all recorded individuals were fluking and thereby considered adults, as well as to avoid detection of another peak in p0. The upper limit of 5 ms was chosen as females in the Caribbean are rather small with a total length below 12 m, corresponding to an IPI of maximum 4.95 ms according to Gordon (1991). Furthermore, previous analysis on clicks from adults in the Caribbean have never revealed IPIs larger than 4 ms (Schulz et al., 2008; Schulz et al., 2011).

J. Acoust. Soc. Am. **144** (1), July 2018

Bøttcher et al.   367

2. Method 2: Cepstrum

Cepstrum analysis is a standard signal analysis technique picking up replicas within a signal and has been used for IPI analysis in different forms (Goold, 1996; Teloni et al., 2007; Antunes, et al., 2010). We used the complex cepstrum (following Antunes et al., 2010) which Fourier transforms the signal, takes the logarithm to the amplitude spectrum, reapplies the phase information and inverse Fourier transforms it. As was the case with the envelope method the IPI was calculated as the time from the largest positive peak around the marking (± 3 ms) to the largest positive peak 2–5 ms after p0.

3. Method 3: Cross-correlation

As a third approach, we used cross-correlation (Goold, 1996; Schulz et al., 2009; Schulz et al., 2011; Antunes et al., 2010) of the first pulse of each click with the rest of the click. This calculated the correlation between a template waveform (0.4 ms before the peak of p0 to 1 ms after) and the rest of the waveform in the click at different points in time. The time to the highest correlation coefficient 2–5 ms after the first peak corresponds to the time difference between p0 and p1 and therefore the IPI.

D. Analysis of IPI differences

1. Comparing IPIs between recording techniques

The effect of recording technique on the IPI was examined in two different ways. First, IPIs derived from a range of towed (2005, 2008, 2010, 2015) and Dtag recordings (2014, 2015) were compared. As towed recordings were no more than three minutes long, only the first three minutes of echolocation clicks marked on Dtag recordings were included. Second, two individuals in singleton clusters in 2016 were recorded with a Soundtrap and a Dtag in synchrony. Using ICIs between the echolocation clicks from the synchronous recordings to accurately synchronize the clicks, the exact same clicks were compared in the IPI analysis. Only clicks that met all three criteria (signal-to-noise ratio, no clipping, and no buzzes) in both recording techniques were kept for IPI comparison.

2. Comparing IPIs derived from codas and echolocation clicks

To examine whether there is a difference in IPI between coda clicks and echolocation clicks, IPIs from all focal coda and echolocation clicks from the same Dtag recordings from 2014 and 2015 were compared for each whale. In addition, the IPIs of both click types from Dtag recordings were compared with the depth at which they were emitted as recorded by the pressure sensor of the Dtag.

III. RESULTS AND DISCUSSION

To examine the stability of the IPI of sperm whale clicks as well as the potential influence of different recording techniques, recordings by Dtag, towed hydrophones, and Soundtraps were collected. Both Dtag and towed hydrophone recordings were available from nine photo-identified sperm whales (Table I). The age of these individuals is unknown, but they were all fluking at the time of recording and thus classified as adults. After clicks that did not meet the criteria (signal-to-noise ratio >20 dB, no clipping and no buzzes) were excluded, 6710 echolocation clicks from towed hydrophone recordings (from 57 recordings) and 18 733 echolocation clicks from Dtag recordings (from 14 Dtag deployments) were available for analysis. In addition, 460 echolocation clicks recorded on the Soundtrap were accepted from whales #4 and #5. These were all recorded in synchrony with Dtag recordings. In addition, 217 codas recorded on 11 of the Dtag deployments were accepted. The 217 codas consisted of 963 clicks in total.

A. Which method renders the smallest variation in IPI estimation?

To test which method is best for estimating IPIs from echolocation clicks, three commonly used techniques were applied (Fig. 1). The envelope method had the lowest interquartile range (IQR) for most whales as well as the lowest average IQR (0.23). The cepstrum and cross-correlation methods had the highest IQR for 10 and 9 whales, respectively, with an average IQR of 0.33 and 0.32 (Table II). Therefore, we conclude that the envelope method is the most precise method for automatic estimation of IPIs from sperm whale echolocation clicks, regardless of recording technique.

TABLE I. Number of accepted clicks and recordings for towed, Dtag, and synchronous recordings, respectively. Number of Dtags for each whale is given as number of Dtags including echolocation clicks and, in brackets, number of Dtags including codas.

| Whale	Towed recordings		Tag recordings					Synchronous recordings	
	# dives/ recordings	# clicks	# tags (with codas)	# dives	# echo. clicks	# codas	# coda clicks	# dives	# clicks
#1	4	221	1(1)	7	1341	26	134	—	—
#2	6	558	2 (1)	18	2385	18	64	—	—
#3	6	1022	1 (1)	6	632	12	75	—	—
#4	9	641	1 (0)	7	2021	0	0	5	414
#5	1	45	1 (1)	3	571	22	107	2	46
#6	13	2543	1 (1)	7	2070	28	138	—	—
#7	10	1021	2 (2)	13	4385	15	74	—	—
#8	2	46	2 (1)	13	2366	11	7	—	—
#9	6	613	3 (3)	24	2962	85	364	—	—
Total	57	6710	14 (11)	98	18733	217	963	7	460
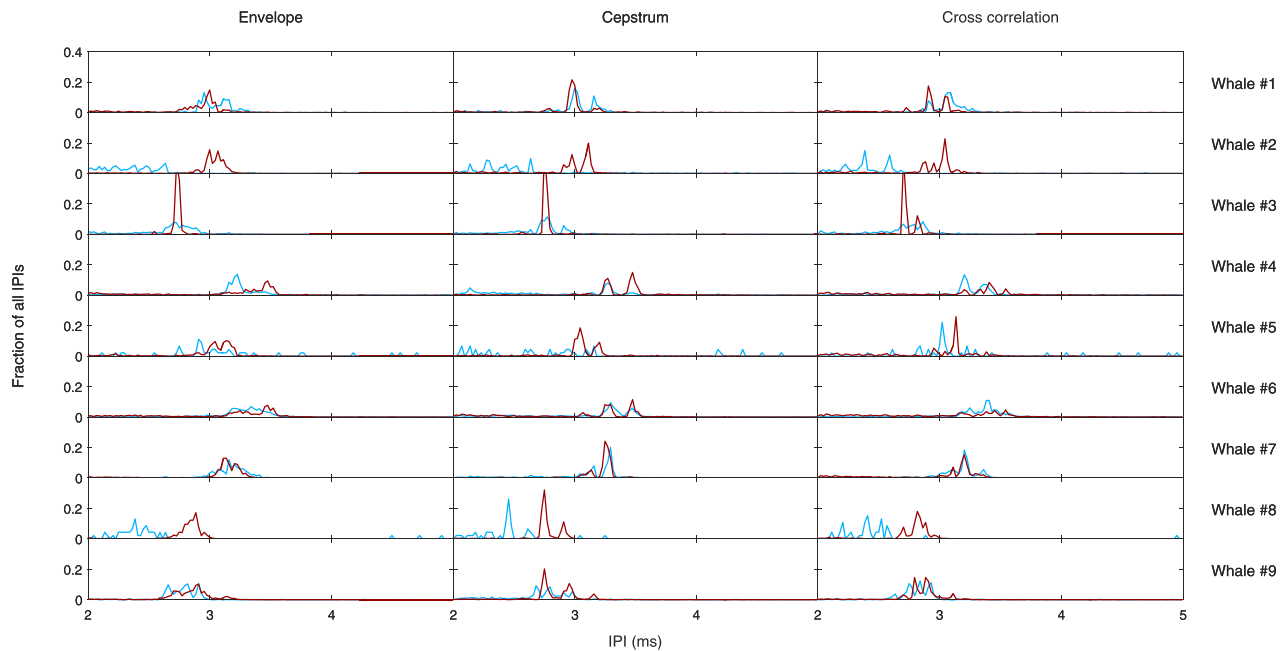
FIG. 1. (Color online) Echolocation click IPI dispersion for each individual using the three different IPI methods (envelope, cepstrum, and cross-correlation) for both towed hydrophone (blue/light) and Dtag (red/dark) recordings. The y-axis is normalized for each of the recording types.

A previous evaluation of methods to estimate IPI suggested that the cepstrum is more suitable (Antunes *et al.*, 2010). The Antunes study used waveform averaging across all clicks before calculating the IPI, as well as different quality criteria for including clicks, and both of these differences may explain the conflicting results.

Once a method has been found that renders the smallest dispersion, it raises the question of how to parameterize the IPI

TABLE II. Echolocation click IPI estimates in ms for the nine whales for each IPI method (envelope, cepstrum, and cross-correlation) as well as the measure of central tendency (median and mode). The median of the envelope estimates is in bold, as this method renders the smallest IQR.

		Envelope			Cepstrum			Cross-correlation		
	Whale	Median	Mode	IQR	Median	Mode	IQR	Median	Mode	IQR
Towed	#1	**3.04**	2.95	0.17	3.00	3.00	0.19	3.06	3.08	0.15
	#2	**2.35**	2.63	0.32	2.43	2.63	0.35	2.38	2.39	0.32
	#3	**2.72**	2.70	0.22	2.75	2.77	0.23	2.74	2.86	0.27
	#4	**3.23**	3.22	0.15	2.98	3.27	0.98	3.23	3.20	0.17
	#5	**3.00**	2.91	0.31	2.79	2.45	0.83	3.02	3.02	0.29
	#6	**3.29**	3.34	0.23	3.29	3.29	0.59	3.38	3.38	0.23
	#7	**3.19**	3.16	0.17	3.23	3.29	0.26	3.21	3.20	0.15
	#8	**2.40**	2.39	0.19	2.46	2.45	0.19	2.42	2.41	0.25
	#9	**2.77**	2.91	0.16	2.75	2.68	0.28	2.82	2.84	0.17
Tag	#1	**2.96**	3.00	0.21	2.98	2.98	0.06	2.93	2.91	0.32
	#2	**3.03**	3.00	0.10	3.00	3.11	0.18	3.02	3.04	0.18
	#3	**2.73**	2.73	0.03	2.76	2.75	0.02	2.71	2.70	0.03
	#4	**3.34**	3.47	0.64	3.44	3.47	0.23	3.30	3.41	0.91
	#5	**3.07**	3.13	0.15	3.04	3.04	0.16	3.10	3.13	0.57
	#6	**3.28**	3.47	0.68	3.27	3.47	0.84	3.16	3.54	0.98
	#7	**3.14**	3.13	0.12	3.26	3.25	0.13	3.18	3.20	0.48
	#8	**2.85**	2.88	0.09	2.76	2.75	0.13	2.82	2.82	0.12
	#9	**2.86**	2.91	0.19	2.79	2.75	0.21	2.88	2.88	0.13
Average				0.23			0.33			0.32

distribution of a given recording. A Jarque-Bera hypothesis test of normality (Jarque and Bera, 1980) showed that the distributions of IPIs for all three IPI methods and both recording types were not normally distributed (p < 0.02 for all whales). Hence, mean may not be a good measure of central tendency.

The mode is generally used for categorical data sets; however, it has also been used in previous IPI analysis (Antunes *et al.*, 2010; Schulz *et al.*, 2011; Gero *et al.*, 2016). These previous studies used a bin size to determine the modal estimate in milliseconds by rounding off to two decimal places corresponding to a resolution of 10 μs. The researchers then defined a good IPI estimate as one in which greater than 50% of clicks in a sequence were within 0.05 ms of the modal estimate and excluded estimates from click sequences in which this condition was not met. Yet, the minimum sampling rate in this study is 44.1 kHz and therefore does not support a resolution of 0.01 ms. From the perspective of the mode, a change in sampling rate could thus in fact lead to the determination of a different IPI without it being different. For this reason, this study has used a bin size of 22.7 μs (1 ms/44.1 samples per ms). While the mean is rendered problematic by the lack of normality of the distributions and mode has problems with different bin sizes, the median is the one measure out of the three that is the least susceptible to outliers in potential small data sets. Therefore, the median will be used as the measure of central tendency in the remainder of this paper. The modal estimate is on average 0.04 ms larger than the median and is included here for comparison with previous papers (Antunes *et al.*, 2010; Schulz *et al.*, 2011; Gero *et al.*, 2016).

B. Are IPIs for the same clicks consistent between recordings on the whales and with far-field recordings?

An animal-borne sound recording Dtag provides a fixed recording aspect for any click series from the focal whale. In

J. Acoust. Soc. Am. **144** (1), July 2018

Bøttcher *et al.* 369

contrast, far-field recordings may impose variation in the IPI due to changing recording aspects. Following from this, we show a variation in IPI between the towed hydrophone and Dtag recordings (Fig. 1) with the absolute difference between Dtag and towed IPI estimates being up to 0.68 ms when using IPIs calculated as medians of the envelope method (whale #2). However, several confounding factors may further influence these results. The most obvious potential problem is the time difference of several years between recording types for some of the whales. If the recorded individual was not fully grown at the time of the first recording (see below), the IPI may have changed between the time when the towed hydrophone recordings were made and the Dtags were deployed (average time interval from first to last recording: 6 years, range: 5–10 years).

Hence, to test if the degree to which IPI estimates of clicks recorded in the far field are explained by source variations in the IPI or variation imposed by changing recording aspects, synchronous recordings are needed. We therefore performed recordings with a Dtag and a Soundtrap in synchrony in May 2016 (Table III, Fig. 2) on two sperm whales (#4 and #5). Only clicks where the criteria were met for both recording techniques were included, ensuring that the exact same clicks were compared. The median IPIs were calculated using the envelope method. For whale #5 where only 46 clicks from 2 dives were available, the Dtag median IPI estimate is 0.92 ms (43%) larger than the Soundtrap median IPI estimate. For whale #4, providing 414 clicks from 5 dives, the difference in median IPI estimates for the two recordings was 0.08 ms (modal difference of 0.23 ms). Thus, more clicks drive down the uncertainty on the IPI recorded in the far field and suggests that automated pulse recognition is unusable on small data sets. Either way these limited data suggest that clicks recorded with tags on a sperm whale render a different IPI distribution and median than if the same clicks are recorded in the far field. Because the tags were placed well behind the sound producing nose, the recording aspect should be close to that of a hydrophone lowered in the fluke print of a diving whale. Furthermore, even if the tag was close enough to the nose to be more off the body axis, it should result in shorter IPIs compared to those in the far field and not longer, as seen here. The implication is that Dtag recordings from a tagged whale cannot necessarily be coupled to far-field recordings via the same IPI if a lot of animals are present with similar IPIs as can easily be the case for large groups of adult females.

TABLE III. Median IPIs calculated using the envelope method for synchronous recordings of echolocation clicks from two whales. 414 and 46 clicks were recorded in synchrony for whale #4 and #5, respectively.

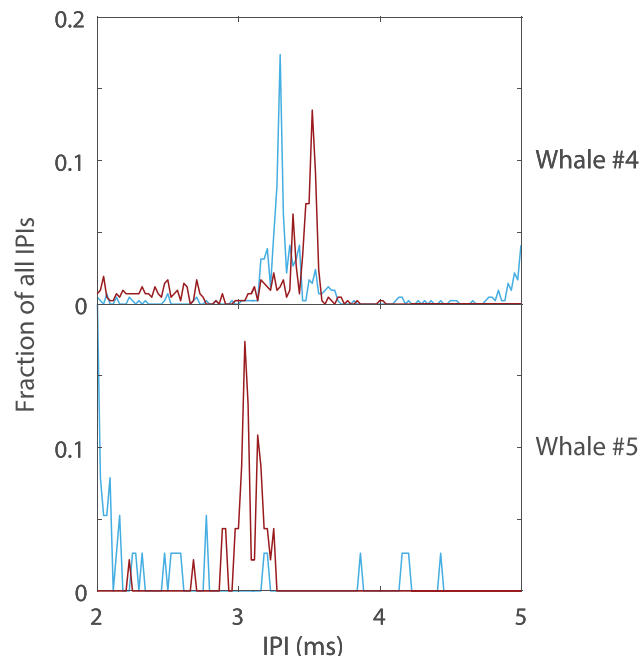|		Envelope	
		Median	IQR
Whale #4	Soundtrap	3.31	0.27
	Tag	3.39	0.81
Whale #5	Soundtrap	2.14	0.57
	Tag	3.06	0.13

FIG. 2. (Color online) Distributions of echolocation click IPI estimates from envelope method for synchronous recordings by a Soundtrap (blue/light) and Dtag (red/dark).

C. IPI variation in far-field recordings of the same whales

The ability to reliably assign clicks to an individual of a certain size requires that the IPI estimates are stable over time intervals where growth can be ignored. To test this, we therefore compared IPI estimates from clicks from the same individual recorded with towed hydrophones on consecutive days. This was done for whale #1, where all towed hydrophone recordings were made within 6 days of each other and included between 24 and 104 clicks each (mean: 56.5 clicks/recording). Surprisingly, the results show variation in median IPI between recordings ranging from 2.96 to 3.10 ms (Fig. 3) despite that all recordings were made in the fluke print of the diving whale, reducing aspect related effects. The recorded whale did not change size over the six days of recording, and yet the IPIs suggest a variation in length of 0.2 m (calculated using equation from Gordon, 1991). Variations in IPIs in the same order have also been seen over short time spans in other studies (Gordon, 1991; Growcott et al., 2011; Schulz et al., 2011). It is implied that even over short time intervals, it is not possible to reliably tell similarly sized sperm whales apart unless they have IPIs that differ in the order of 0.2 ms, corresponding to a size difference of 0.3 m. Whether that relates to small differences in aspect changes of descending whales on different days or source modulation cannot be determined.

D. IPI changes over longer time periods as proxies for growth

Given the substantial short-term variations in IPI, we next sought to address the question of whether IPI over longer time periods allow for estimations on growth for immature animals and reliable size classification of physically
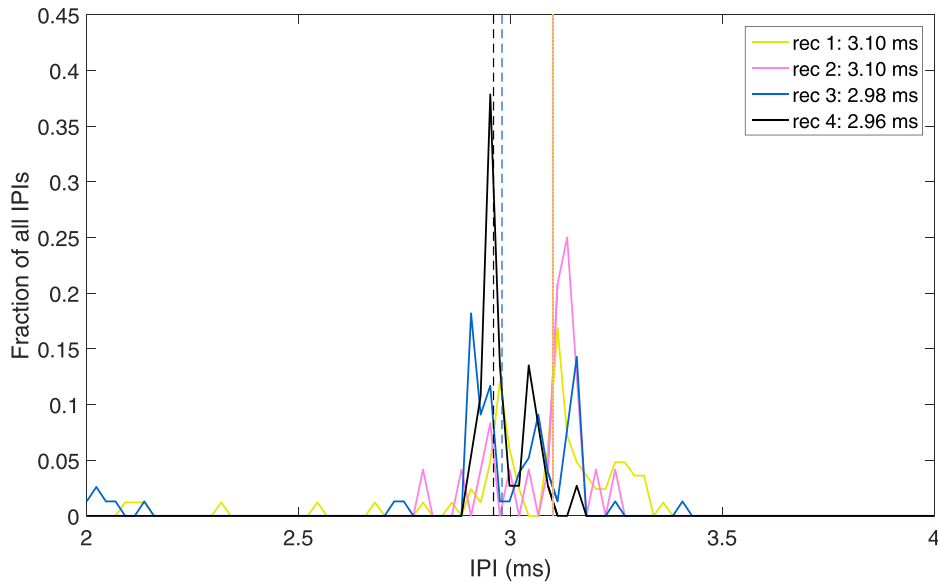
Bøttcher *et al.*

FIG. 3. (Color online) Distribution of IPI estimates from echolocation clicks using the envelope method for the four towed hydrophone recordings of whale #1 plotted as individual recordings. The median for each recording is provided in the legend and indicated by dashed vertical lines.

mature animals. By quantifying IPIs from clicks recorded with towed hydrophones from different years, we could provide measures of variability of IPIs over long time spans. Towed array recordings across years existed from six whales. Between 6 and 13 recordings with between 4 and 368 clicks in each recording (mean: 128 clicks) were available from each whale. The IPI for most of the whales was relatively stable across years, but for especially the whale with the lowest IPI (whale #2), the IPI increases with time (Fig. 4), rising from 2.2 ms in 2008 to 2.6 ms in 2010. This difference of 0.4 ms corresponds to a growth of 0.5 m, following the equation of Gordon (1991) of the relationship between total body length and IPI. This growth in two years corresponds very well with the growth rate plots for adult females of this size (Best, 1970; Lockyer, 1981; Evans and Hindell, 2004). Hence, both the large dispersion of IPIs from towed hydrophone recordings and the difference between the median IPI between towed and Dtag recordings for whale #2 is likely to be highly influenced by growth. However, for some of the other whales very little of the variation in IPIs within years (whales #6 and #7, Fig. 4) or

between years (whales #3 and #6) can be explained by growth. Thus, in agreement with findings made in other studies (Rhinelander and Dawson, 2004; Miller et al., 2013) we conclude that IPIs may be able to speak to individual growth rates if the study is conducted over long enough time intervals for young, growing animals to minimize effects of the substantial uncertainty in the IPI measurements.

Thus, the recordings of sperm whale echolocation clicks both on and off the clicking animals over both short and long time intervals do show that larger whales with long IPIs consistently have long IPIs and that smaller whales consistently have shorter IPIs, but that IPI estimation has limited precision which results in an IPI value span up to 0.4 ms between recordings within the same year (Fig. 4). The largest difference between data values is observed between recordings of whale #3 and #6 in 2010 and 2005, respectively. A closer look at these recordings revealed that they all consisted of at least 44 accepted clicks and with no correlation that could be allocated to growth. Such absolute changes correspond to relative variations of some 12%–15% in IPI. Given that great care has been taken in this and other
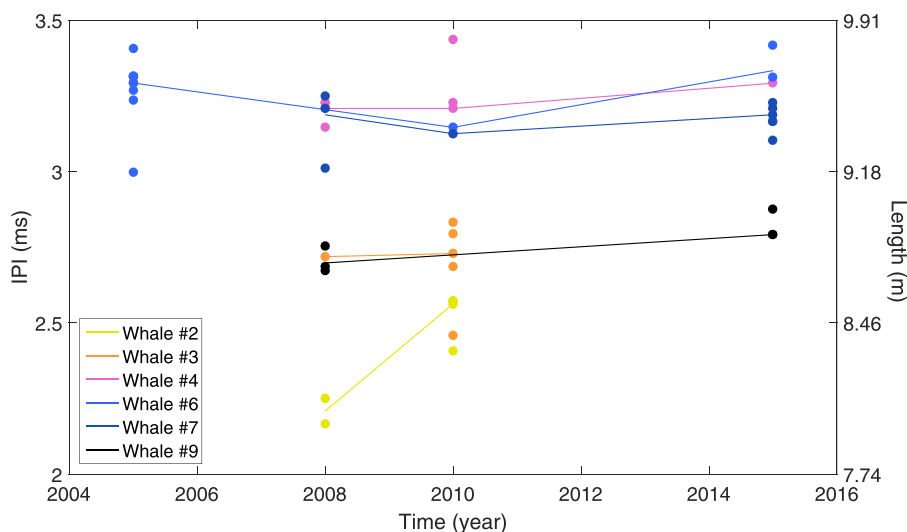

FIG. 4. (Color online) The change in echolocation click IPI (envelope method) within and between years measured by a towed hydrophone for the six whales where towed recordings were available for more than one year. • denotes the median echolocation click IPI from a single recording (containing between 4 and 368 clicks), whereas lines are drawn between the median of all echolocation click IPI estimates from one year to another. The calculated length from a given IPI using the equation given by Gordon (1991) is indicated on the right y-axis.

J. Acoust. Soc. Am. **144** (1), July 2018

Bøttcher et al.   371

studies to record in the same aspect (Antunes *et al.*, 2011; Growcott *et al.*, 2011; Schulz *et al.*, 2011; Miller *et al.*, 2013; Gero *et al.*, 2016) and that several analytical methods have been tested to drive down the variability, it is tempting to hypothesize that some of this variation is in fact source generated. If so, it follows that either the distance between the reflective air sacs or the sound speed in the spermaceti oil in concert should be able to change by up to 12%–15% in clicking sperm whales. A number of studies have shown that both pressure and temperature will affect the sound speed in spermaceti oil (Flewellen and Morris 1978; Goold *et al.*, 1996). To test the effects of pressure we therefore proceeded to extract IPIs in Dtag recordings on diving sperm whales.

E. IPI changes with depth

Figure 5 is an example of a dive profile and IPIs of whale #9 showing that there is an apparent negative correlation between IPI and the depth at which the click was emitted. This runs counter to the conclusion from analysis of clicks from a single sperm whale foraging dive, where Madsen *et al.* (2002a) concluded that the derived IPIs appeared stable from 100 to 700 m. That study, however, suffered from poor signal to noise ratios that may have prevented detection of smaller IPI changes. As the sound velocity in the spermaceti organ will increase with increasing pressure and decreasing temperature (Flewellen and Morris, 1978; Goold *et al.*, 1996), the IPIs are expected to decrease with depth. However, over the depth range considered here, pressure effects would only lead to a change in sound speed and hence IPI of <3% at any given temperature of relevance (Goold *et al.*, 1996). Temperature, on the other hand, is expected to provide a larger effect, but given that spermaceti oil solidifies at 28 °C (Clarke, 1978b), preventing any multi-pulses to be formed (Møhl, 2001), the maximum temperature change that is possible is from 37 °C to 29 °C, rendering less than 2% change in sound speed and hence IPI (Goold *et al.*, 1996). Further, given the substantial problems (Madsen 2002b) with the cooling theory of Clarke (1970), it is very unlikely that the temperature of the spermaceti oil deviates much from 37 °C in diving sperm whales. Nevertheless, all nine whales tagged with Dtags display depth related trends

in their IPIs (Fig. 6). We therefore argue that the most parsimonious explanation for the source variation in measured IPIs in diving sperm whales relate to changing distances between the two reflective air sacs that generate the multi-pulse structure. The sperm whale nose is supported anterio-ventrally by an amphitheater shaped skull, but the large majority of it, protruding anteriorly, is made of soft tissues and air sacs, allowing for conformation changes that may easily change the distance between the reflective air sacs by 10%. In fact, the upper part of the nasal complex is innervated by a large and complex set of longitudinal muscles that if contracted will pull the soft parts of the nose back towards the skull, perhaps to change the radiation pattern of the powerful echolocation clicks to modify the acoustic field of view (Møhl *et al.*, 2003) as known for smaller toothed whales (Wisniewska *et al.*, 2015). Thus, the IPI changes observed in diving sperm whales (Figs. 5 and 6) may be a passive consequence of biosonar adaptations to changing prey fields and auditory scenes over the course of a foraging dive.

Looking at the distribution of IPIs it is clear that all tag recordings suffer from a large degree of variation. Particularly the tendency of IPIs to be close to the lower boundary of 2 ms is distinctive and is concluded to be an artefact of the automatic IPI detection as clicks with no clear second pulse (p1) will lead to a lower IPI as the maximum peak within 2–5 ms after p0 is then likely to be the trailing part of p0. Attempts to truncate data by automated methods to eliminate clicks with no clear p1 have not been successful and hence all detected IPIs are plotted resulting in a reduced median estimate. In addition, a bimodal distribution of IPIs seems apparent for several whales (in particular whale # 3 and # 5). No explanation of this distribution has been found. Next, we proceeded to evaluate whether coda clicks have the same IPIs as the echolocation clicks from the same whale.

F. Difference in IPI between coda clicks and echolocation clicks

Coda clicks used for acoustic communication have a different pulse structure than echolocation clicks with a smaller decay rate between the individual pulses (Madsen *et al.*, 2002b). This seems to impact the dispersion of the calculated
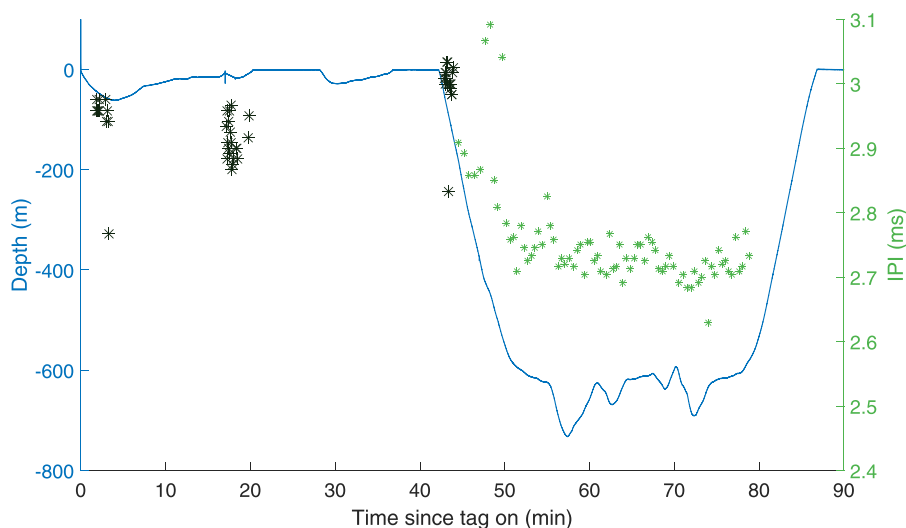

FIG. 5. (Color online) A dive recorded by one of the Dtags (whale #9). The dive profile is plotted in blue as a line whose axis is along the left side of the figure, IPI of echolocation clicks in light green, and median IPI of each coda in dark green/black, and whose axis is along the right side of the figure. The IPIs of echolocation clicks are plotted as the median of 50 IPIs to the mean time of these 50 IPIs.
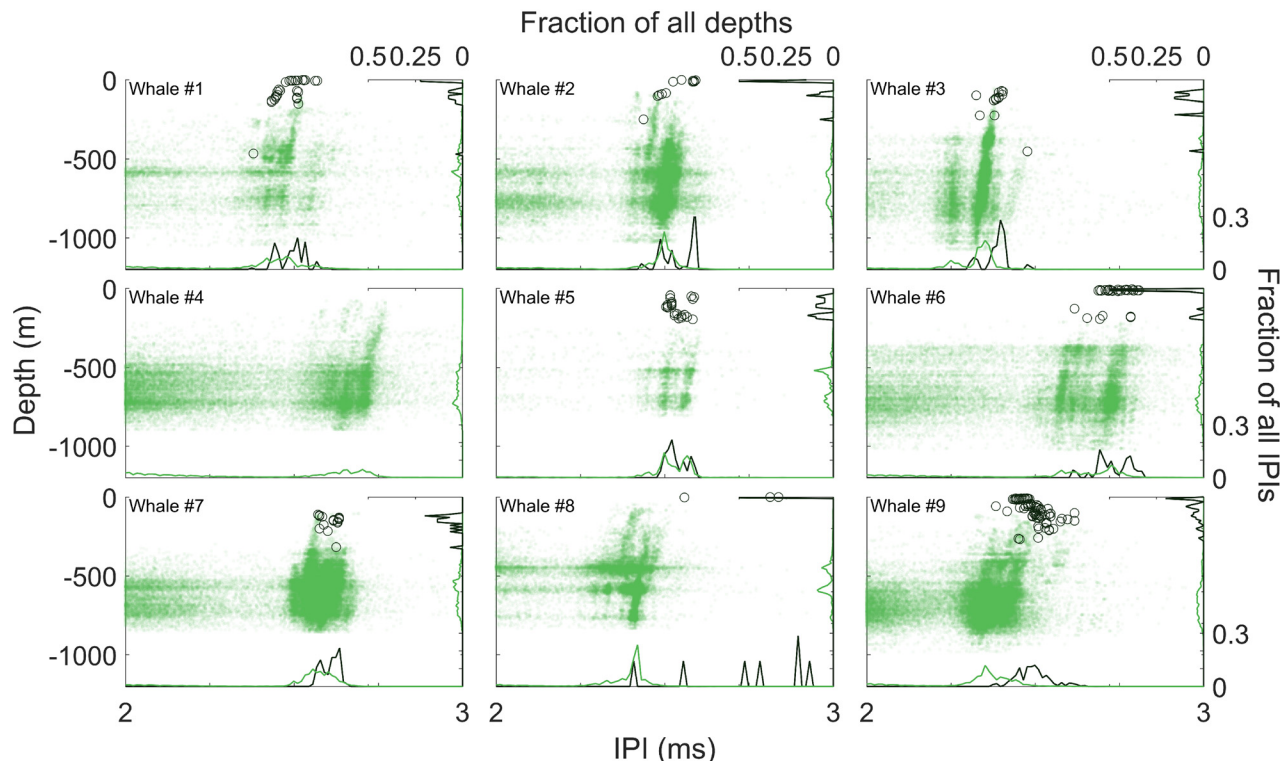
FIG. 6. (Color online) IPIs (bottom *x* axis) of echolocation clicks (light green dots) and codas (dark green circles) in relation to the depth at which they were emitted (left y-axis). The plots include echolocation clicks and codas from all Dtags. Dispersion of IPIs (bottom histogram, right y-axis) and depths (right histogram, top *x* axis) for echolocation clicks (light green) and coda clicks (dark green) are shown.

IPIs as the average IQR is lower for coda clicks (<0.17 ms) than echolocation clicks (>0.22 ms) regardless of which of the three methods used to estimate IPI. Especially IPIs estimated with cross-correlation revealed a lower dispersion with a mean IQR of 0.05 ms.

The difference in pulse structure between click types has been proposed to arise from conformational changes in the sound producing nose where inflation of the right nasal passage may trap the sound energy of coda clicks in the spermaceti organ, perhaps giving rise to lower directionality that may benefit the active space of coda communication (Madsen *et al.*, 2002b). As seen in both Figs. 5 and 6 where we have superimposed IPIs from coda clicks on IPI data for echolocation clicks, coda clicks consistently have IPIs that are longer than the echolocation clicks from the same sperm whale. Part of that may be explained by the fact that coda clicks predominantly are produced during social interactions close to the surface and therefore do not suffer from the weak depth effects on the IPIs of the echolocation clicks produced at depth for foraging. However, the coda click IPIs differ substantially, compared with what can be explained by sound speed changes in the spermaceti oil, suggesting that the longer coda IPIs may arise from conformation changes in the nasal complex serving to perhaps dramatically change the radiation pattern during acoustic communication. Irrespectively, the difference in IPIs between coda clicks and echolocation clicks makes it very difficult from far field recordings to assign coda clicks and echolocation clicks to the same sperm whale if recorded in social groups where distributions among adult individual animals are likely to overlap.

IV. CONCLUSION AND PERSPECTIVES

We have shown that recordings of the same whale result in variable estimates of IPI based on recording system used, analytical method applied, between replicate recordings, across depth, and amongst coda and echolocation clicks. As no combination of aspect alignment, depth truncation, and/or signal processing techniques can account for all the observed variation, we conclude that IPIs of echolocation clicks are likely also affected by conformational changes in the soft structures of the nasal complex. This modification may be conducted to adapt the radiation patterns of coda and echolocation clicks that consequently induce a change in IPIs. Regardless of the explanation for the variation, listening conspecifics as well as scientists will have to accommodate errors on the order of 0.2 ms when acoustically assessing the size and individual identity of an echolocating sperm whale. This corresponds to a size difference of 0.3 m, which is often within the differences in length between several females within a social unit. Hence, the applicability of IPI to identify specific individuals is limited without a more precise IPI estimate.

It is therefore also unlikely that similar sized female sperm whales use IPIs as sole cues for identification of individuals, although they may be part of a suite of acoustic cues allowing reliable individual identification if obvious problems of forward masking can be overcome to provide submillisecond time resolution in the sperm whale auditory system. Accordingly, we propose that the multi-pulse structure of sperm whale clicks has not been selected for as a size cue, but rather is a passive consequence of radiating clicks that

J. Acoust. Soc. Am. **144** (1), July 2018

Bøttcher *et al.* 373

must serve the very different purposes of echolocation and communication.

Adler-Fenchel, H. S. (1980). "Acoustically derived estimate of the size distribution for a sample of sperm whales (*Physeter catodon*) in the Western North Atlantic," Can. J. Fish. Aquat. Sci. 37, 2358–2361.

Antunes, R., Rendell, L., and Gordon, J. (2010). "Measuring inter-pulse intervals in sperm whale clicks: Consistency of automatic estimation methods," J. Acoust. Soc. Am. 127, 3239–3247.

Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J., and Rendell, L. (2011). "Individually distinctive acoustic features in sperm whale codas," Anim. Behav. 81, 723–730.

Best, P. B. (1970). "The sperm whale (*Physeter catodon*) off the west coast of South Africa 5, age, growth, and mortality" Invest. Rep. Div. Sea Fish. S. Afr. 79, 1–27.

Caruso, F., Sciacca, V., Bellia, G., De Domenico, E., Larosa, G., Papale, E., Pellegrino, C., Pulvirenti, S., Riccobene, G., Simeone, F., Speziale, F., Viola, S., and Pavan, G. (2015). "Size distribution of sperm whales acoustically identified during long term deep-sea monitoring in the Ionian Sea," PLoS One 10, e0144503.

Clarke, M. R. (1970). "Function of the spermaceti organ of the sperm whale," Nature 228, 873–874.

Clarke, M. R. (1978a). "Structure and proportions of the spermaceti organ in the sperm whale," J. Mar. Biol. Assoc. U. K. 58, 1–17.

Clarke, M. R. (1978b). "Physical properties of spermaceti oil in the sperm whale," J. Mar. Biol. Assoc. U. K. 58, 19–26.

Cranford, T. W. (1999). "The sperm whale's nose: Sexual selection on a grand scale?," Mar. Mammal Sci. 15, 1133–1157.

Cranford, T. W., Amundin, M., and Norris K. S. (1996). "Functional morphology and homology in the Odontoceti nasal complex implications for sound generation," J. Morphol. 228, 223–285.

Evans, K., and Hindell, M. A. (2004). "The age structure and growth of female sperm whales (*Physeter macrocephalus*) in southern Australian waters," J. Zool. Soc. London 263, 237–250.

Flewellen, C. G., and Morris, R. J. (1978). "Sound velocity measurements on samples from the spermaceti organ of the sperm whale (*Physeter catodon*)," Deep. Res. 25, 269–277.

Gero, S., Milligan, M., Rinaldi, C., Francis, P., Gordon, J., Carlson, C., Steffen, A., Tyack, P., Evans, P., and Whitehead, H. (2014). "Behavior and social structure of the sperm whales of Dominica, West Indies," Mar. Mammal Sci. 30, 905–922.

Gero, S., Whitehead, H., and Rendell, L. (2016). "Individual, unit and vocal clan level identity cues in sperm whale codas," R. Soc. Open Sci. 3, 150372.

Goold, J. C. (1996). "Signal processing techniques for acoustic measurement of sperm whale body lengths," J. Acoust. Soc. Am. 100, 3431–3441.

Goold, J. C., Bennell, J. D., and Jones, S. E. (1996). "Sound velocity measurements in spermaceti oil under the combined influences of temperature and pressure," Deep. Res. Part I Oceanogr. Res. Pap. 43, 961–969.

Gordon, J. C. D. (1991). "Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations," J. Zool. 224, 301–314.

Growcott, A., Miller, B., Sirguey, P., Slooten, E., and Dawson, S. (2011). "Measuring body length of male sperm whales from their clicks: The relationship between inter-pulse intervals and photogrammetrically measured lengths," J. Acoust. Soc. Am. 130, 568–573.

Jarque, C. M., and Bera, A. K. (1980). "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," Econ. Lett. 6, 255–259.

Johnson, M. P., and Tyack, P. L. (2003). "A digital acoustic recording tag for measuring the response of wild marine mammals to sound," IEEE J. Ocean. Eng. 28, 3–12.

Lockyer, C. (1981). "Estimates of growth and energy budget for the sperm whale," *Mammals in the Seas*, FAO Series No. 5, Vol. III, pp. 489–504.

Madsen, P. T. (2002a). "Morphology of the sperm whale head: A review and some new findings," Ph.D. thesis, University of Aarhus, Denmark, access: peter.madsen@bios.au.dk

Madsen, P. T. (2002b). "Sperm Whale Sound Production - in the acoustic realm of the biggest nose on record," Ph.D. thesis, University of Aarhus, Denmark, access: peter.madsen@bios.au.dk

Madsen, P. T., Carder, D. A., Au, W. W. L., Nachtigall, P. E., Møhl, B., and Ridgway, S. H. (2003). "Sound production in neonate sperm whales," J. Acoust. Soc. Am. 113, 2988–2991.

Madsen, P. T., Payne, R., Kristiansen, N. U., Wahlberg, M., Kerr, I., and Møhl, B. (2002a). "Sperm whale sound production studied with ultrasound time/depth-recording tags," J. Exp. Biol. 205, 1899–1906.

Madsen, P. T., Wahlberg, M., and Møhl, B. (2002b). "Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: Implications for echolocation and communication," Behav. Ecol. Sociobiol. 53, 31–41.

Mathias, D., Thode, A., Straley, J., and Folkert, K. (2009). "Relationship between sperm whale (*Physeter macrocephalus*) click structure and size derived from video camera images of a depredating whale (sperm whale prey acquisition)," J. Acoust. Soc. Am. 125, 3444–3453.

Miller, B. S., Growcott, A., Slooten, E., and Dawson, S. M. (2013). "Acoustically derived growth rates of sperm whales (*Physeter macrocephalus*) in Kaikoura, New Zealand," J. Acoust. Soc. Am. 134, 2438–2445.

Møhl, B. (2001). "Sound transmission in the nose of the sperm whale *Physeter catodon*. A post mortem study," J. Comp. Physiol. A 187, 335–340.

Møhl, B., Larsen, E., and Amundin, M. (1981). "Sperm whale size determination: Outlines of an acoustic approach," in *Mammals in the Seas*, FAO Series No. 5, Vol. III, pp. 327–332.

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (2003). "The monopulsed nature of sperm whale clicks," J. Acoust. Soc. Am. 114, 1143–1154.

Norris, K. S., and Harvey, G. W. (1972). "A theory for the function of the spermaceti organ of the sperm whale (*Physeter catodon L.*)," NASA Special Publication No. 262, Animal Orientation and Navigation, pp. 397–417.

Rendell, L., and Whitehead, H. (2004). "Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurement," Anim. Behav. 67, 865–874.

Rhinelander, M. Q., and Dawson, S. M. (2004). "Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length," J. Acoust. Soc. Am. 115, 1826–1831.

Schulz, T., Whitehead, H., Gero, S., and Rendell, L. (2011). "Individual vocal production in a sperm whale (*Physeter macrocephalus*) social unit," Mar. Mammal Sci. 27, 149–166.

Schulz, T. M., Whitehead, H., Gero, S., and Rendell, L. (2008). "Overlapping and matching of codas in vocal interactions between sperm whales: Insights into communication function," Anim. Behav. 76, 1977–1988.

Schulz, T. M., Whitehead, H., and Rendell, L. (2009). "Off-axis effects on the multi-pulse structure of sperm whale coda clicks," J. Acoust. Soc. Am. 125, 1768–1773.

Teloni, V., Zimmer, W. M. X., Wahlberg, M., and Madsen, P. T. (2007). "Consistent acoustic size estimation of sperm whales using clicks recorded from unknown aspects," J. Cetacean Res. Manag. 9, 127–136.

Whitehead, H., and Weilgart, L. (1990). "Click rates from sperm whales," J. Acoust. Soc. Am. 87, 1798–1806.

Wisniewska, D. M., Ratcliffe, J. M., Beedholm, K., Christensen, C. B., Johnson, M., Koblitz, J. C., Wahlberg, M., and Madsen, P. T. (2015). "Range-dependent flexibility in the acoustic field of view of echolocating porpoises (*Phocoena phocoena*)," Elife 4, e05651.

Zimmer, W. M. X., Madsen, P. T., Teloni, V., Johnson, M. P., and Tyack, P. L. (2005a). "Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production," J. Acoust. Soc. Am. 118, 3337–3345.

Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (2005b). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," J. Acoust. Soc. Am. 117, 1473–1485.