

# Consistent acoustic size estimation of sperm whales using clicks recorded from unknown aspects

VALERIA TELONI\*, WALTER M.X. ZIMMER†, MAGNUS WAHLBERG\*‡, PETER T. MADSEN\*‡‡

Contact e-mail: vteloni@tiscali.it

## ABSTRACT

The multipulse structure of sperm whale clicks offers a unique way to acoustically estimate body length, as the inter-pulse intervals within the clicks relate to the two-way travel time within and thereby to the size of the hypertrophied nose in this species. Despite its large potential to allow the estimation of length acoustically, the technique has only been used in a few studies to assess the length composition of sperm whale populations. Its limited use may relate to the fact that only some clicks within a click series normally display the regular multipulsed structure required for size estimation. The inter-pulse intervals of usual clicks vary with the recording aspect to the clicking whale and the pulse delays are not necessarily directly related to the length of the spermaceti organ. To overcome these difficulties, a method is provided to estimate sperm whale body lengths, based on averages of cepstra derived from a large number of clicks recorded from whales in unknown recording aspects. This study shows that the two-way travel time in the spermaceti organ can consistently be estimated by a peak in the averaged cepstra when a large number of clicks are analysed. This method is shown to give a consistent estimation of the size of the spermaceti organ when recording the whale in an unknown orientation and also when recordings are heavily influenced by surface reflections.

KEYWORDS: SPERM WHALE; VOCALISATION; SURVEY; ACOUSTICS; MONITORING

## INTRODUCTION

Estimation of length in marine mammals plays an important role in the assessment of stock size and can be used to derive the age and stage of sexual maturity of the individuals (Angliss *et al.*, 1995; Koski *et al.*, 1993). Photogrammetric techniques are used to measure the length of whales at sea (Cosens and Anders, 2003; Dawson *et al.*, 1995; Gordon, 1990; Jaquet, 2006; Miller *et al.*, 2004). These techniques require good weather and considerable effort if a reasonable fraction of a population is to be covered, especially for deep-diving species that spend little time at the surface (Watkins *et al.*, 1999). In addition, there are some biases associated with photographic measurements of body size (Angliss *et al.*, 1995).

The multipulse structure of sperm whale clicks (Backus and Schevill, 1966) offers a unique way to estimate the body length of individuals acoustically. This potential was first appreciated by Norris and Harvey (1972) from combined observations of the click structure and nasal anatomy of this species. As part of their original theory on sperm whale sound production, the authors proposed that the time between the pulses (the inter-pulse interval or IPI) within a single sperm whale click is given by the two-way travel time in the spermaceti organ. Norris and Harvey envisioned that most of the sound energy generated at the *museau de singe*, a valve-like structure of connective tissue, escapes the nose directly and that fractions of sound energy are reflected backwards by an air sac (the distal sac) into the spermaceti organ and subsequently forward by another air sac lining the skull (the frontal sac). These internal reflections repeat themselves, generating a train of pulses with decaying amplitude.

The Norris and Harvey theory has been amended by Møhl *et al.* (1981) and Møhl (2001), who proposed that the bulk of the sound energy is first directed backwards into the spermaceti organ while only a small portion of the initial

sound energy leaks directly into the water, generating the first  $p_0$  pulse (Fig. 1, top panel). The primary sound pulse ( $p_1$ ) is emitted into the water from the junk after reflection from the frontal air sac and is followed by the  $p_2$  pulse after one more round through the spermaceti organ. This so-called bent horn model has been corroborated by more recent findings (Madsen *et al.*, 2002; Møhl *et al.*, 2003; Zimmer *et al.*, 2005a; Zimmer *et al.*, 2005b).

The underlying concept for acoustic length estimation of sperm whales is that the IPIs are given by the distance between the sound reflectors and the speed of sound in the spermaceti oil, measured by Flewellen and Morris (1978) and Goold *et al.* (1996). Knowing the allometric relationship between the size of the nose and the overall body size of the sperm whale (Clarke, 1978), the length of the whale can be estimated acoustically. Alternatively, the relationship between IPI and body length may be derived directly by acoustic and photogrammetric studies of the same individuals (Gordon, 1991; Rhineland and Dawson, 2004). Thereby one may envisage a simple technique to survey large areas for the size composition of sperm whale populations.

Even though the potential for acoustic length estimation was proposed more than 30 years ago, the technique has only been explored in few studies (Adler-Fenchel, 1980; Drouot *et al.*, 2004; Gordon, 1991; Leaper *et al.*, 1992; Marcoux *et al.*, 2006; Pavan *et al.*, 2000; Rendell and Whitehead, 2004; Rhineland and Dawson, 2004). One of the main reasons for this could be that the pulses within clicks seldom display the clear pattern of regular intervals needed for consistent measurement. Adler-Fenchel (1980) found that only about 11% of the total number of clicks met the criteria for analysis based on signal-to-noise ratio and phase relationship between consecutive pulses within the clicks. Gordon (1991) reported frequent incidences of 'anomalous clicks', with a less well-defined pulse structure. He suggested that variations in click structure might be

\* Zoophysiology, Dept. of Biological Sciences, University of Aarhus, Build. 1131, 8000 Aarhus C, Denmark.

† NATO Undersea Research Centre, V.le San Bartolomeo 400, 19126 La Spezia, Italy.

‡ Fjord&Bælt, Margrethes Plads 1, 5300 Kerteminde, Denmark.

‡‡ Woods Hole Oceanographic Institution, MA 02543, USA.

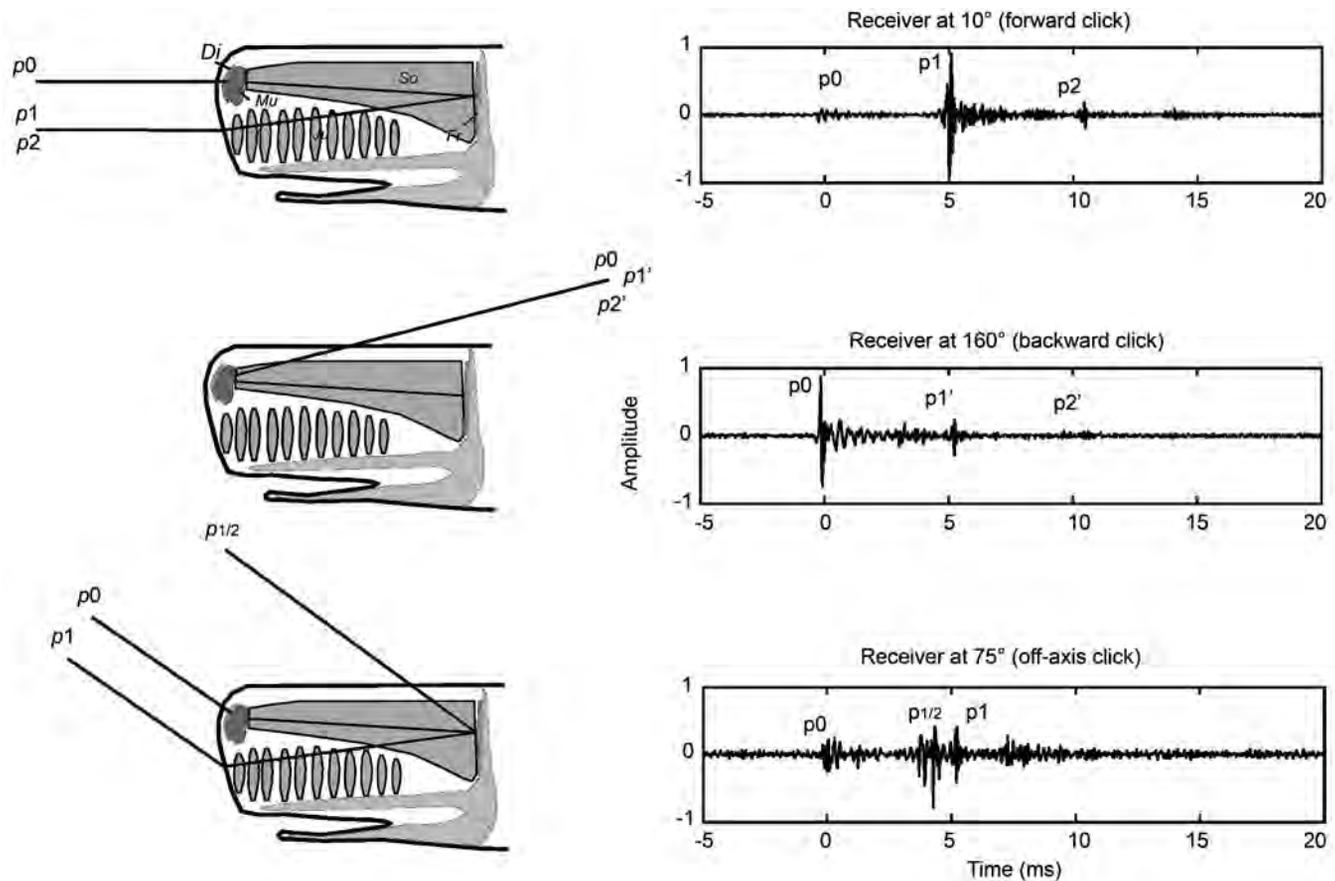


Fig. 1. Schematic drawing of a sperm whale head (left) and corresponding multi-pulse structure of usual clicks (right), with the receiver at different bearings (from Zimmer *et al.*, 2005a). Top panel is a nearly on-axis click (*So*: spermaceti organ; *Ju*: junk; *Mu*: museau de singe; *Di*: distal sac; *Fr*: frontal sac). Middle panel is a click recorded at a caudal aspect. Bottom panel is a usual click recorded at an angle of 75°. Nasal passages are omitted for simplicity. The pulses  $p_1$  and  $p_2$  are represented with the same line, where  $p_2$  is subjected to an additional internal reflection from the distal to the frontal sac (not shown here) before leaving the junk with the assumed same path as  $p_1$ . Clicks were selected from whale no.2 ( $n=1,997$ , Table 1).

caused by changes in the reflective characteristics of the head or the relative orientation of the whale and the receiver. In addition, Goold (1996) found few clicks with a pulse structure well-suited for measuring IPIs and he therefore stated that no reliable measurements of IPI could be derived from single sperm whale clicks. He proposed that the failure of the methods based on individual clicks might be the results of off-axis distortion from a directional source.

Rhineland and Dawson (2004) examined IPI stability in photographically identified individuals and found that the measured inter-pulse intervals were stable over short periods of time. Their recordings were taken right after the whale's fluke up from a location on top of the footprint and were limited to the first minutes of the dive to minimise any changes in the IPI.

In summary, previous attempts to measure inter-pulse intervals in clicks have been made with manual measurements on individual clicks, where clicks that do not comply with the clear multi-pulse pattern are removed from the analysis (Goold *et al.*, 1996; Rendell and Whitehead, 2004; Rhineland and Dawson, 2004).

Zimmer *et al.* (2005a) have recently shown that a geometric model consistent with the bent horn theory of sperm whale sound production (Møhl *et al.*, 2003) can explain the varying multi-pulse structure of off-axis clicks. Using clicks recorded from known aspects to the whale, they demonstrate that the  $p_1$  pulse indeed exits from the frontal area of the junk, whereas the initial  $p_0$  pulse exits from the nose at the site close to the sound-producing museau de singe (Madsen *et al.*, 2003). Further, they show

that an intermediate pulse (named  $p_{1/2}$ ) seen in off-axis recordings, is generated by the reflection of the initial pulse from the frontal sac (Fig. 1, bottom panel). The  $p_{1/2}$  can have delays between 0, merging with the  $p_0$  when recorded behind the whale (Fig. 1, middle panel), and the two-way travel time of the spermaceti organ, merging with  $p_1$  (Fig. 1, top panel) when recorded right in front of the whale (Zimmer *et al.*, 2005a). These findings have made it clear that some of the pulse structures of clicks, such as the  $p_0$ - $p_{1/2}$  and  $p_0$ - $p_1$  intervals, vary with the recording aspect of the clicking whale. Therefore, sperm whale length cannot readily be estimated correctly from the pulse timings of a single click recorded off the body axis of the whale, without knowing its relative aspect. Further complications may be introduced by interference with surface-reflected paths confusing the original inter-pulse interval patterns.

In this paper, the findings of Zimmer *et al.* (2005a) are exploited to test a method for estimating the inter-pulse intervals of sperm whale usual clicks, using single hydrophone recordings. From Zimmer *et al.*'s findings and the predictions of the bent horn model (Møhl *et al.*, 2003), it is hypothesised that other inter-pulse intervals within a click, such as the  $p_1$ - $p_2$ , remain constant, independent of aspect. However, even though the regular inter-pulse intervals pattern will be found in all clicks including those recorded off-axis directions, they might be obscured by additional aspect-dependent peaks. Rather than being based on inter-pulse interval measurements of a single click, or short sequences of clicks recorded with the whale at a known aspect, a method is presented here that calculates the

mean of cepstra (Bogert *et al.*, 1963) derived from long click sequences. The study shows that the two-way travel time in the spermaceti organ can be estimated from a distinct peak in the averaged cepstrum after processing of a few hundreds clicks. The method is robust for whales recorded in different aspects, for recordings with interfering surface-reflected paths, as well as for recordings of limited bandwidth and dynamic range, and these findings are discussed in the light of the potential for passive acoustic monitoring of sperm whale populations with single hydrophone recordings.

## MATERIALS AND METHODS

In this paper, the term 'click' refers to a 'normal' sperm whale click. Further, the term 'inter-pulse interval' is used to describe any interval between two pulses in the click, and the term 'nominal IPI' as the longitudinal two-way travel time in the spermaceti organ of the sperm whale nose.

The datasets used for the analysis are listed in Table 1. The Mediterranean dataset consists of two sperm whales that were recorded in the Ligurian Sea (northwestern part of the Mediterranean Sea) in 2000 and 2001 during sea trials organised by the NATO Undersea Research Centre (NURC). The passive sonar system developed at NURC consists of a horizontal line array of 128 hydrophones, a real-time digital beamformer and a sonar display system. The hydrophone array was towed at a depth of about 80m, just below the thermocline. The hydrophones of the towed array were set to an effective saturation level of 140 dB<sub>peak</sub> re: 1µPa and sampled with 16 bit resolution at 31.25kHz, providing a maximum bandwidth of about 15kHz. To suppress the low frequency flow and ambient noise, the data were filtered with a 3kHz high-pass filter. Click sequences were recorded from two sperm whales tracked for 7 and 10 complete foraging dives, respectively. To develop the technique described here, recordings from a single hydrophone were focused on. The aspect information for clicks in Fig. 1 is taken from the work of Zimmer *et al.* (2005a).

The Indian Ocean dataset consists of a single sperm whale dive recorded off the Seychelles from the research vessel *Odyssey* of the Ocean Alliance ([www.oceanalliance.org](http://www.oceanalliance.org)). Recordings were made while drifting close to the fluke up of a single male sperm whale performing foraging dives and producing usual clicks and buzzes. Recordings were performed with a towed two-element array. For the present recordings the array was allowed to sink to an estimated depth of about 70m. Analogue signals from the hydrophones were amplified and band-pass filtered with a 2-pole cut-off and corner frequencies (-3dB) at 1.6 and 12kHz. The

amplified, filtered output signal was interfaced with a PC desktop computer and digitised with a stereo 16-bit sound card sampling at 48kHz.

The Faroe Islands dataset was recorded in 1975 in the Skálafjordur from a sperm whale trapped in the fiord. Recordings were made with a B&K 8100 stationary hydrophone deployed at a depth of 6m, and stored on a Nagra IV-D tape recorder. The frequency response was flat to about 20kHz.

The North Atlantic datasets were collected off Andenes (Northern Norway) during summers 2000 and 2005. The first sperm whale was recorded in 2000 from five independent platforms with hydrophones at depths between 5 and 30m. For details on the recording geometry see fig. 3 in Møhl *et al.* (2003). The signals were recorded on digital tape recorders (Sony TCD-D3, 7, 8 and 10) sampling at 48kHz. The platforms were spaced more than 300m apart. The second sperm whale was recorded in 2005 from two separate platforms with hydrophones (B&K 8101 and Reson 4034) deployed at 30m and 15m depth, respectively. The two platforms were more than 500m apart while the animal was diving. The signals were recorded on digital tape recorders (Sony TCD-D7) at a sampling frequency of 48kHz. A digital tag recorder (Dtag, Johnson and Tyack, 2003) attached to the whale provided inter-click interval information of click sequences that allowed for identification of the same whale on the two separate platforms.

All sperm whales recorded were within an estimated 1-2km from the hydrophones and the depth of the whale varied from near surface to deep water, sometimes within the same click sequence. For all whales, only usual clicks were selected for the analysis. Clicks from creaks have a lower intensity compared to usual clicks, rendering analysis difficult or impossible (Madsen *et al.*, 2002). The number of clicks analysed for each whale is reported in Table 1. Clicks were detected by a custom-written click detector with *MatLab* (*Math works 6.0*).

Cepstrum analysis (Bogert *et al.*, 1963; Oppenheim *et al.*, 1968) was used to determine the repeated patterns within the interpulse intervals of usual clicks. The real-valued cepstrum analysis (where the term 'cepstrum' is simply 'spectrum' with the first four letters in reverse order) takes as input the time series  $x_t$  and constructs the cepstrum  $C_t$  by the nonlinear transformation:

$$C_t = \left| \text{FFT}^{-1}(\log|\text{FFT}(x_t)|) \right| \quad (1)$$

where FFT is the fast Fourier transform and  $\text{FFT}^{-1}$  is the inverse fast Fourier transform and denotes the absolute value (Zimmer *et al.*, 2003). The cepstrum converts the

Table 1

Datasets used for the measurements of inter-pulse intervals of sperm whale usual clicks. Clicks within each dataset are known to be from the same whale. Clicks are consecutive from the same dive recorded in open ocean, except for the whale in the Faroe Islands which was recorded in a shallow fiord. Clicks from datasets 1, 2 and 4 are all usual clicks recorded during a single dive. The North Atlantic dataset from 2005 consists of two distinct click sequences (141 and 831 clicks) recorded from the same whale during the same dive from two separate platforms. Estimated nominal IPIs are from the cepstral averages of all clicks.

Data set	Number of clicks	Origin	Year	Recording type	Hydrophone depth (m)	Sampling frequency (kHz)	Estimated nominal IPI (ms)
Whale no. 1	2,303	Mediterranean Sea	2000	Single hydrophone from towed array	80	31.25	5.71 ± 0.05 <sup>§</sup>
Whale no. 2	1,997	Mediterranean Sea	2001	Single hydrophone from towed array	80	31.25	5.44 ± 0.1 <sup>§</sup>
Whale no. 3	252	North Atlantic	2000	5 drifting hydrophones	5 and 30	48	6.86 (0.02 <sup>#</sup> )
Whale no. 4	141 + 831	North Atlantic	2005	2 drifting hydrophones	15 and 30	48	6.98 (0.02 <sup>#</sup> )
Whale no. 5	2,549	Indian Ocean	2002	Single hydrophone from towed array	70	48	7.70 ± 0.1 <sup>§</sup>
Whale no. 6	292	Faroe Islands	1975	Drifting hydrophone	6	20*	7.58 ± 0.01 <sup>§</sup>

\*Frequency response of the recording system. <sup>§</sup>Width of the cepstral peak, defined at 50% of its amplitude. <sup>#</sup>IPI deviation between the separate platforms.

logarithm of the power spectrum back into a time domain presentation ( $C_t$ ), where peaks appear at delay times identical to the time between repeating structures (pulses) in the original time signal (sperm whale click). For the cepstral peak to appear, the length of the time window used in the analysis must be at least twice as long as the inter-pulse interval. Here a duration of 50ms was chosen. The click cepstra were subsequently averaged. The Matlab code for averaging cepstra of a click series and subsequent measurement of the nominal inter-pulse interval is provided in Appendix A.

## RESULTS

The cepstral analysis was run for six whales and the results are listed in Table 1. The averaged cepstra from whale no.1 gave an estimated nominal IPI of 5.71ms. For whale no.2, an on-axis click was selected ( $10^\circ$  aspect, see Fig. 1) and the nominal IPI was measured as the interval between  $p1$  and  $p2$  peaks in the Hilbert transformed click. This value fitted with the averaged cepstral peak at a value of 5.44ms (dotted line in Fig. 2). The averaged cepstra for an increasing number of consecutive clicks taken from whale no.2 are shown in Fig. 2. The average of 5 and 10 clicks shows several peaks from which the nominal IPI is difficult or impossible to discriminate. A predominant peak starts to emerge after averaging 100 clicks. By using all the 1,977 clicks available for the dive, a distinct peak has stabilised at the end of the plateau. The shape of the curve reflects the degree of variation of the inter-pulse intervals within the clicks. The peak indicates repeated measurements with the same value, whereas the plateau to the left of the peak represents the varying aspect-dependent inter-pulse intervals (e.g.  $p0-p1/2$ ).

Cepstrum statistics for whale no.3 are shown in Fig. 3 (top panel). Stations  $r0$ ,  $r1$  and  $r6$  gave a clear peak at a value of 6.86ms, while the averaging of clicks from stations

$r3$  and  $r4$  gave a less clear result, but centred around the above value. The two click sequences from whale no.4 were processed separately and the averaged click cepstrum is shown in Fig. 3 (bottom panel). The black curve corresponds to 141 clicks recorded from one platform and the grey curve to 831 clicks recorded from the second one. The two peaks of the estimated nominal IPI, which were obtained by averaging clicks from different stages of the same dive and from independent platforms, matched at a value of 6.98ms.

The averaged cepstra for an increasing number of clicks from whale no.5, for which no information about the size of the whale or its relative orientation were available, are shown in Fig. 4. In this case averaging 100 clicks is not enough to generate a clear peak, which starts to emerge from 1,000 clicks and onwards. The cepstrum statistics of the first 500 clicks where the measurement might be taken at the end of the plateau, would give a nominal IPI of 7.65ms. This would lead to a difference in whale length of about 8cm (Rhinelander and Dawson, 2004), compared to the nominal IPI derived from all clicks (7.70ms, Table 1).

The averaged cepstra for whale no.6 are shown in Fig. 5. The distance between  $p1$  and  $p2$  was measured from one of the Hilbert transformed clicks (inset, Fig. 5), and the peak matching with this interval is at 7.6ms. The two broader peaks around 2 and 3ms correspond to the surface reflections in the recordings. They clearly stand out in the clicks and can easily be discriminated from the peak of the nominal IPI, but they are smeared due to variations in the recording geometry during the click sequence.

## DISCUSSION

### Recording aspect and inter-pulse intervals of clicks

Fig. 1 shows how the multipulse structure of a sperm whale click varies with the recording aspect. Considering that sperm whales usually dive deep (Papastavrou *et al.*, 1989)

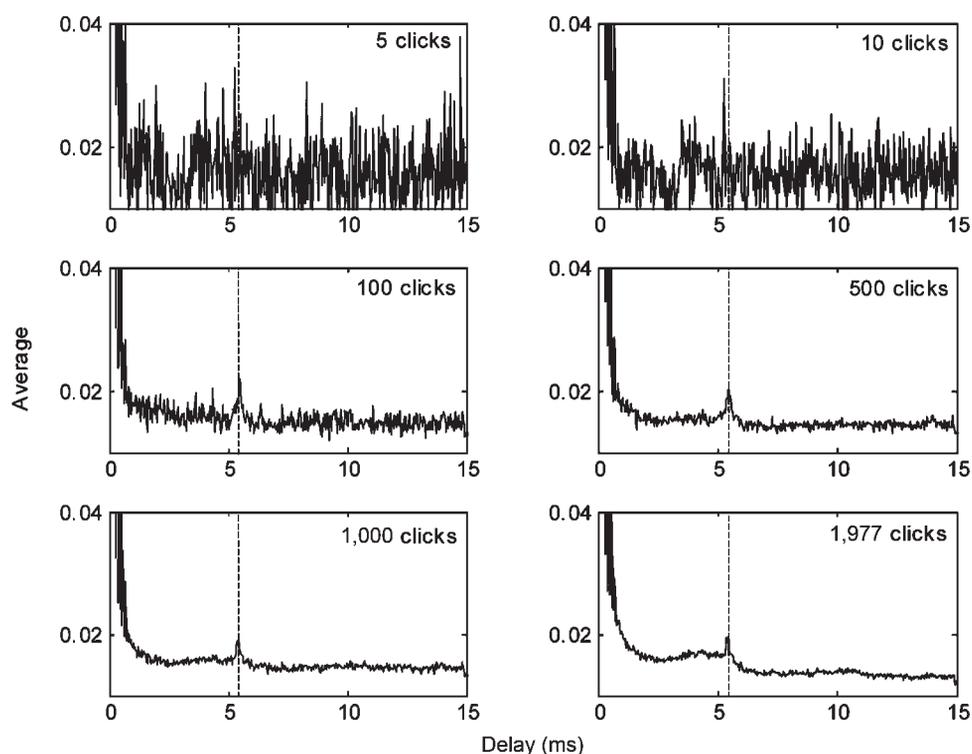


Fig. 2. Average values of the cepstrum for 5, 10, 100, 500, 1,000 and 1,977 consecutive clicks from the same dive from whale no.2 ( $n=1,977$ , Table 1). The dashed line corresponds to the nominal IPI measured by hand from an on-axis click between  $p1$  and  $p2$ . The delay in ms on the  $x$  axis is termed quefrency in cepstrum analysis (Bogert *et al.*, 1963).

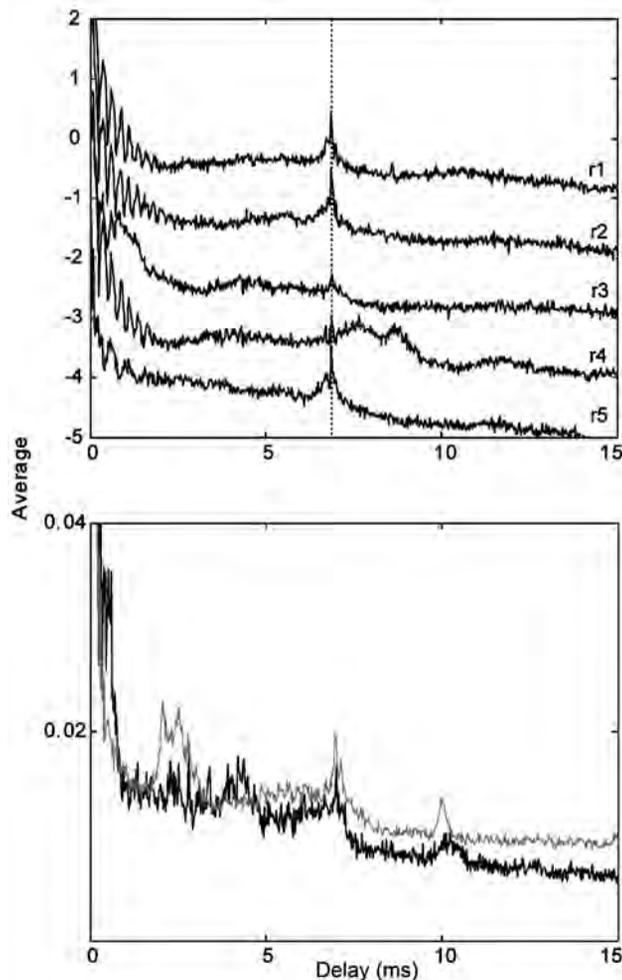


Fig. 3. Cepstral averages for clicks recorded in the North Atlantic. Top panel: clicks are from the same whale recorded simultaneously from 5 platforms located in different directions from the whale (whale no.3, Table 1). Peaks from stations  $r1$ ,  $r2$ ,  $r3$  and  $r5$  are marked with a dashed line and match at 6.86ms. Averages are shifted vertically for clarity. Bottom panel: clicks are from the same whale but recorded at different stages of the same dive from two separate platforms ( $n=141$  black curve and  $n=831$  grey curve, whale no.4, Table 1). The peaks of the nominal IPI match at 6.98ms, while the other peaks before 5ms and at 10ms are other reflections. The delay in ms on the x axis is termed quefrency in cepstrum analysis (Bogert *et al.*, 1963).

and at a steep angle (Miller *et al.*, 2004) and normally stop clicking during the ascent phase of the dive (Jaquet *et al.*, 2001; Madsen *et al.*, 2002; Teloni *et al.*, 2003), the chance of recording on-axis clicks in front of the whale (Fig. 1, top panel) when working with surface hydrophones is small (Møhl *et al.*, 2003). In addition, if the high intensity and large bandwidth of the  $p1$  pulse (Møhl *et al.*, 2003) are not considered when choosing the recording settings, the signals will most likely be clipped in the dynamic as well as in the frequency range. When recording from behind (Fig. 1, middle panel), the  $p0$  pulse is the most intense pulse, as it leaks out of the spermaceti organ with a backward-directed beam in the opposite direction of the  $p1$  pulse projected out in front of the whale (Zimmer *et al.*, 2005b). The second pulse seen in this caudal aspect has made one round through the spermaceti organ and then re-radiated from the whale at a site very close to the  $p0$  pulse. This pulse is called the  $p1'$  pulse and the consecutive pulse the  $p2'$  pulse, to indicate that these pulses are not emitted through the junk like their relatives  $p1$  and  $p2$ . Thus, the pulses recorded at the receiver in aspects larger than 90 degrees are not the  $p1$  and  $p2$  pulses

leaving the junk, but the pulses (named  $p1'$  and  $p2'$ ) that leak from the spermaceti organ after reflection at the distal sac. It is seen that the inter-pulse intervals can easily be derived when recording behind the animal, as is often the case for IPI measurements made over the footprint of a diving whale (Goold and Jones, 1995; Gordon, 1987; Rhinelander and Dawson, 2004). Thus, the multipulse structure stands out clearly when recorded on the body axis, either in front of the animal or behind it, and the nominal IPIs can be derived.

As part of the energy of the pulse reflected by the frontal sac might leak from the spermaceti organ (Zimmer *et al.*, 2005a), the multipulse structure is still present, but the nominal inter-pulse interval pattern is cluttered by additional aspect-dependent pulse delays (Fig. 1, bottom panel). This may lead to confusion when inspecting a single click, both in the time and cepstral domains (Fig. 6, panels C and D). In addition, surface-reflected paths may interfere with the direct clicks, leaving the analyst with even more confusing pulse patterns to choose amongst.

#### The key to the nominal IPI: averaging cepstra derived from click sequences

These problems in deriving the nominal IPI call for an automated method. The approach derived here relies on the analysis by Zimmer *et al.* (2005a): every click of a sperm whale will contain both inter-pulse intervals corresponding to the nominal IPI and additional IPIs depending on the aspect of the whale with respect to the hydrophone. In addition, there may be inter-pulse intervals caused by interference with surface-reflected paths, and such intervals will depend on the depth of the recording geometry between the hydrophone and the clicking whale. In many clicks, the aspect-dependent interclick intervals may dominate the picture completely, but according to the outline above, the nominal inter-pulse interval will always be there, in every single click. Thus, from a long series of clicks where the aspect, range and depth of the whale may vary considerably, all aspect- and geometry-dependent effects may be averaged out, and the nominal IPI is left as a reliable estimator of the two-way travel time in the spermaceti organ, which in turn conveys the size of the nose and thereby the size of the animal.

Cepstrum analysis is a convenient tool to extract inter-pulse intervals from sperm whale clicks. It has the advantage of not being sensitive to errors in temporal alignment between extracted clicks. Due to the variations in amplitudes of the various pulses in a click, the click detector may easily misalign the clicks according to a specific pulse. The cepstrum, however, is insensitive to such jitter in the alignment, as long as the time window extracted for every click is large enough to cover the whole click. Thereby time-consuming and tedious hand-processing of clicks is avoided. The application of cepstrum analysis in measuring inter-pulse intervals was first exploited by Goold (1996), but it has been applied in practice only a few times (Pavan *et al.*, 1997; Teloni *et al.*, 2000).

Instead of cepstrum analysis, autocorrelation analysis is a viable alternative for measuring repeated patterns within the clicks (Goold, 1996; Rhinelander and Dawson, 2004). Averaging the autocorrelation function for the clicks in a sequence would render results very similar to those described below for the cepstrum analysis of sperm whale clicks. If sperm whale clicks contain multiple pulses then cepstrum analysis is the preferred method as it designed to estimate the periodicity of multi-pulse structures.

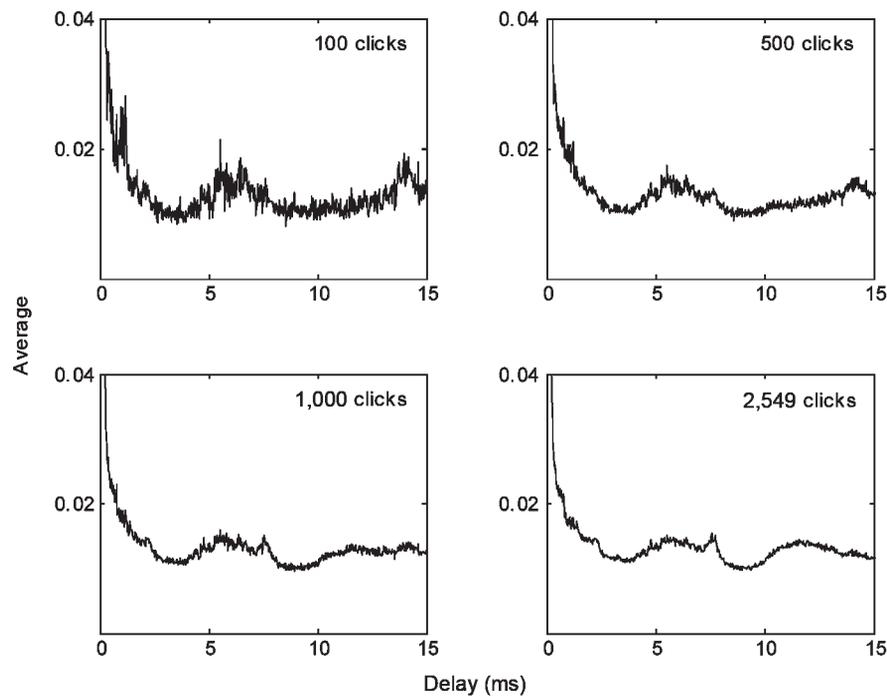


Fig. 4. Average values of the cepstrum for 100, 500, 1,000 and 2,549 clicks. Clicks are the total consecutive usual clicks ( $n=2,549$ , Table 1) from the same dive from whale no.5. The delay in ms on the x axis is termed quefrency in cepstrum analysis (Bogert *et al.*, 1963).

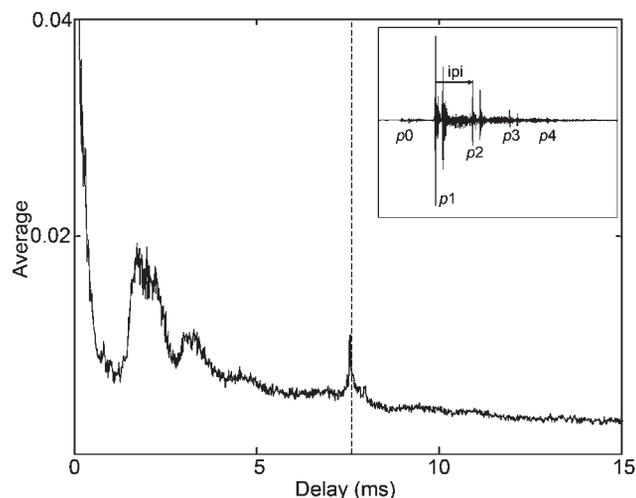


Fig. 5. Cepstral averages for all clicks ( $n=292$ ) recorded from whale no.6 in the Faroe Islands. The solid line crossing the peak at 7.6ms represents the nominal IPI measured as the distance between  $p1$  and  $p2$  (click in the inset, where the pulses are marked as  $p0$ ,  $p1$ ,  $p2$  and  $p3$ ). The two broad peaks at around 2 and 3ms are surface reflections. The delay in ms on the x axis is termed quefrency in cepstrum analysis (Bogert *et al.*, 1963).

### Test of cepstral measurements

Averaging of click cepstra does not always give results as clear as those for whale no.2 (Table 1 and Fig. 2). In the recording from a male sperm whale in the Indian Ocean (Fig. 4), the nominal IPI started to become apparent only after averaging 500 clicks and was never as well-defined as for the Mediterranean whale. Surface reflections are prominent in the click recorded from whale no.5. However, the effect of surface reflection should quite rapidly be averaged out, as the inter-pulse interval would vary with slant angle to the whale and hydrophone depth (Fig. 5).

In some cases, however, it may not be possible to average out the effects of surface reflections. The dataset recorded in the North Atlantic in 2000 was from five synchronised

recording platforms, spaced more than 350m apart, where the hydrophone depths ranged 5-30m. Station  $r4$ , in particular, has a less defined peak followed by a larger hump (Fig. 3, top panel). This platform was recording from a 5m deep hydrophone close to being on top of the whale (see fig. 3a in Møhl *et al.*, 2003). This would have resulted in surface reflections, giving inter-pulse intervals close to the nominal IPI. As the receiver is situated right on top of the whale, the inter-pulse interval created by the surface reflection is changing only slowly with time so that this effect is not easily averaged out. It therefore seems important to use longer hydrophone cables or to vary the depth of the hydrophone during the recording to avoid this effect.

Consistent IPIs can be derived from different platforms with different hydrophones in different aspects to the same whale, as shown from the averaged cepstra of two click sequences recorded from different stages of the same dive and from independent platforms (Fig. 3, bottom panel).

### Dataset criteria: how many clicks are needed to measure the IPI?

In this study it has been shown that the nominal IPI can be estimated from the average of cepstra from sequences of clicks. The shape of the averaged cepstral peak depends on the coverage of different recording aspects around the whale. This is due to the fact that a click cepstrum would show any inter-pulse interval contained in the clicks and some of these intervals have been demonstrated to be aspect-dependent (Zimmer *et al.*, 2005a). The recording aspect is given by the small scale movements of the whale combined with the relative aspect of the whale to the receiver. Therefore, it is difficult to estimate a minimum number of clicks to be recorded in order to average out the aspect-dependent peaks and obtain a reliable estimate of the nominal IPI. In addition, the signal-to-noise ratio and the presence of surface reflections may also affect the effectiveness of the method. Likewise, the width of the averaged peak will also depend on the above-mentioned parameters and may therefore not be easy to assess. For the

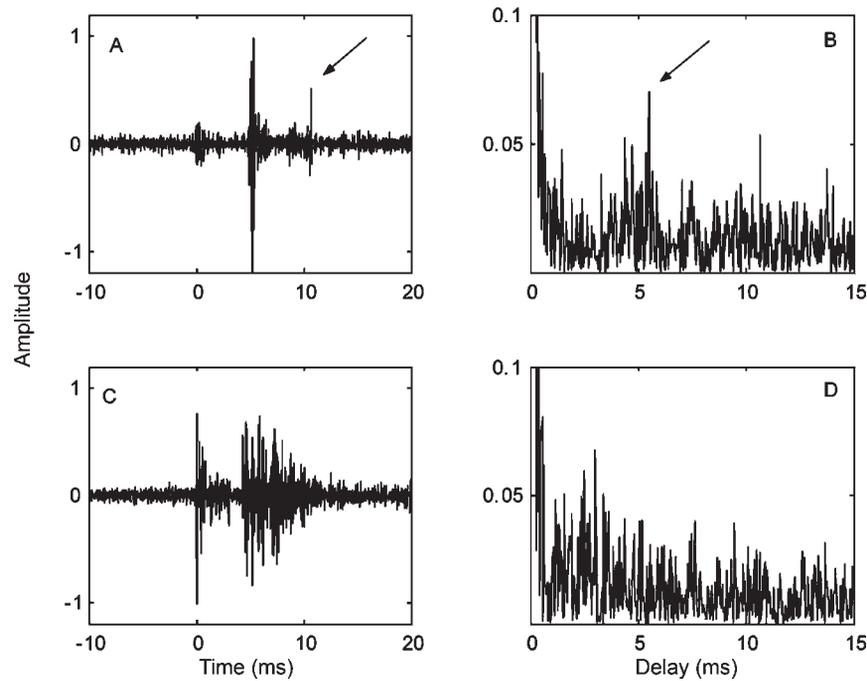


Fig. 6. Waveform (left panels) and corresponding cepstrum (right panels) of an on-axis click (top) and an off-axis click (bottom). Note that no obvious multi-pulse structure is recognisable for the off-axis click. The arrow in the on-axis click shows the  $p_2$  pulse in the waveform and the corresponding peak in the cepstrum. Clicks were selected from whale no.2 ( $n=1,997$ , Table 1). The delay in ms on the  $x$  axis of the right panels is termed quefrency in cepstrum analysis (Bogert *et al.*, 1963).

single hydrophone data sets the width of the cepstral peak (defined at 50% of its maximum) ranges from  $50\mu\text{s}$  to  $100\mu\text{s}$  (Table 1). For the independent measurements of the nominal IPI obtained in various directions from the whale (whales 3 and 4 in Table 1) the maximum IPI deviation is only  $20\mu\text{s}$ .

A complete sperm whale's dive generally contains 1,200-2,000 usual clicks (Madsen *et al.*, 2002; Wahlberg, 2002). Results from high-resolution digital tags recording the whale's movements at a fine scale indicate that sperm whales are consistently moving and rotating while echolocating on prey and considerable depth excursions occur during the bottom phase of a dive (Johnson and Tyack, 2003; Miller *et al.*, 2004; Zimmer *et al.*, 2003). By using all usual clicks from a complete dive, with the whale moving underwater at an unknown distance, bearing and aspect, one can assume that the recorded clicks cover a substantial range of different aspects of the emitting whale in respect to the receiver (Zimmer *et al.*, 2005b).

In general, the shape of the curve differs for different datasets, depending on the number and signal-to-noise ratio of clicks that have been averaged, and different data sets have different convergence rates. The distribution around the peak when increasing the number of clicks (Figs 2 and 4) gives an indication of the quality of the recordings and may be used to judge if the number of clicks is appropriate for the nominal IPI measurement. It should be noted that, according to the theory outlined above, there is no pulse interval longer than the two-way travel time in the spermaceti organ, except for the nominal IPI plus the distance between the monkey lips and junk exit point divided by the speed of sound in water.

The performance of this technique will depend on how quickly the cepstral nominal IPI peak will emerge during averaging and on how accurately the IPI can be derived. Thus, the performance depends on the signal-to-noise ratio and width of the cepstral nominal IPI peak (see above), which will both depend on the signal-to-noise ratio as well

as the bandwidth of the pulses in the click. Depth-dependent effects in the nominal IPI may also influence the width of the peak: as the sperm whale is diving, the sound velocity of the spermaceti oil will change (Goold *et al.*, 1996) and so will the IPI. From the limited data available to date it can be inferred that this effect is less than  $0.2\text{ms}$  for a dive from the surface to a depth of  $1,000\text{m}$  (Madsen *et al.*, 2002; Wahlberg *et al.*, 2003). This translates to an error in length estimation of  $0.6\%$  for IPIs of  $5\text{ms}$  and  $2\%$  for IPIs of  $7\text{ms}$  (Rhineland and Dawson, 2004). The observed smearing of the averaged cepstral peak is therefore possibly an effect of the whale changing its depth while diving.

Variations in the nominal IPI may also occur due to the action of the sound generator. As air is driven past the monkey lips and fills up the distal air sac during a click train, it is expected that the sound path from the frontal and distal air sacs might shorten by a few centimetres, which may cause some  $100\text{s of } \mu\text{s}$  of smearing of the averaged cepstral peak. Smearing of the peak may also occur if most clicks were recorded from an aspect where the dominating aspect-dependent inter-pulse intervals (such as the  $p_0-p_{1/2}$ , or  $p_0-p_1$ ) are close to the nominal IPI throughout the click sequence.

If clicks straight in front of or behind the whale cannot be recorded, a large number of clicks should be used instead. This increases the probability of averaging several aspects of the recorded whale and thereby emphasising the inter-pulse intervals that correspond to the two-way travel time in the spermaceti organ, the nominal IPI. The empirical exploration suggests that 200 to 1,000 clicks seem to be sufficient in the data set used here.

## CONCLUSIONS

Click sequences from different sperm whales have been used to test an automatic and robust method for the acoustic size estimation of the sperm whale nasal complex by

exploiting the modified bent horn model for sperm whale sound production (Møhl *et al.*, 2003; Zimmer *et al.*, 2005a). This method makes use of a click detector followed by cepstrum analysis and therefore does not require visual inspection of individual sperm whale clicks. The nominal IPI is measured from averaged cepstra, as all other inter-pulse interval components are aspect-dependent and thereby reduced through averaging. The method allows measurements largely independent of whale depth, recording equipment or recording aspect to the whale.

Here only usual clicks were used for the analysis, although other types of clicks were present in the recordings. Coda clicks have been used for size estimation of sperm whales (Marcoux *et al.*, 2006; Rendell and Whitehead, 2004), but it is not clear if the production of this type of click follows the same mechanisms as for usual clicks (Madsen *et al.*, 2002). If this is the case, any aspect-dependent inter-pulse intervals that may occur in coda clicks should be taken into account when using codas for size estimation. Moreover, these clicks seem largely emitted by maturing and mature females (Marcoux *et al.*, 2006), which would introduce a bias in application of this technique to male sperm whales, young animals or to populations where coda click are not as abundant as in the Pacific, like in the Mediterranean Sea (Teloni, 2005).

The method presented here offers a reliable tool for acoustic size estimation of the spermaceti organ, which could be further related to the overall size of the whale (Clarke, 1978). We did not calculate the length of the whales as the objective of this study was to test a consistent method for measuring IPIs that the published regressions (Gordon, 1991; Rhineland and Dawson, 2004) are based upon. More data on the relationship between the IPI and total length for a large range of whale sizes are required to apply the acoustic measure of IPIs to population studies, but a consistent method for reliable derivation of the nominal IPI required for this is now available. Long click sequences from diving sperm whales are needed in order to improve the probability that the whale may be recorded from different aspects. Considering that sperm whales emit a usual click on average every second (Whitehead and Weilgart, 1990), it would be necessary to record an animal for approximately fifteen minutes in order to record a thousand clicks. Recordings from a single hydrophone are sufficient for this work as information on the range and aspect of the whale is not required. This method can be applied to acoustic surveys, where long sequences of sperm whale clicks are collected, avoiding time-consuming hand-processing and allowing for size estimation of clicking whales for which the relative orientation is unknown.

A common problem with acoustic size estimation is that of several whales clicking simultaneously. In this case, the technique presented here needs to be augmented with the capability to discriminate between different whales, e.g. by resolving the bearing to the individual whales with two or more hydrophones. Automatic real time detection software (for example *RainbowClick* from [www.ifaw.org](http://www.ifaw.org)) that allow the observer to assign sequences of clicks to individual whales might facilitate the analysis in case of multiple animal recordings.

#### ACKNOWLEDGMENTS

We thank NRV *Alliance*, RV *Odyssey*, RV *Narhvalen* and *Iolaire* crews and B. Møhl, M. Carron, R. Payne, I. Kerr and P. Tyack for their support. M. Johnson kindly provided Dtag data from the Andenes fieldwork and gave helpful

comments to previous versions of the manuscript. B. Møhl and M. Amundin kindly provided the recordings from the Faroe Islands. We thank P. Miller and M. Fage Christoffersen for their contribution to the tagging operation and the Hvalsafari A/S (Andenes) for logistic support. We thank B. Dahl for proof reading. VT was supported by the Office of Naval Research Grant N00014-03-1-0929. MW was supported by the Carlsberg Foundation. PTM was supported by the Danish Natural Science Research Council through a Steno Fellowship. The manuscript benefited greatly from the comments of the anonymous reviewers.

#### REFERENCES

- 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: 2358-61.
- Angliss, R.P., Rugh, D.J., Withrow, D.E. and Hobbs, R.C. 1995. Evaluations of aerial photogrammetric length measurements of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). *Rep. int. Whal. Commn* 45: 313-24.
- Backus, R.H. and Schevill, W.E. 1966. *Physeter* clicks. pp.510-28. In: Norris, K.S. (eds). *Whales, Dolphins and Porpoises*. Univ. Calif. Press, Berkeley. 789pp.
- Bogert, B.P., Healy, M.J.R. and Tukey, J.W. 1963. The quefrency analysis time series for echoes: cepstrum, pseudoautocovariance, cross-cepstrum, and saphe cracking. pp.209-43. In: Rosenblatt, M. (eds). *Symposium on Time Series Analysis*. Wiley, New York.
- Clarke, M.R. 1978. Structure and proportions of the spermaceti organ in the sperm whale. *J. Mar. Biol. Assoc. UK* 58: 1-17.
- Cosens, S. and Anders, B. 2003. Size- and age-class segregation of bowhead whales summering in northern Foxe Basin: a photogrammetric analysis. *Mar. Mammal Sci.* 19(2): 284-96.
- Dawson, S.M., Chessum, C.J., Hunt, P.J. and Slooten, E. 1995. An inexpensive, stereophotographic technique to measure sperm whales from small boats. *Rep. int. Whal. Commn* 45: 431-36.
- Drouot, V., Gannier, A. and Goold, J. 2004. Diving and feeding behaviour of sperm whales (*Physeter macrocephalus*) in the northwestern Mediterranean Sea. *Aquat. Mamm.* 30: 419-26.
- Flewellen, C. and Morris, R.J. 1978. Sound velocity measurements on samples from the spermaceti organ of the sperm whale (*Physeter catodon*). *Deep Sea Research* 25: 269-77.
- Goold, J. 1996. Signal processing techniques for acoustic measurement of sperm whale body lengths. *J. Acoust. Soc. Am.* 100: 3431-41.
- Goold, J., Bennel, J.D. and Jones, S.E. 1996. Sound velocity measurements in spermaceti oil under the combined influences of temperature and pressure. *Deep Sea Research* 43: 961-69.
- Goold, J.C. and Jones, S.E. 1995. Time and frequency domain characteristics of sperm whale clicks. *J. Acoust. Soc. Am.* 98(3): 1,279-91.
- Gordon, J.C.D. 1987. The behaviour and ecology of sperm whales off Sri Lanka. PhD thesis, Darwin College, University of Cambridge. 347pp.
- Gordon, J.C.D. 1990. A simple photographic technique for measuring the length of whales from boats at sea. *Rep. int. Whal. Commn* 40: 581-88.
- Gordon, J.C.D. 1991. Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations. *J. Zool. (Lond.)* 224: 301-14.
- Jaquet, N. 2006. A simple photogrammetric technique to measure sperm whales at sea. *Mar. Mammal Sci.* 22: 862-79.
- Jaquet, N., Dawson, S.M. and Douglas, L. 2001. Vocal behaviour of male sperm whales: why do they click? *J. Acoust. Soc. Am.* 109: 2254-59.
- Johnson, M. and Tyack, P.L. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE J. Ocean. Engr.* 28: 3-12.
- Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1993. Reproduction. pp.239-74. In: Burns, J.J., Montague, J.J. and Cowles, C.J. (eds). *The Bowhead Whale. Special Publication No.2*. The Society for Marine Mammalogy, Lawrence, Kansas. 787pp.
- Leaper, R., Chappell, O. and Gordon, J. 1992. The development of practical techniques for surveying sperm whale populations acoustically. *Rep. int. Whal. Commn* 42: 549-60.
- Madsen, P.T., Carder, D., Au, W.W., Nachtigall, P., Møhl, B. and Ridgway, S. 2003. Sound production in neonate sperm whales. *J. Acoust. Soc. Am.* 113: 2988-91.

- Madsen, P.T., Payne, R., Kristiansen, N.U., Wahlberg, M., Kerr, I. and Møhl, B. 2002. Sperm whale sound production studied with ultrasound time/depth recording tags. *J. Exp. Biol.* 205: 1899-906.
- Marcoux, M., Whitehead, H. and Rendell, L. 2006. Coda vocalisations recorded in breeding areas are almost entirely produced by mature female sperm whales (*Physeter macrocephalus*). *Can. J. Zool.* 84(4): 609-14.
- Miller, P.J.O., Johnson, M.P., Tyack, P. and Terray, E.A. 2004. Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus*. *J. Exp. Biol.* 207: 1953-67.
- Møhl, B. 2001. Sound transmission in the nose of the sperm whale *Physeter catodon*. A post mortem study. *J. Comp. Physiol.* 187: 335-40.
- Møhl, B., Larsen, E. and Amundin, M. 1981. Sperm whale size determination: outline of an acoustic approach. *FAO Fisheries Series No. 5 (Mammals in the Seas)* 3: 327-32.
- 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: 1143-54.
- Norris, K.S. and Harvey, G.W. 1972. A theory of the function of the spermaceti organ of the sperm whale (*Physeter catodon*). pp.397-417. In: Galler, S.R., Schmidt-Koering, K., Jacobs, G.J. and Belleville, R.E. (eds). *Animal Orientation and Navigation*. NASA, Washington DC. Special Publication 262.
- Oppenheim, A.V., Schafer, R.W. and Stockham, T.G. 1968. Nonlinear filtering of multiplied and convolved signals. *Proc. IEEE* 56: 1264-91.
- Papastavrou, V., Smith, S.C. and Whitehead, H. 1989. Diving behaviour of the sperm whale, *Physeter macrocephalus*, off the Galapagos Islands. *Can. J. Zool.* 67(4): 839-46.
- Pavan, G., Hayward, T.J., Borsani, J.F., Priano, M., Manghi, M., Fossati, C. and Gordon, J. 2000. Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985-1996. *J. Acoust. Soc. Am.* 107(6): 3487-95.
- Pavan, G., Priano, M., Manghi, M. and Fossati, C. 1997. Software tools for real-time IPI measurements on sperm whale sounds. *Proceedings IOA, Loughborough, UK* 19: 157-64.
- Rendell, L. and Whitehead, H. 2004. Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurements. *Anim. Behav.* 67: 865-74.
- Rhineland, 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: 1826-31.
- Teloni, V. 2005. Patterns of sound production in diving sperm whales in the northwestern Mediterranean. *Mar. Mammal Sci.* 21: 446-57.
- Teloni, V., Zimmer, W.M.X., Fossati, C., Manghi, M., Pavan, G. and Priano, M. 2000. Variability of temporal and spectral click characteristics of sperm whales (*Physeter macrocephalus*). *Eur. Res. Cet.* 14: 91-95.
- Teloni, V., Zimmer, W.X.M., Johnson, M. and Tyack, P.L. 2003. Inter click interval variation of deep-diving sperm whales in the Mediterranean Sea. *European Cetacean Society Conference Guide and Abstracts* 17: 84-85.
- Wahlberg, M. 2002. The acoustic behaviour of diving sperm whales observed with a hydrophone array. *J. Exp. Mar. Bio. Eco.* 281: 53-62.
- Wahlberg, M., Madsen, P.T., Møhl, B. and Payne, R. 2003. Acoustic size estimation of sperm whales: effects of directionality and pressure. 1st Conference on Acoustic Communication by Animals, 27-30 July 2003, Maryland.
- Watkins, W.A., Daher, M.A., DiMarzio, N.A., Samuels, A., Wartzok, D., Fristrup, K.M., Gannon, D.P., Howey, P.W., Maiefski, R.R. and Spradlin, T.R. 1999. Sperm whale surface activity from tracking by radio and satellite tags. *Mar. Mammal Sci.* 15(4): 1,158-80.
- Whitehead, H. and Weilgart, L. 1990. Click rates from sperm whales. *J. Acoust. Soc. Am.* 87(4): 1798-806.
- Zimmer, W.M.X., Johnson, M.P., D'Amico, A. and Tyack, P. 2003. Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*). *IEEE J. Ocean. Engr.* 28: 13-28.
- Zimmer, W.M.X., Madsen, P.T., Teloni, V., Johnson, M.P. and Tyack, P. 2005a. Off-axis effects on the multi-pulse structure of sperm whale usual clicks with implications for the sound production. *J. Acoust. Soc. Am.* 118: 3337-45.
- Zimmer, W.M.X., Tyack, P., 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: 1473-85.

Date received: March 2006

Date accepted: March 2007

## Appendix A

```

% Teloni, Zimmer, Wahlberg, Madsen © 2005
% Sperm whales acoustic size estimation
%next three lines to be adjusted
disk='disk name\';
detFile=[disk 'click detect\click detector file.txt'];
wavFile=[disk 'wav\audio file.wav'];
%
%load detection file
a=load(detFile);
tcl=a(:,1); %first column has click times
% sort times in case there are out of order
[Y,I]=sort(tcl);
%
% get click-time and inter click interval
tcl=Y;
ici=diff(tcl);
%retrieve sampling frequency fs
[x,fs]=wavread(wavFile,1);
%extract click and store in big matrix
sta=0.01; %interval before detection [s]
sto=0.04; %length of data to be extracted [s]
off=0; %offset for click time [s]
xx=[]; %matrix where to store click data
nd=floor(sto*fs); %number of samples to extract
%
ncl=length(tcl); %tot number of clicks
for ii=1:ncl
    t=tcl(ii)+off;
    N1=max(1,floor((t-sta)*fs));
    N2=N1+nd-1;
    x=wavread(wavFile,[N1 N2]);
    if length(xx)==0
        xx=zeros(nd,ncl); %allocate storage to speed-up execution
    end
    xx(:,ii)=x(:,1); %store data (here first channel of multi-channel dataset)
end
%
%
Tc=(0:nd-1)/fs*1000; %elapsed time [ms]
C=abs(ifft(log(abs(fft(xx))))); %cepstrum
Z=mean(C,2); %averaged cepstrum
plot(Tc,Z)
ylim([0 0.1])
xlim([0 20])

```