Available online at www.sciencedirect.com
ScienceDirect
Deep-Sea Research I 54 (2007) 1421–1444
DEEP-SEA RESEARCH PART I
www.elsevier.com/locate/dsri

Instruments and Methods

# Recording and quantification of ultrasonic echolocation clicks from free-ranging toothed whales

P.T. Madsen[a,b,*], M. Wahlberg[a,c]

[a]*Zoophysiology, Department of Biological Sciences, Build. 1131, University of Aarhus, Denmark*
[b]*Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA*
[c]*Fjord & Bælt, Margrethes Plads 1, 5300 Kerteminde, Denmark*

Received 8 October 2006; received in revised form 15 March 2007; accepted 2 April 2007
Available online 18 May 2007

**Abstract**

Toothed whales produce short, ultrasonic clicks of high directionality and source level to probe their environment acoustically. This process, termed echolocation, is to a large part governed by the properties of the emitted clicks. Therefore derivation of click source parameters from free-ranging animals is of increasing importance to understand both how toothed whales use echolocation in the wild and how they may be monitored acoustically. This paper addresses how source parameters can be derived from free-ranging toothed whales in the wild using calibrated multi-hydrophone arrays and digital recorders. We outline the properties required of hydrophones, amplifiers and analog to digital converters, and discuss the problems of recording echolocation clicks on the axis of a directional sound beam. For accurate localization the hydrophone array apertures must be adapted and scaled to the behavior of, and the range to, the clicking animal, and precise information on hydrophone locations is critical. We provide examples of localization routines and outline sources of error that lead to uncertainties in localizing clicking animals in time and space. Furthermore we explore approaches to time series analysis of discrete versions of toothed whale clicks that are meaningful in a biosonar context.
© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Toothed whale; Ultrasound; Recording; Click; Hydrophone; Array; Echolocation

## 1. Introduction

Echolocating toothed whales emit ultrasonic clicks to acquire information about their environment and to find food by reception and analysis of echoes returning from ensonified objects in the water column and the sea bed (Au, 1993). The performance of a biosonar system is dictated partly by the source properties of the emitted clicks (Au, 1997, 2004) in that high-amplitude signals may ensonify more distant targets and a higher directionality (Fig. 1A) reduces the number of unwanted echoes (also called clutter). The temporal and spectral properties of the clicks determine the information that can be derived from returning echoes and how well the echoes can be detected in noise and clutter. The ability to resolve the location of a target in time and space as well as its size, shape and material follow partly from the properties of the signal waveform (Au, 1993). A broad bandwidth of

*Corresponding author. Zoophysiology, Department of Biological Sciences, Build. 1131, University of Aarhus, Denmark.
*E-mail address:* peter.madsen@biology.au.dk (P.T. Madsen).

0967-0637/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr.2007.04.020

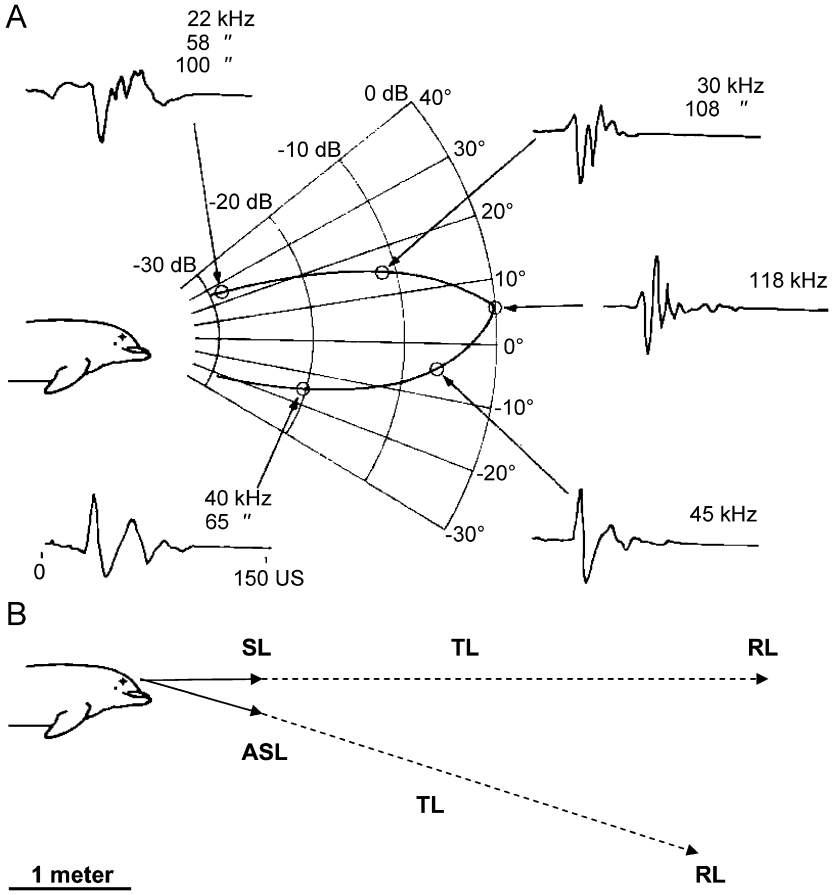


Fig. 1. (A) Radiation pattern of a click from an echolocating dolphin (Au, 1993) showing off-axis distortion. (B) The passive sonar equation applied to a clicking dolphin. SL is the back-calculated sound pressure level (Received level (RL)+transmission loss (TL)) 1 m from the source and on the acoustic axis. ASL is the sound pressure back-calculated to 1 m in an unknown aspect to the animal.

the source signal improves these abilities (Brill et al., 1992; Roitblat et al., 1995), whereas the detectability in broad band noise improves with diminishing band width (Madsen et al., 2005b). The detection range increases with the transmitting and receiving directionality, which in turn depends on the dominant wavelengths of the sonar signal relative to the transmitting aperture. Shorter wavelengths emitted from a given transmitting aperture increase the directionality on both the transmitting and receiving side. Higher frequencies reduce the incoming noise levels with masking potential (Wenz, 1962), but will also lead to a higher absorption of sound energy in the medium (Urick, 1983). Thus, there is a trade off in terms of the range, noise levels, clutter reduction and resolution for a given biosonar as dictated by the transmitted clicks, the receiving directionality and the sound producing structures of the echolocating animal emitting them. For example, the source properties of sperm whale echolocation clicks with very high-source levels (SL) at 15–20 kHz show that these signals hold a strong potential to operate in a long range biosonar system (Madsen et al., 2002b; Moehl et al., 2003), whereas the low amplitude, 130 kHz clicks from small toothed whales such as harbor porpoises and pygmy sperm whales preclude anything but use in short range biosonar systems (Au et al., 1999; Madsen et al., 2005b). Source properties of echolocation clicks are also relevant for passive acoustic monitoring with towed arrays (Barlow and Taylor, 2005; Leaper et al., 2002), sonobuoys (Levenson, 1974) and automated porpoise detectors (e.g. TPODS, Carlstroem, 2005) for optimizing detection routines and classification of species on the basis of acoustic cues.

Echolocating animals point the sound beam on the target of interest (Au, 1993), so for a meaningful

evaluation of the consequences for performance of a biosonar system, it is of paramount importance that the source properties are derived for clicks on or close to the acoustic axis of the sound beam (Fig. 1A). It is not only the acoustic intensity that differs greatly with the recording aspect relative to the animal, but also the temporal and frequency properties of the signal changes (Fig. 1A). Clicks recorded on the acoustic axis of the transmitting aperture are generally of broad bandwidth and short duration. As the recording aspect to the acoustic axis increases, the frequency content decreases, as the high-frequency components are more directional than the low-frequency ones. Sound is not radiated from a point source in the toothed whale nasal complex, but from an extended aperture of many small sources, which means that the duration of the click increases, as sound radiated by various parts of the nasal complex arrives with increasing delays off the acoustic axis. For some species of toothed whales, such as porpoises, dwarf and pygmy sperm whales and dolphins of the genus *Cephalorhynchus* producing narrow band high-frequency signals, these distorting effects are small (Au et al., 1999), because of the narrower bandwidth.

For on-axis measurements in captivity the animal can be trained to hold station in a hoop or on a bite plate and transmit clicks towards an array of hydrophones in front of it (Au et al., 1974). Such investigations have provided critical information on the performance of toothed whale biosonar, and have formed the basis for formulating meaningful parameters and models for emission and use of toothed whale sonar signals (Au, 1993). However, as shown for bats (Surlykke and Moss, 2000), it can be questioned if data on the transmission system from trained animals studied in captivity are representative of the signals free-ranging animals produce while using their biosonar for orientation and food finding in natural habitats (Au, 1993; Au et al., 2004; Au and Herzing, 2003; Madsen et al., 2004a, b). With the exception of rare open water experiments (e.g. Au et al., 1974; Murchison, 1980), animals in captivity are normally recorded in small tanks where their sonar signals are of much lower amplitude than those recorded in the field, and with a much lower frequency emphasis and therefore directionality (Au, 1993). Secondly, housing and training of captive toothed whales are costly, rendering derivation of source properties of clicks from all toothed whale species by means of capture and training impractical at best. Accordingly, there are good arguments for quantifying toothed whale sonar signals in the wild in habitats and in behavioural settings such as foraging for which the biosonar systems have evolved.

Field recordings are logistically challenging and must often take place in changing weather conditions and in the presence of interference from other sound sources or with more than one vocally active species present. Moreover, estimation of source parameters of directional, high-pressure transients, like toothed whale clicks, from moving sources, requires multiple receivers connected to calibrated recorders with sufficient bandwidth and dynamic range to handle such signals.

The use of hydrophone arrays for bioacoustic research at sea was pioneered in the 1960s and 1970s by W. Watkins and W. Schevill who recorded a large number of cetacean and pinniped species (Watkins and Schevill, 1972), but employed little quantitative analysis on the recordings. Later on, Moehl et al. (1990) used a deep vertical array of hydrophones to localize echolocating narwhals in the wild and to estimate source parameters of their clicks. This study documented for the first time that free-ranging toothed whales could generate source sound pressure levels just as high as or higher than documented for trained animals engaged in long-range target detection experiments (Au et al., 1974). It was also demonstrated that foraging free-ranging toothed whales, in analogy with echolocating bats, produce high-repetition buzzes during the final stages of prey capture (Miller et al., 1995).

Up until recent years, multi-channel recorders with sufficient bandwidths were available only in the form of analog tape recorders (Diercks et al., 1973; Weber, 1963). Unfortunately, high-speed analog recording systems have limited dynamic range ($<35$ dB) and are large, expensive and cumbersome to handle. In addition, analysis of recordings stored on magnetic tape is hampered either by the time-consuming processes of analog signals analysis or by subsequent less-than-real time digitization of the analog signals (see Watkins and Daher, 1991). The development of high-speed digital recorders and PC-based signal processing of discrete signals have made recording and derivation of source properties of toothed whale clicks much less expensive and less cumbersome, and therefore accessible to a larger community of researchers. The use of digital recording gear to quantify source properties of biosonar sound sources underwater is nevertheless

not straightforward. Researchers at sea are faced with a number of potential pitfalls and technical challenges in concert with questions of how to approach the analysis of discrete versions of directional, ultrasonic transients.

Here, we provide a technical overview and a critical evaluation of methods to record and quantify ultrasonic echolocation clicks from free-ranging toothed whales at sea using digital, multi-channel recording systems. We outline array configurations, specifications, and calibration of hydrophones and amplifiers suited for recording and localization of ultrasonic transients from toothed whales, and we discuss signal processing approaches to provide quantitative parameters of discrete versions of toothed whale clicks. We address common problems and provide recommendations that will hopefully facilitate the use of experimentally comparable approaches to bioacoustics, biosonar research and passive acoustic monitoring.

## 2. Recording and localization

### 2.1. Background

Sound consists of particle motions that create a pressure wave propagating away from the sound source at a sound speed determined by the properties of the medium. A sound field is therefore made up by a particle velocity component ($v$) and a pressure component ($p$), and their product defines the acoustic intensity ($I$):

$$I = p \cdot v.$$

The particle velocity is given by the pressure divided by the acoustic impedance of the medium. In the acoustic-free field, far from the sound source and any reflecting boundaries, the acoustic impedance is the product of the sound velocity $c$ and the density of the medium ($\rho$). Under those circumstances the sound intensity may therefore be calculated as

$$I = p^2/(\rho c).$$

Usually, we express the sound intensity in decibel units:

$$10\log_{10}(I/I_0),$$

where $I_0$ is the intensity of a plane wave with an rms sound pressure of 1 μPa in water.

Accordingly, we quantify the sound pressure as

$$\text{dB re } \mu\text{Pa} = 20\log_{10}(p/p_0),$$

where $p_0$ is reference sound pressure of 1 μPa quantified in the same way as the measured pressure $p$. The pressure wave amplitude is normally decreasing in magnitude as the sound pulse propagates away from the source. The transmission loss (TL) is defined as the ratio between the acoustic intensity 1 m in front of the animal ($I_{1\,\mathrm{m}}$) to the acoustic intensity at a distance $r$ from the animal ($I_r$). By transforming to logarithmic units this ratio becomes a difference instead (Urick, 1983):

$$\mathrm{TL} = 10\log_{10}(I_r/I_{1\,\mathrm{m}}) = 20\log_{10}(P_r/P_{1\,\mathrm{m}}) = \mathrm{SL} - \mathrm{RL}.$$

SL is the source level, or the acoustic intensity measured at, or back-calculated to, 1 m on the acoustic axis of the animal in decibel units, $20\log_{10}(p_{1\,\mathrm{m}}/p_0)$, where $p_{1\,\mathrm{m}}$ is the acoustic pressure 1 m from the source, and $p_0$ is a reference pressure of 1 μPa in water. The received level (RL) is defined as $20\log_{10}(p_r/p_0)$, where $p_r$ is the acoustic pressure at a distance $r$ from the source. RL is normally the sound pressure that is impinging on the recording hydrophone. Thus, if the TL is known, the SL can be estimated from measurements of the RL (Urick, 1983):

$$\mathrm{SL} = \mathrm{RL} + \mathrm{TL}.$$

The apparent simplicity of this expression is seductive as it contains two terms that must be interpreted with great care. First, the received level must be measured in an unambiguous way with relevance for the hearing/sonar system of the animal in question. Secondly, the TL must be known. In its most basic form the TL consists of two parts: geometric spreading and absorption. Geometric spreading is caused by the sound energy being distributed over an expanding surface analogous to the ever-expanding ring of a wave created by dropping a stone into the water. In a boundary-free, iso-velocity medium, the sound wave is expanding with the distance $r$ to the sound source as if on the surface of a sphere. The surface area of the sphere is increasing by $r^2$, and the acoustic intensity is thereby reduced proportional to $r^{-2}$. This is the so-called spherical spreading (or inverse square loss) where the sound intensity in decibel units decreases by $20\log_{10}(r)$. If the sound energy is channeled between two reflective surfaces that are close together relative to the duration of the sound pulse, the sound wave is expanding as cylindrical spreading by which the acoustic intensity decreases by $10\log_{10}(r)$. In many real situations, the geometric

transmission loss will be in-between cylindrical and spherical for long or continuous sounds (Richardson et al., 1995). However, from a geometric calculation for ranges and bottom depths of normal interest (up to 100 m range and more than 1 m away from the surface and the bottom), the spherical spreading is a reasonable approximation for the short-duration and very directional toothed whale signals (Fig. 2, Villadsgaard et al., 2007). Nevertheless, measurements of the TL of toothed whale signals transmitted with the same directionality as the animals should be conducted in the habitat in question to test this assertion, especially when recording in very shallow water (relevant for some dolphin species) or at long ranges (>1 km) of e.g. sperm whale echolocation clicks.

Absorption is a complicated process caused by molecular interactions induced by the pressure fluctuations. It may be looked at as sound being lost by friction, e.g. sound energy being transformed into heat. Absorption ($\alpha$) is highly frequency and temperature dependant, and in the frequency range of interest for toothed whale biosonars it can be approximated by (Au, 1993)

$$\alpha = ABf^2/(B^2+f^2)\,\mathrm{dB/m},$$

where $f$ is the frequency (in Hz) and the other parameters are given by $A = 48.83\times10^{-8} + 65.34\times10^{-10}T\,\mathrm{s/m}$, $B = 1.55\times10^{7}(T+273.1)\exp(-3052/(T+273.1))\,\mathrm{Hz}$ and $T$ is the temperature in degrees Celsius.

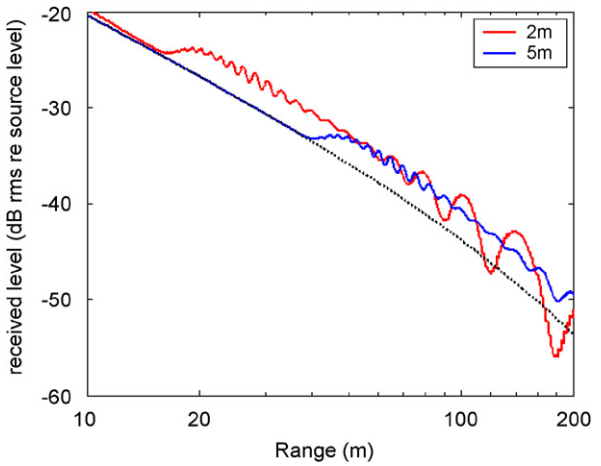


Fig. 2. Modeling the received level of a harbor porpoise click emitted at 1 m depth and received at 5 m depth at different ranges. Bottom depth is 10 m. The dotted line indicates a transmission loss according to spherical spreading and absorption of 38 dB/km. The model does not take into account that the signal is directional, so the effect of surface interference is exaggerated.

For broadband signals, higher frequencies will attenuate faster than lower ones. Thus, the transmission will low-pass filter the signal and thereby change its temporal and spectral structure. This is important to consider when measuring toothed whale signals at long ranges. Consequently, absorption can be ignored for sperm whale clicks at 15 kHz recorded at 500 m (<1 dB absorption), but must be considered for a porpoise click at 130 kHz at the same (>20 dB absorption at 500 m) and much shorter ranges.

Combining geometric spreading and attenuation, the TL may be quantified as

$$\mathrm{TL} = 20\,\log(r) + \alpha r.$$

This formula assumes spherical spreading which is applicable within a certain range depending on the signal duration and depth of the hydrophone, the sea floor and the clicking animal. On top of this, the transmission loss is affected by several other processes. A varying sound velocity of the medium will cause the sound waves to bend, very much like optical rays through a lens. Such refraction may significantly affect the TL at larger ranges (Medwin and Clay, 1998), but for distances of interest here (normally less than 100 m, except for sperm whales) the problem is small, when keeping the hydrophones away from the surface. Sound will also be scattered by objects in the medium that are large enough relative to the dominant wavelengths of the sound to provide efficient backscatter.

### 2.2. Recording gear

A recording chain for underwater recordings normally consists of a hydrophone, a preamplifier and a band-pass filter connected to a digital recorder (Fig. 3). The hydrophone is normally a piezoelectric ceramic pressure to voltage transducer that generates a given voltage (V) per Pascal (Pa) of sound pressure that is impinging on it (Levin, 1973). Given the large dynamic range of received sound pressures and the general use of the dB scale, hydrophone sensitivity is specified in dB re 1 V per μPa. For example, a hydrophone with a sensitivity of −180 dB re 1 V per μPa provides 1 nV per 1 μPa of sound pressure impinging on it ($20\log_{10}(10^{-9}\,\mathrm{V}/1\,\mathrm{V}) = -180\,\mathrm{dB}$ re 1 V), and a hydrophone with a sensitivity of −200 dB re 1 V per μPa would be 20 dB (10 times) less sensitive. Another way of looking at

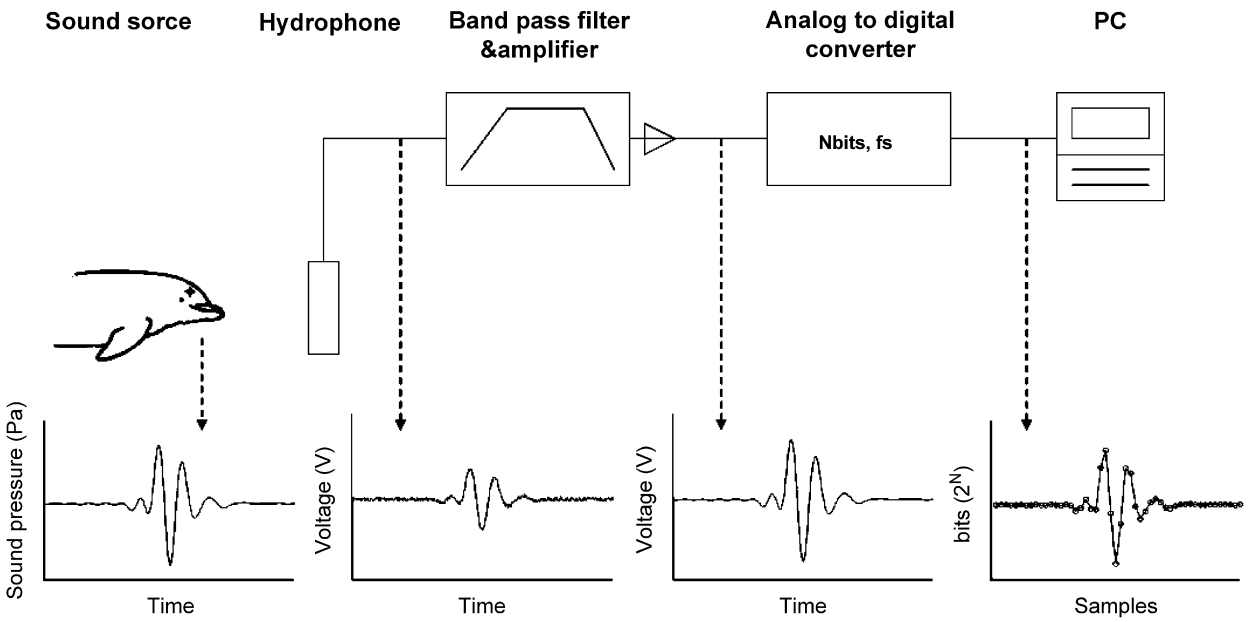


Fig. 3. Clicking dolphin and a recording chain with a hydrophone, an amplifier and an analog-to-digital-converter (ADC). The dolphin produces an analog pressure waveform that is converted to an analog voltage waveform by the hydrophone. The amplifier unit filters and amplifies the analog voltage input before digitization in the ADC that forms a discrete version of the click.

it is simply that the hydrophones will generate 1 V if exposed to 180 and 200 dB re 1 μPa, respectively. Hence, an output of 0.5 V (−6 dB re 1 V) from a hydrophone with a sensitivity of −180 dB re 1 V per μPa means that a sound pressure of 174 dB re 1 μPa (180 dB re 1 μPa −6 dB) is impinging on it. Thus, the sound pressure level, $z$, impinging on a hydrophone is simply given by 20 times the logarithm to the base 10 of the voltage, $y$, plus the sound pressure level, $x$, it takes to generate 1 V in the hydrophone, minus any gain (in dB) used:

$$z\,\text{dB re 1 μPa} = |x\,\text{dB re 1 V/μPa}| + 20\log_{10}(y\,\text{V}/1\,\text{V}) - \text{Gain}.$$

The sensitivity of hydrophones is frequency dependant with a maximum sensitivity at the resonance frequency of the piezoelectric ceramic element. Generally for good hydrophones, the frequency response is acceptably flat ($<\pm 2$ dB) below the resonance frequency, whereas it falls off rapidly at higher frequencies (Fig. 4a). It is therefore important to use a hydrophone that has a resonance frequency well above the frequency range of interest (Fig. 4a). For recording ultrasonic clicks of toothed whales, this calls for hydrophones with small piezoceramic elements, which in turn are quite insensitive (less voltage per μPa). Another argument for using a small or spherical hydrophone element is that non-spherical hydrophones become directional when the wavelength of the sounds impinging on it is comparable to or smaller than the physical dimensions of the hydrophone element (Fig. 4b,c). By using hydrophones with small spherical or cylindrical piezoceramic elements their receiving beam pattern is close to omnidirectional also at high frequencies (Fig. 4b,c), and can therefore be used to measure ultrasonic transients from toothed whales without correcting for directional sensitivity.

The low sensitivity of small piezoceramic elements can to some degree be countered by introducing a preamplifier next to the element as an integrated part of the hydrophone. This means that the hydrophone effectively becomes more sensitive, and that the output can be driven in long cables without having the capacitance of the element and the resistance of the cable affect the sensitivity and the frequency response of the hydrophone. The use of small hydrophone elements means that the resulting signal-to-noise ratio (SNR) from the preamplifier will be lower, and therefore such hydrophones cannot be used to record natural ambient noise levels at low-wind speeds (Wenz, 1962). Thus, there is a trade off between choosing small hydrophone elements for obtaining a sufficient flat recording bandwidth with omnidirectional

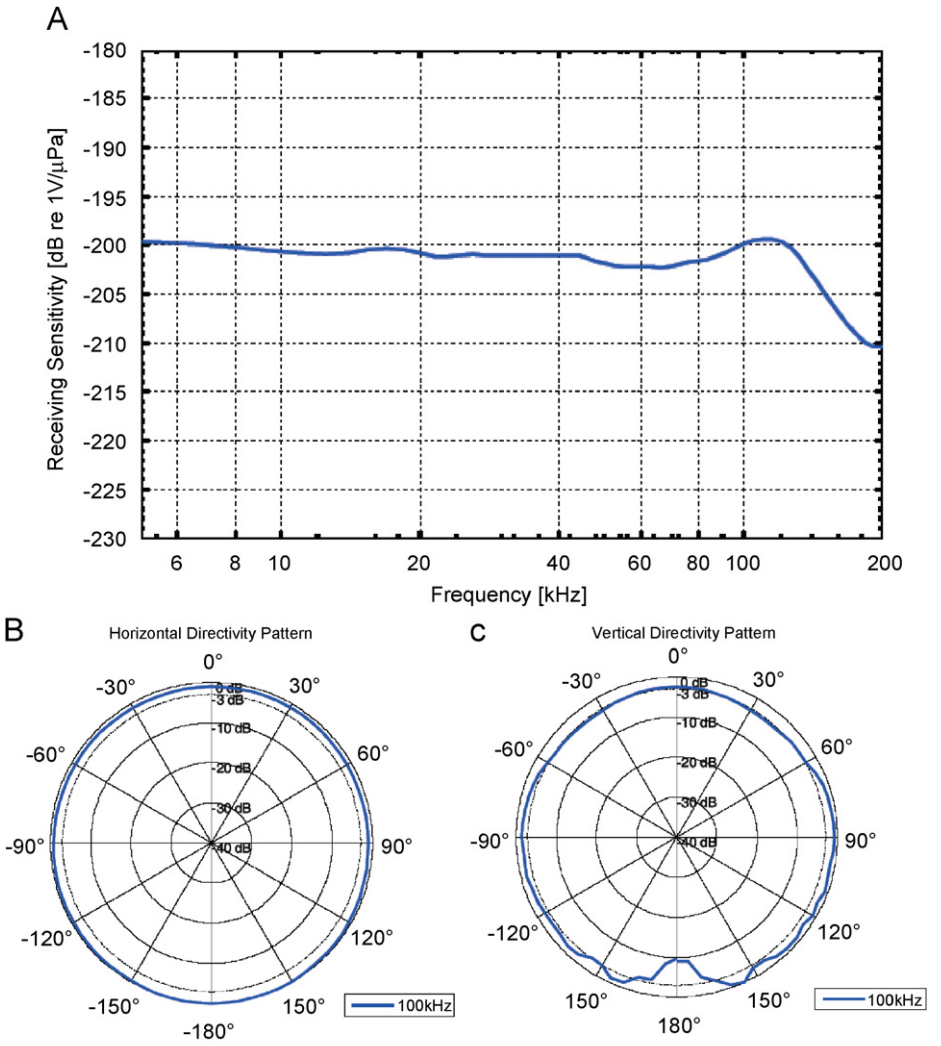


Fig. 4. (A) Sensitivity of a spherical hydrophone (Reson TC4043) as a function of frequency. Note that it is reasonably flat up to the resonance frequency of the element around 120 kHz. (B) Receiving directivity in the horizontal plane of the same hydrophone. (C) Receiving directivity in the vertical plane of the same hydrophone. Courtesy of Reson (www.reson.com).

sensitivity, and at the same time maintaining an adequate SNR in the recordings.

Hydrophones should come with specification charts stating their sensitivity as a function of frequency and direction (Fig. 4A). However, with or without calibration charts it is good practice to check the overall sensitivity through calibration ideally prior to and after every recording. Recording of a calibration signal through the entire recording chain and generation of a wave file with a known sound pressure level is reassuring for subsequent derivation of absolute sound measures of toothed whale clicks. Calibration can be done with a pistonphone, by insert voltage calibration, relative to a hydrophone of known sensitivity or by reciprocity calibration (Urick, 1983). It should be done in a way that ensures that the hydrophone is calibrated in the frequency range of interest.

We will not deal with analog recorders here (see Weber, 1963; Evans, 1973; Watkins 1980; Watkins and Daher, 1992 for use of analog tape recorders), but focus on digital recordings of high-sampling rates. A digital recorder is based on an analog to digital converter (ADC) that creates a discrete time series from an analog voltage input (Fig. 3). The maximum frequency, the so-called Nyquist frequency, at which a signal can be unambiguously described requires at least 2 samples per

pressure/voltage sinosoidal cycle (the Nyquist sampling theorem), and theoretically any frequency below the Nyquist rate can be represented unambiguously. For reasons listed below, it is prudent to sample fast enough that the Nyquist frequency ($f_s/2$, where $f_s$ is the sampling frequency) is at least 1.5 times higher than the highest frequency of interest.

The effective dynamic range of a recording system provides the range of resolved amplitudes that can be covered between the system noise floor and saturation (clipping). The noise floor is defined by the combination of the noise contributions from each active and passive component in the recording system from the hydrophone up to and including the ADC, and can even include quantization noise from possible audio compression on the digital side. Usually, however, the noise floor is dominated by one contribution, almost invariably, either the preamplifier immediately following the hydrophone or the ADC, when dealing with insensitive hydrophones and amplifiers set to handle the high-pressure transient clicks from toothed whales. Clipping occurs whenever the signal is too large to be represented by the voltage or digital numbers available. It can occur anywhere in the recording circuit but is most likely to happen in the ADC. The peak-to-peak dynamic range of an ADC (i.e. the ratio of the largest signal that can be represented without clipping to the maximum quantization noise) is theoretically given by $2^N$, where $N$ is the number of bits (*n*bits in Fig. 3). For a sine-wave input, the RMS dynamic range is $2^N + 3.5$ dB. Each bit therefore provides 6 dB of dynamic range. For example, an 8-bit ADC provides 256 quantization steps ($2^8$, 48 dB pk–pk dynamic range), whereas a 16-bit ADC provides 65 376 points ($2^{16}$ or 96 dB). In reality, no ADCs will achieve their theoretical dynamic range, and performance could be 10 dB or worse depending on other noise sources in the ADC, aliased noise due to inadequate anti-alias filtering, as well as noise coupled through power supplies and ground circuit. Nonetheless, an ADC with more bits will generally provide a larger dynamic range, which will allow the recording system to handle larger fluctuations in the received sound pressure levels without adjusting the gain settings, provided that the dynamic range is not limited by other parts of the recording chain. The price to pay for more than 8 bits is that the size of the wave file is doubled for the same sampling rate up to and including 16 bits. For example 500 ksamples/s at 8 bits generates 0.5 Mb/s per channel, whereas a 12, 14 or 16 bit ADC will generate 1 Mb/s at the same sampling rate (in the computer, 10, 12 or 14 bits will be stored as a 16-bit number) per channel. There are 24 bits and higher ADCs available, but to our knowledge these rarely offer more than 100 dB of dynamic range, i.e. an effective number of bits of 16–17 and there is little advantage in preserving the superfluous bits in most situations, while the amount of data that has to be written to a hard drive becomes very big. If the gain setting in the recording chain is adjusted properly, it is our experience that the 80–90 dB of real dynamic range of a 16-bit ADC can handle the range, source and directionality induced fluctuations in received levels of toothed whale clicks effectively. Received levels of toothed whale echolocation clicks may in some cases exceed 200 dB re μPa (pp) which means that it is not possible to use the same 12- or 16-bit ADC with the same gain settings to make simultaneous recordings of the low ambient noise levels at ultrasonic frequencies as measured by Wenz (1962).

It is of paramount importance to measure the dynamic range of the entire recording chain under realistic conditions (e.g. with partly discharged batteries, etc.) because limitations in the dynamic range that go undetected can lead to erroneous conclusions on the dynamics and properties of the toothed whale sound generator. For example, the received peak-to-peak amplitudes of clicks recorded from an approaching toothed whale might appear to be independent range of the animal if there is clipping anywhere in the recorder. Calculating the SL of the animal from these measurements could lead to the interpretation that the animal was reducing its SL as it approached the hydrophone array by 20 log(range), so-called automatic gain control (see Au and Benoit-Bird, 2003 who did not suffer from this problem (W. Au pers. comm.)), even if, in reality, the SL was invariant. It should be noted that it is not sufficient to look for +/− full-scale numbers in the ADC output to detect clipping although this is always a good thing to do as well. If clipping occurs before the ADC, e.g. in a pre-amplification stage, the flattened peaks of the signal will become rounded after passage through the anti-aliasing filter and so may appear as more-or-less normal clicks when digitized.

Localization of the sound source with multiple receivers requires synchronized sampling of a number of sound channels in the ADC. This can be achieved either by sharing a single high-speed

ADC among the channels (a process known as multiplexing), through a sample-and-hold amplifier-ADC device; or by having multiple ADCs, each dedicated to a single recording channel. The problem with a multiplexed file from $N$ channels is that there will be a phase delay of (1/overall sampling rate) $\times N$ s between the channels that must be compensated for when looking at time-of-arrival differences (TOADs) of the same click on different channels.

Depending on the hydrophone sensitivity and the voltage clipping level of the ADC, it is often necessary to amplify the hydrophone output before digitization. This is normally done in an external conditioning box with variable gain and filter settings. If the hydrophone is without a built-in preamplifier, the conditioning box should have a suitable high-input impedance to avoid unwanted high-pass filtering of the signal and a low-output impedance to avoid attenuation when coupled to the ADC. Before digitization, a band-pass filter is necessary. Low frequency ambient noise must be reduced to avoid saturating the recording system and if signal components above the Nyquist frequency ($f_s/2$) are not filtered out, energy at frequencies above the Nyquist frequency will be folded back during digitization creating aliasing ambiguity and noise. For recording of ultrasonic clicks from toothed whales, a high-pass filter at 1 kHz is often suitable. If there is no preamplifier built in to the hydrophone, a one-pole (−6 dB/octave) high-pass filter can be made in the conditioning box simply by choosing its input impedance properly. Given a hydrophone with a capacitance $C$ (provided by the manufacturer), and a desired −3 dB cut off frequency of the filter of $f_0$, the input resistance of the conditioning box should be $R = 1/(2\pi f_0 C)$. Anti-aliasing is partly facilitated by the roll-off in hydrophone sensitivity at high frequencies, but should always be augmented with a low-pass filter located as close as possible to the ADC, to avoid picking up high-frequency noise in the cables between the low-pass filter and ADC. It is tempting to use a steep, high-order anti-aliasing filter to maximize the usable bandwidth below the Nyquist frequency while still filtering the frequencies above the Nyquist frequency efficiently, but that can cause ringing, dynamic range limitations and phase matching problems. These difficulties are substantially avoided by using a lower order (e.g. a 4–6th order Butterworth type) low-pass filter that starts well below the Nyquist frequency. A good rule of thumb is to sample at least three times faster than the highest frequency of interest. For some toothed whale signals a bandwidth of 200 kHz is required, which in turn calls for a sampling rate of 500–600 ksamples/s per channel to get a faithful discrete representation of the signal, while avoiding aliasing problems. An elegant way to eliminate aliasing problems is to use a sigma–delta converting ADC. This type of converter, common in modern audio recording equipment and PCs, samples at a very high rate and then combines samples digitally to reduce the sampling rate using a steep digital anti-alias filter.

### 2.3. Acoustic localization and array configurations

Localization of a vocalizing animal is important in many behavioral studies and critical for deriving TL for source parameter estimation. One way to accomplish this is to record the sound of the animal with an array of receivers (Fig. 5). A signal emitted by the source will arrive at the various receivers with different time delays, due to different propagation path lengths from the source to the receivers (Spiesberger and Fristrup, 1990). For a certain source location relative to the array, there will be a certain set of time delays. It is possible to use these time lags to calculate the location of the source, and thereby the spatial relationship between the source and the receivers needed to estimate the TL.

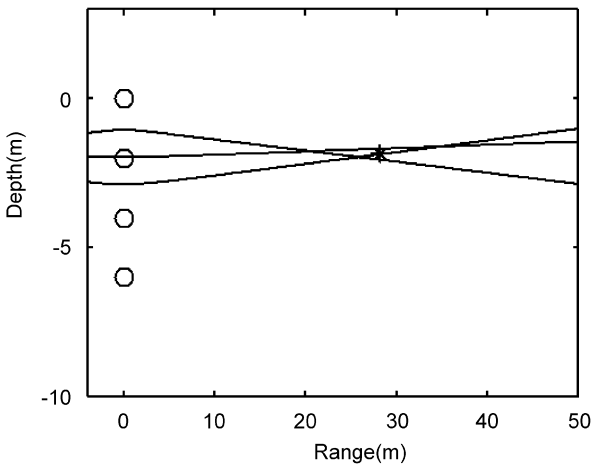


Fig. 5. Localizing a clicking dolphin with a 4-hydrophone linear array. The time-of-arrival differences at the receivers (marked with circles) generate three independent hyperbolas. They do not intersect in a single point, due localization errors as described in the text. The analytical source location (see appendix) is marked with a star and lies within the hyperbola crossings.

Consider first the problem of localizing a sound source in a 2-dimensional (2-D) situation (Fig. 5). From each pair of receivers the signal TOAD can be measured. Each TOAD restricts the possible location of the source to a hyperbolic curve, having its axis in the direction of the line connecting the two receivers. With one more receiver, another hyperbolic curve can be generated, and the source is ideally restricted to the intersection of the two curves.

Adding the third receiver will actually add *two* hyperbolic curves, as there will be two TOADs generated by the additional reception of the click, one relative to the first and one relative to the second receiver. However, it follows that one of these TOADs will be a function of the other one, and provide no new information on sound source location.

While horizontal 2-D localization may be feasible for some terrestrial situations, it is usually not feasible for toothed whales moving in a 3-dimensional (3-D) world (Wahlberg et al., 2001). For localization in three dimensions, each TOAD restricts the source to a hyperboloid surface. Using the line of argument as in the 2-D example above, three such surfaces are needed to restrict the source to one point. Thus, three receivers are needed to pinpoint the source location in two dimensions, whereas four receivers are needed in three dimensions. In some source–receiver constellations there may be two source locations generated by each set of TOAD, and here an extra receiver is needed to resolve the ambiguity (Spiesberger, 2001).

It is important to realize that if only a minimum number of receivers are used, there will be no explicit information available on how well the localization system performs. The hyperbolas or hyperboloids will always (except for some pathological cases) render a single and therefore at first glance a precise result (a perfect intersect), independent of the error in data fed into the localization algorithm. The situation is similar to the one using only two Cartesian coordinates to determine the slope of a line. Using linear regression the line will perfectly intersect the two data points. However, we will have no chance of knowing how accurate the derived slope is. This can be achieved only from increasing the number of data points. Similarly, more receivers are needed to determine how well the hyperboloids intersect. A well-defined intersect from many hyperboloids provides considerably more confidence in the derived source location, compared with only having one such intersect. The drawback of this is that more receivers are needed, adding logistical challenges in the field and more data to process.

The precision in the calculated source location is dependant on the information fed into the localization algorithm. For good localization, precise information is needed on the receiver locations, the TOADs and the sound velocity of the medium. The meaning of 'good' and 'precise' is determined by how accurate the investigator needs to know the location of the vocalizing animal. For smaller arrays (e.g. Au et al., 2004; Schotten et al., 2003; Madsen et al., 2004a, b; Villadsgaard et al., 2007), inter-receiver locations may be obtained with a measuring rod, and channel synchronization is obtained from connecting all receivers to the same recorder. For large arrays, the receiver locations may be deduced using signals emitted at known locations (Watkins and Schevill, 1972; Wahlberg et al., 2001). Recently, the Global Positioning System (GPS) has added an elegant way in which 2-D receiver locations and synchronization may be obtained from an arbitrary number of receivers located at any distance from each other using timing pulses from GPS satellites (Moehl et al., 2001).

Many toothed whale species forage at great depths, and biosonar signals may not be recordable from individuals close to the surface (Madsen et al., 2005a, b). To perform recordings at the depths where the animals are using their biosonar, hydrophone arrays may need to be deployed at great depths. This poses some problems for the hydrophone array design (see Moehl et al., 2003; Heerfordt et al., 2007) and the means by which they are deployed.

In Table 1, different array configurations are outlined, rendering different number and kinds of source coordinates relative to the receivers of the array. In general, the more receivers that are used, the more information can be obtained about the source location. Besides outlining the number of receivers required, Table 1 also shows that the geometric arrangement of the receivers is critical for the localization information that can be derived. In general, optimal source location is obtained for inter-receiver spacing of the same order of magnitude as the source-to-array distance, and for sources within the array (Moehl et al., 2003), or in the direction perpendicular to the array's geometric extension. For source parameter estimates, only the range from the receivers to the source is actually

Table 1
Summary of array types

| | $N_r$ | Geometry | Array name | $N_{coordinates}$ | Examples |
|---|---|---|---|---|---|
| ○ | 1 | Point | Single | – | – |
| ○ ○ | 2 | Line | Stereo | 1 | (Bearing) |
| ○ ○ ○ | 3 | Line | Linear | 2 | ($x$, $y$), range |
| ○<br>○ ○ | 3 | Plane | 2-D MINNA | 2 | ($x$, $y$, bearing, range) |
| ○ ○ ○<br>○ ○ ○○ | >3 | Plane | 2-D ODA, or 2/3-D | >2 | $x$, $y$, ($z$, bearing, range) |
| ○ ○<br>○ ● ○ | 4 | Volume | 3-D MINNA | 3 | ($x$,$y$,$z$, bearing, range) |
| ○ ○ ○ ● ○<br>○ ○ ● | >4 | Volume | 3-D ODA | >3 | $x$,$y$,$z$, bearing, range |

$N_r$: number of receivers, MINNA: minimal receiver number array, ODA: over-determined array. $N_{coordinates}$ indicates the number of source coordinates that are possible to derive with the array. In the first column, unfilled circles indicate receivers in the horizontal plain, while filled circles are located on a different vertical coordinate. Parenthesis around examples of coordinates indicates that the coordinates may be ambiguous, e.g. when the source may be located on either side of the array plane.

needed, and therefore a vertical linear array of sufficient aperture will do (e.g. Moehl et al., 1990; Madsen et al., 2004a, b, Heerfordt et al, 2007). Linear sparse arrays are favorable in the sense that the region of accurate localization extends quite far out from the array on its broadside axis. However, if information is also needed on animal movement and the directionality of the clicks at short ranges, a planar array may in some cases be more useful (e.g. Au et al., 2004; Rasmussen et al. 2004).

Additional virtual hydrophones may be constructed by using surface-reflected source to receiver paths that can improve the vertical resolution of the localization. In calm waters, the surface-reflected path to the hydrophone may be viewed as a signal reaching a receiver situated above the water surface at a height corresponding to the hydrophone depth. Additional reflections from the bottom, surface-bottom, etc. may provide additional virtual hydrophones (Thode et al., 2002). Thereby, hyperboloids can be generated from such reflections so that the location of the whale can be estimated from single-hydrophone recordings (e.g. Aubauer et al., 2000; Laplance et al., 2004). Likewise, hyperboloids from virtual hydrophones can be used to improve the localization precision with multi-hydrophone arrays (Moehl et al., 1990; Wahlberg et al., 2001; Thode et al., 2002).

The location of the virtual hydrophones, and thereby their usefulness, is critically dependent on surface waves which may disturb the perfect mirroring of the sound path, and stability in hydrophone depth (Urick, 1983). Another issue with surface-reflected paths is that, as noted below, if the animal is pointing its beam upwards so that the surface-reflected path is much stronger than the direct path, we may easily misinterpret the reflected path as being the direct one. For minimum receiver number arrays (MINNAs) there is no easy way to mediate such errors (Table 1). For over-determined arrays (ODAs) the problem is usually quite easily recognizable as it often results in one or several hyperbolas pointing towards a completely different point of intersection than most of the other ones (Table 1). Almost anyone working with sound localization of whales has experienced whales localized below the sea bed or above the water surface. Such errors can be caused by an erroneous interpretation of click multipaths or by undetected tilt in an assumed vertical array. Such obvious errors should be regarded as an excellent chance both to control and debug the localization algorithms used, but also to carefully scrutinize the interpretation of the recorded signals.

The coordinates of the source relative to the receivers may be derived from the hyperbola intersects or through calculations. It is prudent to do both and trust localizations only where the two independent methods render comparable results. For 3-D localization the hyperboloid plots can be projected on the vertical plane crossing one of the receivers and the best hyperboloid intersect (see Wahlberg et al., 2001).

For calculating the source location, different approaches may be used. Most of them rely on the following basic scheme. Each signal impinging

on a hydrophone results in an equation which we write as

$$|\mathbf{r}_i - \mathbf{s}|^2 = (cT_i)^2.$$

In this equation, $|\mathbf{r}_i - \mathbf{s}|$ is the distance between the source ($\mathbf{s}$) and the $i$th hydrophone ($\mathbf{r}_i$). Bold types denote vectors, so the source has coordinates $\mathbf{s} = [s_x, s_y, s_z]$, where the last element is omitted in 2-D cases. Similarly, each receiver has coordinates $\mathbf{r}_i = [r_{ix}, r_{iy}, r_{iz}]$, where we again omit the last element in 2-D applications. The sound velocity is denoted $c$, and $T_i$ is the time it takes the signal to travel from the source to the receiver. There will be a total of $N$ such equations, where $N$ is the number of receivers in the array.

When recording signals with an array, we may determine the receiver locations and the sound velocity, but we do not know the time of arrivals, $T_i$. What is known, however, are the TOAD ($\tau_i$) between any receiver and e.g. receiver no. 1, $\tau_i = T_i - T_1$. If the first equation is subtracted from the $N-1$ other equations, and if we choose the coordinate system so that the first receiver is in the origin (0,0,0), we end up with the following set of equations:

$$2\mathbf{r}_i\mathbf{s} + \mathbf{2}c^2\tau_i \mathrm{T}_1 = -c^2\tau_i^2 + |\mathbf{r}_i|^2, \quad i = 2 \ldots N.$$

For MINNA systems this equation may be solved for $s_x$, $s_y$ and $s_z$ through elimination, rendering the source coordinates. For ODA systems there will be more equations than unknowns, and least-squares or other smoothing techniques may be used to derive the source location (see Wahlberg et al., 2001; Spiesberger, 2001 for details). Another, and sometimes more reliable, approach for ODA systems is to split up the localization system into smaller MINNA 'cells', and derive an averaged source location from the MINNA locations (Spiesberger and Wahlberg, 2002; Wahlberg, 2003; Spiesberger, 2005).

Localizing consecutive vocalizations greatly helps to validate the data, improve source localization precision and sort out outliers. In addition, with data from ODA systems, possible biases in some TOAD or other array data can be identified to help ruling out erroneous measurements and further increase the localization performance.

As noted above and in Table 1, different array geometries will render different numbers and types of source coordinates, and with a precision depending on many factors. For measurements of the source properties of toothed whale echolocation signals we find the linear and star-shaped arrays particularly useful. The linear array has the advantage of being able to cover a relatively large volume of the water within which it can localize sound sources with adequate precision, and they are also easier to tow and deploy at great depths. Three hydrophones are needed to derive the bearing and range to the sound source, which is sufficient for assessing the TL and thereby the SL. Adding an extra hydrophone improves the reliability in the derived source location and thereby increases the effective range of the array. The distance between the hydrophones should be kept as large as possible, but not larger than the same signal that can be recorded on all receivers. Also, the larger the inter-receiver distances, the more difficult it is to keep track of the inter-receiver distance, and the general shape of the array.

A disadvantage with the linear array is that it is not possible to estimate the animals' 3-D swimming direction using consecutive clicks and thereby test if the animal's body axis is generally pointing towards a hydrophone. A star-shaped array is more suited for this (Au et al., 2003; Rasmussen et al., 2002; Rasmussen et al., 2004). The localization algorithm and the on-array-axis localization accuracy for a star-shaped array have been published by Au and Herzing (2003). The drawback with this array type is that animals can only be localized with some confidence within a relatively narrow volume of water right in front of the array aperture. Also, most small planar arrays are MINNA systems that ideally should be augmented by an additional hydrophone to improve the confidence in the derived source locations.

The localization precision will critically depend on the degrees of freedom given by array shape and size, the source–receiver geometry, the signal type and the SNR of the received clicks. Before determining which array is needed for what kind of signal at which range, it is important to consider the required accuracy. For SL measurements of interest here, a precision within a decibel is rarely crucial. As noted above it is above all the TL estimates that determine the accuracy at which the SL can be computed. For spherical spreading with an attenuation no higher than that for harbor porpoise signals (0.038 dB/m), a variation of 2 dB in TL (and thereby in SL) corresponds to a maximum absolute error in range of about 20% out to a distance of 100 m (sloping dotted line in Fig. 6). The slopes of the lines are the effect of attenuation. All

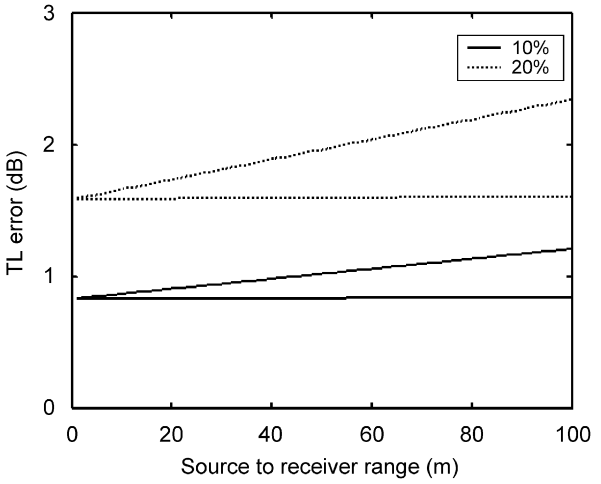


Fig. 6. Maximum error in calculated transmission loss for a harbor porpoise signal (sloping lines) and a sperm whale (almost horizontal lines) located at a range of up to 100 m distance, with a ranging absolute error of 10% and 20%, corresponding to received level errors of 1–2 dB when only considering spherical spreading.

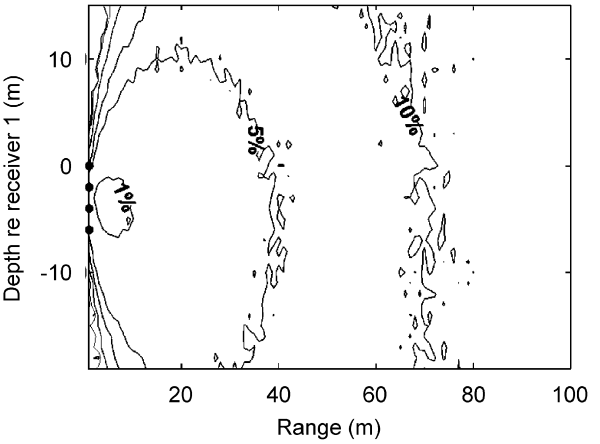


Fig. 7. Modeling the source localization error for a 4-hydrophone array, assuming a good signal-to noise ratio on all four receivers. The sound velocity is assumed to be known within 10 m/s, the receiver locations within 5 mm, and the TOADs are assumed to be measurable within 10 μs. The errors are given in units of percentage error (calculated as the average error out of 100 simulations) relative to the range to receiver 1 (the top receiver). Receivers are denoted by filled circles. The figure does not take into account any erroneous horizontal shift in receiver locations (array bending).

toothed whales known up to date have center frequency equal to or lower than the one found for harbor porpoise clicks, and the signals will accordingly suffer from the same or less attenuation. The biosonar signals with lowest frequency emphasis known is generated by the sperm whale, having an attenuation of about 1–2 dB/km at 15 kHz. The effect on the TL of sperm whale signals is also depicted in Fig. 6 as the almost horizontal lines. All other known toothed whale clicks will be found within the 'V' of the harbor porpoise and sperm whale data depicted in Fig. 6.

Having determined the ranging accuracy needed, we may determine which array configuration is needed using error modeling. Such modeling can be done with many different techniques such as linear error propagation modeling (Wahlberg et al., 2001) or numerical modeling (Spiesberger and Wahlberg, 2002). Numerical modeling usually renders the most reliable results (Spiesberger and Wahlberg, 2002; Wahlberg, 2003). Note that source localization precision is not only a function of the range to the receiver, but also to the source's aspect to the array (Fig. 7). Conveniently, another nice feature with the linear array is that it will provide robust bearing estimates so that sources located close to the array axis may be extracted for further analysis with a high degree of confidence.

Needless to say, it is clear that acoustic localization is a rather complicated process producing a large number of possible error sources for wrongly assessing the range between the whale and the receivers. It is therefore essential that both algorithms and hardware are regularly checked through calculating the location of sound sources produced at known locations around the array.

### 2.4. Analysis of array data

Array data can be used to track and monitor vocalizing animals for use in behavioral and abundance studies (e.g. Wahlberg, 2002; Leaper et al., 2002; Thode et al., 2002). Here we will address how array data from a calibrated hydrophone array can be used to derive source parameters of echolocation clicks from toothed whales. Having obtained array recordings, the signal sources may be localized with some precision and the TL and SL can be determined. The derived sound levels will most likely vary tremendously for two reasons. First, toothed whale echolocation signals are highly directional, with a 3 dB beam width in some species smaller than $\pm 5^{\circ}$ (Au, 1993). Therefore, depending on whether the receiver is aligned with the animals' acoustic axis, the received sound level may vary by more than 40 dB. Also, echolocating animals are known to vary their output levels by more than 20 dB in a single click train (Madsen et al., 2002a).

This poses a problem since we wish to record the on-axis signature of the clicks when assessing their sonar potential. However, from merely deriving the back-calculated sound pressure level to 1 m from the source, there is no way of telling a faint on-axis signal from a powerful off-axis one. Signals recorded off-axis may be relevant for passive acoustic monitoring for some species such as sperm whales, but have little value when evaluating the performance of toothed whale biosonar systems. Secondly, recording of clicks at varying degrees off-axis can lead to erroneous classification of seemingly different sonar signals from the same species or animal into different click types, that may in fact be caused by varying degrees of off-axis distortion. Biosonar signals must therefore be recorded on or close to the acoustic axis of the transmitting aperture of the clicking toothed whale to quantify the clicks with respect to performance of the biosonar system and to render inter- and intra specific comparisons and classifications meaningful. Statistical treatment of clicks in the light of biosonar performance makes sense only for on-axis clicks. Subsequently, simultaneous recording of off-axis versions of the same click recorded on-axis can provide an array of useful information of the sound-generating system.

Acknowledging these problems, the concept of the Apparent Source Level (ASL) was introduced by Moehl et al. (2000) (see Fig. 1). The ASL is defined as the acoustic intensity back-calculated to 1 m in any direction from the sound source. While this does not facilitate classification of on-axis signals, it provides a way to report the acoustic level back-calculated to 1 m, without knowing in which direction relative to the sound source the signal was recorded. The term SL should only be used for sound pressure levels back-calculated to 1 m on the acoustic axis of the animal (Urick, 1983). Quite often in the bioacoustic literature SLs are reported where it would be more appropriate to use the ASL instead.

Several criteria have been used to determine whether or not a signal is recorded on-axis. First, the relative amplitude difference between clicks recorded at the various receivers may be used to define the on-axis clicks as signals of highest amplitude as compared to adjacent receivers (Au et al., 2004) or within a certain amplitude difference, such as 3 dB, compared to the same click on other receivers (Au and Herzing, 2003). This criterion is useful only for clicks recorded within a distance where there will be a measurable reduction in received level due to directionality on adjacent hydrophones. The amplitude information in a click train, within both the same receiver channel and adjacent ones, can be used to define on-axis signals as the one with the highest back-calculated sound pressure level, assuming that the animals sometimes scan their sound beams across the array by which one of the hydrophones are ensonified by an on-axis version of a click (Madsen et al., 2004b). The latter approach can be used in conjunction with a threshold for ASL that is considered to be high enough to represent an SL (Moehl et al., 2000, 2003), but that will inevitably introduce a bias of only quantifying clicks with high SLs (Moehl et al., 2003). Surface reflections may also be used to improve the confidence that the analyzed signal is on-axis. If the surface-reflected version of the click has a higher amplitude than the direct path, there is a good reason to believe that the animal acoustic axis is not pointing toward any of the receivers (Madsen et al., 2004a, b).

While the equal levels approach (e.g. RL's within 3 dB on the receivers) is likely too generous by accepting upwards two-third of all recorded clicks as being recorded on-axis (Au et al., 2004), it is also clear that a minimum level approach will exclude many low-amplitude clicks recorded on-axis (Moehl et al., 2003).

For broadband signals (as are most toothed whale clicks) the spectral and temporal properties give additional clues to whether or not the signal is recorded on-axis as clicks recorded off-axis will suffer from low-pass filtering and spectral notches induced by interference between time delayed pulse components radiated from an array of different point sources in the toothed whale nasal complex (Au, 1993; Madsen et al., 2004a, b; Beedholm and Moehl, 2006). The acoustic localization information from tracking the animal in a series of clicks may also hint to when the acoustic axis is most likely pointing towards the array, assuming that its acoustic axis is identical to the animal's body axis as expressed by the pointing vector of the swimming direction and that localizations are accurate enough to derive a reliable velocity vector (Rasmussen et al., 2004). Irrespective of the technique used to determine if a signal is on-axis, it is important to stress that no technique so far developed seems completely satisfying and none of the methods have been thoroughly ground truthed under controlled conditions. Accordingly there is always a risk of including

off-axis clicks or excluding on-axis clicks in the analysis. We recommend that one use several of the above-mentioned criteria and be very specific about the implemented methodology and its caveats when reporting and discussing the data. More specifically, for vertical arrays, we recommend that a click is classified as being recorded on-axis, when it is the highest in a scan passing the hydrophones and that the back-calculated ASL is the highest on one of the hydrophones in the array (Madsen et al., 2004b).

## 3. Analysis of discrete signals

This section addresses how the source parameters of clicks can be analyzed on the basis of discretely sampled amplitude values in a way that is meaningful for quantifying the sound generator in an aquatic biosonar system. An analog to digital converting (ADC) recorder creates a discrete time series by sampling and storing amplitude values of an analog voltage signal. The sampling rate defines how often the analog signal is sampled per second. The faster the sampling rate, the higher frequencies can be reproduced unambiguously (see Nyquist sampling theorem above). Time and frequency domains are linked via the Fourier transform and both domains are useful to quantify when assessing the performance and properties of toothed whale biosonars (Au, 1993, 2004). In the time domain, relevant parameters include click repetition rates, sound pressure and energy measures, and duration. In the frequency domain measures of frequency emphasis and bandwidth should be computed.

### 3.1. Repetition rate, amplitude and duration measures

An often used proxy for the maximum range at which the toothed whale is acoustically searching for prey is the interclick interval (ICI), because toothed whales in captive trials normally use ICIs given by the two-way travel time to the target and a fixed, short processing or lag time (Au, 1993). Although the ICI may reflect the maximum time an echolocating toothed whale is prepared to wait for a returning echo (Au, 1993), the actual two-way travel time to the target may be much shorter than the ICI for animals echolocating in the wild (Madsen et al., 2005a, b). The short durations and well-defined onsets of clicks make derivation of ICI straightforward and any click detector or measures by hand would normally suffice. If the time–bandwidth product (duration times bandwidth) of the signal is sufficient, higher temporal resolution in ICIs can be achieved by using a high SNR on-axis click from the same recording as a matched filter.

An absolute dB measure must be accompanied by a reference value along with information on how the magnitude of the sound pressure was quantified. Measures of magnitude for aquatic biosonars are due to lack of standardization variously reported in terms of peak-to-peak, the root of the mean of the pressure squared (RMS), peak-equivalent RMS and energy flux density measures (Au et al., 1974; Moehl et al., 1990; Au, 1993; Madsen et al., 2004a, b). It is important to emphasize that sound pressure is always given in Pa ($N/m^2$), and that different notations of pressure on a dB scale simply represent different ways of measuring the pressure in Pa relative to a reference pressure of 1 μPa measured in the same way. However, for the same transient waveform, pressure levels in decibels may vary by 15 dB or more depending on how it was measured, rendering meaningful comparisons difficult without detailed information about how the pressure amplitude was quantified. For a sine wave the ratio between peak-to-peak and rms is 9 dB, but for aperiodic signals, such as toothed whale clicks, the difference between peak-to-peak and RMS can be 12 dB or more. The most common and straightforward measure for quantifying the magnitude of toothed whale clicks is the peak-to-peak sound pressure level (Au et al., 1974; Au, 1993) that can be read directly from an oscilloscope. However, since the mammalian ear operates as an energy detector, it is also relevant to include the rms and energy flux density measures of biosonar signals (Au et al., 1999).

A peak-to-peak sound pressure level can be derived in a straightforward manner from a digitized version of the click (Fig. 8) provided that the sensitivity, the gain of the recording chain and the maximum peak voltage of the ADC are known. Let us consider a transient from a delphinid toothed whale (Fig. 8a). It is impinging on a hydrophone with a sensitivity of −206 dB re 1 V/μPa and is amplified by 20 dB before digitization by an ADC with a peak clipping voltage of ±5 V. It follows that a 1 V input to the ADC from the hydrophone and amplifier corresponds to a received sound pressure of 186 dB re 1 μPa (206–20 dB re 1 μPa). An amplitude of unity (+1 or −1), corresponding to 5 V into the ADC, in the discrete version of the click after digitization in the ADC, must therefore

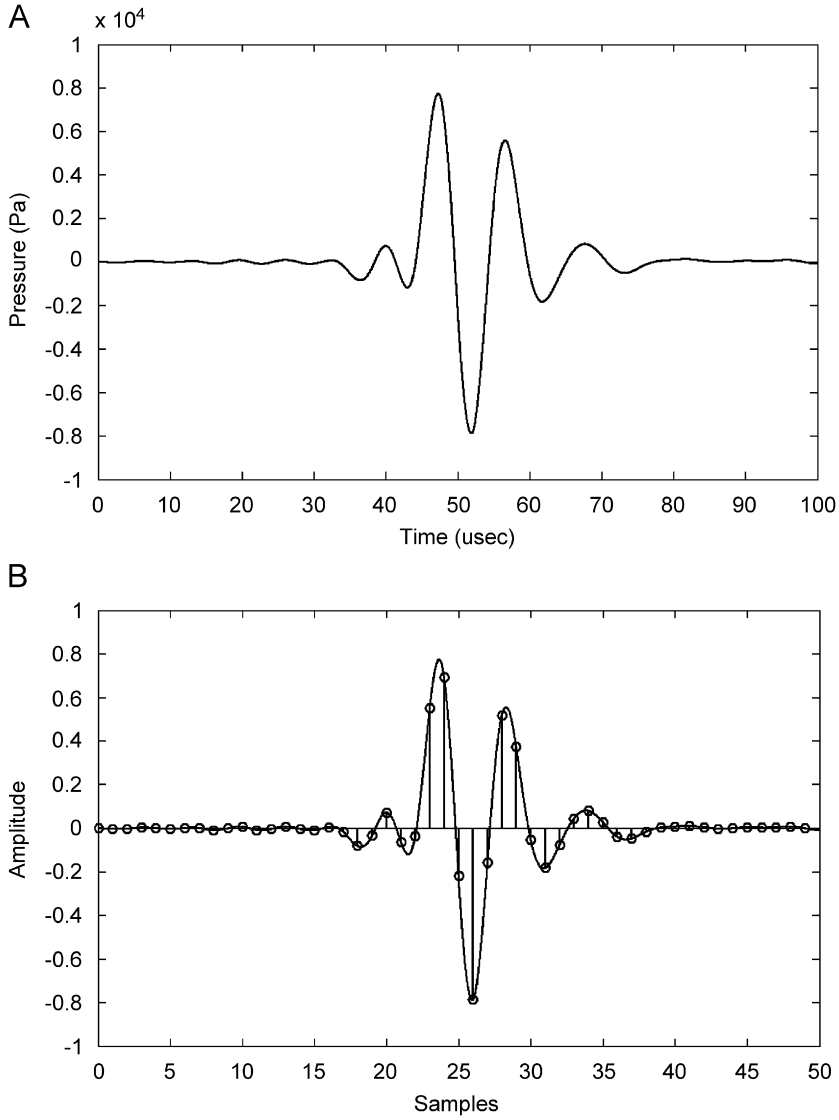


Fig. 8. (A) Toothed whale click as received by the hydrophone. (B) Discrete version of the click in (A) sampled at 500 ksamples/s. (C) Cumulated relative energy in the interpolated (10 times) waveform of (B) to define 95% duration (dotted lines). (D) Peak–peak and rms measures computed on the discrete version of the click in (A).

equal a received sound pressure of 200 dB re 1 μPa on the hydrophone (206 dB re 1 μPa/V−20 dB + 20 log(5 V/1 V)). It follows that a peak-to-peak amplitude from the discrete signal of 1.57 (0.78 + 0.79 in Fig. 8d) is equivalent to a peak-to-peak sound pressure level of 20 log(1.57) + 200 dB re 1 μPa = 204 dB re 1 μPa (pp).

To approximate the rms value of a click, Moehl et al. (1990, 2003) used the peak equivalent rms (peRMS) measure that compares the peak pressure of a click with the rms measure of a calibration signal. The peRMS is still a peak measure and is by definition 9 dB less than the peak-to-peak pressure, if the calibration signal is sinusoidal. True rms measures of toothed whale clicks will depend on the size of the chosen window over which the squared pressure is averaged. This will almost always render a discrepancy between peak-to-peak and rms higher than 9 dB. We therefore strongly advocate calculation of the rms sound pressure rather than using the peRMS measure as a proxy for rms.

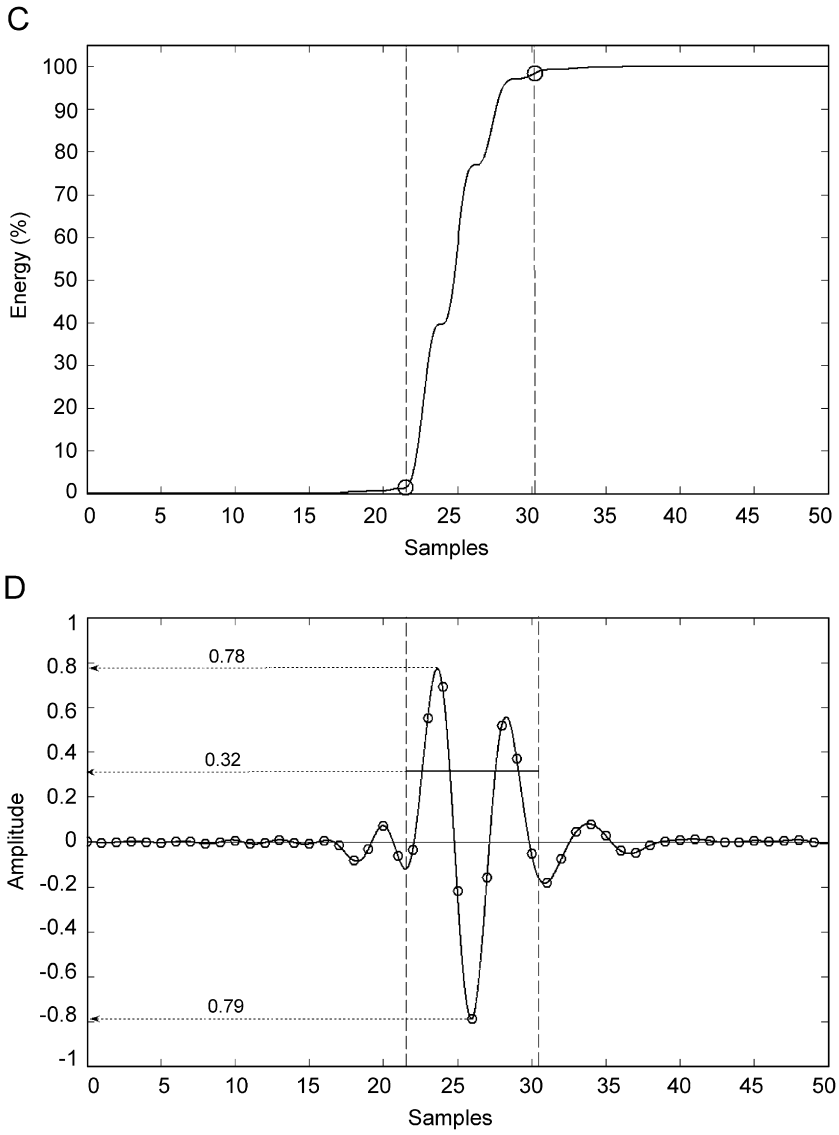


Fig. 8. (*Continued*)

The rms pressure of a plane wave in a time window from 0 to $T$ is given by

$$p_{\mathrm{rms}} = \sqrt{\frac{1}{T}\int_0^T p^2(t)\,\mathrm{d}t},$$

$$\mathrm{dB\ re\ 1\ \mu Pa\ (rms)} = 20\ \log \sqrt{\left(\frac{1}{T}\int_0^T p^2(t)\,\mathrm{d}t\right)},$$

where $p(t)$ is the instantaneous pressure (Urick, 1983).

The RMS of a discrete signal with an amplitude variable $s$ consisting of $N$ samples can therefore be computed as follows:

$$10\ \log\left(\frac{1}{N}\sum_1^N s^2\right).$$

The length of the analysis window is critical for RMS measures of transient signals, since the duration determines the window over which the pressure squared should be averaged (Madsen, 2005). For the same transient waveform, the rms level will decrease with increasing duration of the averaging window. Various techniques have been used to define the averaging window for toothed

whale clicks. The so-called D duration, which is given by the −10 dB end points relative to the peak of the envelope of the waveform, has been applied to determine the durations of narwhal clicks (Moehl et al., 1990). The envelope is computed by taking the absolute value of the analytical signal (consisting of the signal as its real part, and the Hilbert-transformed signal as its imaginary part). As a variation of this approach, Moehl et al. (2003) used −3 dB end points relative to the peak of the envelope when computing rms measures of p1 pulses in sperm whale clicks. However, the use of the −3 dB definition will render 2–3 dB higher rms measures than any other approach (Madsen, 2005) since only the highest amplitude values of the clicks are considered. We advocate using either the D-duration or a window enclosing a high, fixed proportion of the energy of the click, as both measures render about the same rms values (±1 dB) for good SNR (Madsen, 2005). In the latter approach, the duration of transients is determined by using the relative energy in a window that incorporates the entire signal waveform along with short samples of noise on either side. For short duration, high-SNR clicks from toothed whales, a 95 or 97% energy approach has been implemented (Au, 1993; Madsen et al., 2004a, b) (Fig. 8c). For example, for a 95% energy window, the onset of the signal would be defined as the time at which 2.5% of the signal energy is reached, and the termination of the signal is defined as the time at which 97.5% of the signal energy was reached. Using this method we find that the click depicted in Fig. 8a has a duration of 17 μs compared to 16 μs derived with the −10 dB re peak of envelope approach. For very short clicks lasting only a few samples, it is prudent to interpolate (e.g. by a factor of 10) to achieve a better time resolution (Madsen et al., 2004a, b).

For the click of Fig. 8a, the rms amplitude of the waveform in the window making up 95% of the relative energy is computed to be 0.32 (Fig. 8d). Hence, the received rms sound pressure level can be calculated to be 20 log(0.32)+200 dB re 1 μPa=190 dB re 1 μPa (rms). There is typically a difference of 12–15 dB between the rms and the peak–peak pressure of most toothed whale clicks depending on their crest factor (Au, 1993; Madsen et al., 2004a, b).

The energy of a sound pulse is given by the intensity integrated over the pulse duration, where the intensity is proportional to the time-averaged pressure squared for a plane wave in an unbounded medium (Urick, 1983, Au, 1993). Hence, the energy flux density (dB re. 1 $\mu Pa^2$ s) of transients can be derived by 10 log to the time integral of the squared pressure (sum of squared pressures for the discrete version of the signal) over the duration of the pulse (Young, 1970), which for the same duration, $T$, is the rms level (in dB) + 10 log($T$):

$$\text{Energy flux density (dB re. 1 } \mu Pa^2 \text{ s)} = 10 \log \int_0^T p^2(t)\,dt = 10 \log\left(\frac{1}{T}\int_0^T p^2(t)\,dt\right) + 10\log(T),$$

where $T$ is the window length in seconds.

Accordingly, we can compute the received energy flux density of the click in Fig. 8 by integrating the rms sound pressure over the signal duration:

$$190\,\text{dB re 1 } \mu\text{Pa (rms)} + 10\log(17\times10^{-6}) = 142\,\text{dB re. 1 } \mu Pa^2\,\text{s}.$$

The energy flux density on a dB scale (dB re. 1 $\mu Pa^2$ s) can be converted to J/$m^2$ by dividing the squared pressure on a linear scale by the specific impedance $Z$ (which in the free acoustic field is given by the sound speed times the density) of the medium: 142 dB re. 1 $\mu Pa^2$ s=(158 $Pa^2_{rms}$ s/$m^2$)/(1500 m/s × 1025 kg/$m^3$)=0.10 mJ/$m^2$.

Subsequently, the estimated ASL or SL can be generated by adding the TL (TL = 20 log($r$) + $\alpha r$) to the received levels computed above. Thus, if the dolphin was 12 m from the array and the click had the energy centered around 100 kHz, the SL would be

$$204\,\text{dB re 1 } \mu\text{Pa (pp)} + 20\log(12\,\text{m}) + 12\,\text{m}\times 0.03\,\text{dB/m(absorption at 100 kHz)} = 226\,\text{dB re 1 } \mu\text{Pa (pp)}.$$

For the click in Fig. 8 this corresponds to SLs of 212 dB re 1 μPa (rms) and 164 dB dB re. 1 $\mu Pa^2$ s (16 mJ/$m^2$).

### 3.2. Frequency domain

The spectral properties of a digitized click can be unfolded with the Discrete Fourier Transform (DFT) that links the time and the frequency domains of a digitized signal. The most efficient calculation of a DFT is the Fast Fourier Transform (FFT), which can be used if the signal length (in samples) is a multiple of 2. The frequency resolution

in Hz of a spectrum (spectral bin width) is given by the sampling rate divided by the FFT size. Thus, a 256-point FFT on a click sampled at 500 kHz provides a frequency resolution in the spectrum of 1.95 kHz. At normal sampling frequencies, discrete versions of toothed whale clicks will only be 10–100 samples long, which either leads to derivation of the frequency spectrum on the basis of a much larger window or to a spectrum with very coarse frequency resolution. A smaller bin width on a short signal can at first glance be achieved through zero padding (interpolation), where the signal is extended by a number of samples with a value of zero, but the actual resolution has not improved. In Fig. 9 we have computed the power spectrum of the click displayed in Fig. 8. The spectral characteristics of the signal are quantified from a 256-point FFT window symmetrical around the peak of the click envelope (the absolute value of the Hilbert transformed version of the signal). The peak frequency ($f_p$, kHz) is defined as the center frequency of the band with the highest amplitude of the spectrum. The centroid frequency ($f_0$, kHz) is defined as the point dividing the spectrum in halves of equal energy (Au, 1993), which is a much more robust measure of the frequency emphasis of a broad band click than the peak frequency (Madsen et al., 2004a, b). Delphinid clicks often have bimodal spectra with two peaks more than an octave apart, but with small and varying amplitude differences across the spectrum. For such clicks the peak frequency can come out very differently for similar clicks, whereas the centroid frequency is much more similar as it uses the power distribution as a function of frequency in the clicks. The bandwidth (BW) of the signals can be parameterized by the −3 dB BW (kHz) and −10 dB BW (kHz) and by the centralized root mean square bandwidth (RMS-BW, kHz) (Au, 1993) providing a measure of the spectral standard deviation around the centroid frequency ($f_0$) of the linear spectrum (Fig. 9). The RMS-BW can be used as a proxy for the frequency window over which the animal integrates both signal energy and noise (Moehl et al., 2003).

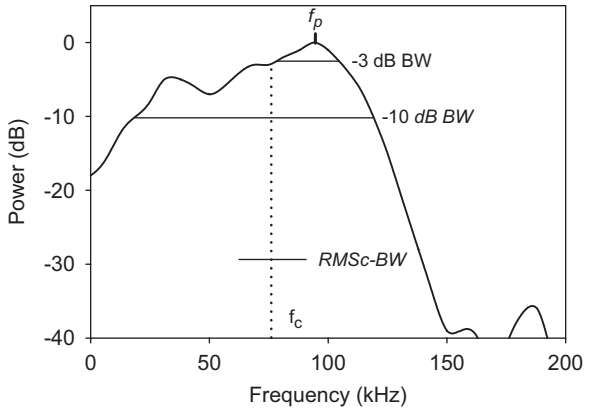


Fig. 9. Power spectrum generated from 256-size FFT. Binwidth of 2 kHz. $f_c$: centroid frequency, $f_p$: peak frequency and RMSc-BW: centralized root-mean-square band width.

### 3.2.1. DI estimation

If we can assure that the same signal has been recorded both on and off the acoustic axis in known angles, this information may be used to estimate the beam pattern (the acoustic intensity as a function of the angle to the acoustic axis) (Moehl et al., 2003; Rasmussen et al., 2004). With no information on animal orientation relative to the array, it must be assumed that the beam pattern is rotational symmetric around the acoustic axis, which is a reasonable approximation for most species measured to date (Au et al., 1986, 1987; Zimmer et al., 2005a, b, Beedholm and Moehl, 2006) with the exception of *Pseudorca crassidens* (Au et al., 1995). Knowing the beam pattern given by the ASL as a function of off-axis angle we may calculate the transmission directionality index (DI). The DI quantifies the directionality of a signal, and is defined as the ratio between the SL intensity on the acoustic axis and the SL intensity of a hypothetical omnidirectional sound source (where ASL = SL) radiating the same acoustic power as the sound source in question (Urick, 1983). A DI of 20 dB means that the SL is 20 dB higher than what would be obtained from an omnidirectional sound source emitting the same power. With the beam pattern of pressure ($p$) sampled and extrapolated at $N$ number of angles $v_1$ to $v_N$, spaced $\Delta v$ apart, the directivity index can be approximated by (Moehl et al. 2003)

$$\mathrm{DI} = 10\log\left(2\Big/\sum_{i=1}^{N} p_i \sin(v_i \Delta v)\right).$$

The angular spacing $\Delta v$ should be reasonably small to obtain a good approximation of DI with this equation. A useful theoretical framework for modeling of radiation patterns of echolocation clicks is that of a circular piston (Au et al., 1978, Au, 1993). The sound transmission of a toothed whale forehead can be modeled by a flat piston of equivalent aperture with the same DI (Au, 1993),

and the off-axis waveforms can be predicted with some accuracy by convolving the on-axis waveform with the transfer function of a flat piston as a function of off-axis angle (Beedholm and Moehl, 2006). The beam pattern given by the pressure, $p$, as a function of off-axis angle ($\vartheta$) can described by a first-order Bessel function using the wave number $k$ times the equivalent aperture radius $a$ (Au, 1993; Zimmer et al., 2005b):

$$p(ka\,\sin(\vartheta)) = p_0 \frac{2J_1(ka\,\sin(\vartheta))}{ka\,\sin(\vartheta)}.$$

The broad band beam pattern can either be derived by integrating $p(ka\,\sin(\vartheta))$ with respect to frequency, or approximated with a pure tone signal at the centroid frequency of the radiated click in question. The problem of using a single tone approximation is that it leads to deep notches in the beam response that will never be generated for a broad band signal (as are most toothed whale clicks). The $ka$ product $((2\pi/\lambda)\cdot a$, where $\lambda$ is the wavelength and $a$ is the radius of the transmitter) is a measure of the relationship between the effective transmission aperture and the radiated wavelength (here approximated by the centroid frequency of the radiated click). This measure comes in handy when assessing the relationship between the half power beam width ($\text{Beam}_{-3\,\text{dB}}$), the frequency emphasis of the echolocation click and the effective aperture of a sound transmitting anatomical structure.

For $ka \gg 1$, which applies for known echolocating toothed whales, the DI can be approximated by (Urick, 1983)

$$\text{DI} \approx 20\,\log(ka),$$

and the half power beam width can be estimated as (Zimmer et al., 2005b)

$$\text{Beam}_{-3\,\text{dB}} \approx 185^\circ/ka.$$

Consequently, the half power beam width is linked to the DI by (Zimmer et al., 2005b)

$$\text{Beam}_{-3\,\text{dB}} \approx 185^\circ \times 10^{-\text{DI}/20}.$$

When the equivalent aperture of a toothed whale sound projection system is estimated, the dimensions normally come out as being smaller than the physical dimensions of the nasal complex of the whale, and the equivalent aperture estimate is larger for larger animals which leads to a higher DI for the same frequency in larger animals (Au et al., 1978, 1987, 1995). However, it is far from evident how the functional morphology of the nasal complex leads to the observed radiation pattern, and there is at present no well founded functional rationale behind the notion that the toothed whale sound generator should behave as a flat piston in an infinite baffle. The piston model should therefore be seen as a helpful tool to approximate and predict radiation patterns from toothed whales, but not as a de facto paradigm that can be forced on all observed radiation patterns (Au et al., 1995).

## 4. Other techniques

Since derivation of source parameters in principle only requires knowledge on the range and the orientation of a clicking toothed whale with respect to a single calibrated receiver, it is possible to acquire the needed information without the aid of arrays of real and/or virtual hydrophones. Two recently developed techniques using onboard tags have provided elegant ways to derive source parameter estimates of echolocating toothed whales. Multisensor, archival tags, called Dtags (Johnson and Tyack, 2003) record sound with 16-bit resolution and at sampling rates up to 192 kHz along with animal orientation parameterized by 3-axis accelerometers and magnetometers sampled at 50 Hz.

Zimmer et al. (2005a) used such an onboard tag in conjunction with a towed hydrophone system to derive source parameters of echolocation clicks from a tagged sperm whale. Knowing the depth of both the tagged whale and the towed hydrophone allowed for derivation of range between the whale and the receiver by exploiting multipath delays between the direct and surface-reflected versions of the clicks. The relative orientation of the whale with respect to the hydrophone in the far field was derived from the absolute pointing vector of the tagged whale and the depth and estimated GPS position of the towed hydrophone. This setup provided recordings of clicks in known aspects and ranges to the whale, generating the first 3-D radiation pattern of echolocation clicks from a toothed whale in the wild (Zimmer et al., 2005a).

In a second study with Dtags, Zimmer et al. (2005b) used the novel approach of having two tagged echolocating Cuviers beaked whales recording each other during deep foraging dives. Both animals were tagged with Dtags recording clicks of the tagged whale and clicks of nearby conspecifics, including the other tagged whale. The time delays between emission and reception of clicks from one whale to the other and vice versa allowed for

derivation of the range between the two tagged whales. The 3-axis orientation sensors of the tags provided the relative orientation of the whales for each click, which along with range between the whales allowed Zimmer et al. (2005b) to compute ASL as a function of off-axis angles. Subsequently, the off-axis angle versus ALS data provided the basis for estimating SL, DI and the spectral properties of the clicks. The advantage of having tagged animals recording each other is that they often dive to the same depth when foraging and that they stay fairly close together. A possible bias is that they may try to avoid ensonifying each other with high-sound pressure levels, which may lead to underestimation of the source parameters (Zimmer et al., 2005b) or that no on-axis clicks are recorded.

## Acknowledgments

We dedicate this paper to Drs. Whitlow W.L. Au and Bertel Moehl who have pioneered quantitative recording and analysis of toothed whale echolocation clicks and formulated many of the concepts presented in this paper. We are greatly indebted to M. Johnson, W.M.X. Zimmer, B. Moehl, K. Beedholm, F.H. Jensen, B.K. Nielsen, J. Spiesberger, W.W.L. Au and two anonymous reviewers for helpful discussions and/or constructive critique on earlier versions of the paper. We also wish to acknowledge the skilled engineering contributions of N.U. Kristiansen. This work was supported by a Steno Fellowship from the Danish National Science Foundation to PTM, a grant from the Carlsberg Foundation to MW with additional support to the authors from Reson, the Novo Nordisk Science Foundation, Aarhus University Research Fund and the Oticon Foundation.

## Appendix A

Table A.1

Table A.1

```
% Source localization and hyperbola plot with linear array implemented in Matlab
% M Wahlberg, Aarhus University, April 2006
%
clear;
c0 = 1494.;                                          % speed of sound (m/s)
R = −4 − [0 2 4 6 ]′;                                % receiver depth coordinate (m)
t = [−0.1657 −0.1657 0]′*10^−3;                      % TOADs for receiver 2 -receiver 1 and so on, see Wahlberg et al. 2001 and this ms

r = (R(2:end)-R(1));                                 % normalized receiver coordinates
step = .01;                                          % step size in hyperbola plots
A = 2 * [r t*c0^2];                                  % source location A matrix, defined in Wahlberg et al. 2001 and this ms
b = −(t.^2)*(c0^2)+r.^2;                             % source location b column vector, defined in Wahlberg et al. 2001 and this ms

m = A                                                % source solution solved by least squares as in Wahlberg et al. 2001 and this ms

so(2) = m(1)+R(1);                                   % source depth coordinate
so(1) = sqrt((c0*m(2))^2−m(1)^2 );                   % source horizontal coordinate
figure (1); set(1,'Color',[1 1 1]);
plot(zeros(4,1),R,'ko');                             % plot receivers
ax = [−3 2*so(1) min([so(2) min(R)])−5 max([ so(2) max(R) % define size of plot
])+5];
hold on
plot(so(1),so(2),'r*');                              % plot source coordinates
for i = 1:length(r),
  a = c0*t(i)/2;                                     % the three parameters a, b, and c used to define the hyperbola curve in line 29

  c = r(i)/2;
  b = sqrt(c^2−a^2);
  if imag(b) == 0,                                   % check if TOAD render physical hyperbola
    if t(i) ~= 0
      y = −sign(r(i))*sign(t(i))*[step:step:2*(ax(4)−ax(3))]   % equally spaced vector for which to calculate a hyperbola below
;
      x = b * sqrt( y.^2/a^2−1);                     % equation of a hyperbola curve with the parameters a and b defined above
```

Table A.1 (*continued*)

```
    else
      y = zeros(1,2);                          % plot straight line instead if TOAD is zero
      x = [0 ax(2)];
    end
    ind = min(find(imag(x) = = 0));            % find index in x vector where hyperbola starts
    x = [fliplr(-x(ind:end)) x(ind:end) ]';    % extend hyperbola to both positive and negative x values
    y = [fliplr(y(ind:end)) y(ind:end)]' + c + R(1);   % extend y vector to two-sided hyperbola curve
    plot(x,y)                                  % plot hyperbola
  end
End
axis(ax);
xlabel('Range (m)');
ylabel('Depth (m)');
% End of program
```

## References


Au, W.W.L., 1993. Sonar of Dolphins. Springer, New York.

Au, W.W.L., 1997. Echolocation in dolphins with a dolphin–bat comparison. Bioacoustics 8, 162–185.

Au, W.W.L., 2004. Echolocation signals of wild dolphins. Acoustical Physics 50 (4), 454–462.

Au, W.W.L., Benoit-Bird, K., 2003. Automatic gain control in the echolocation system of dolphins. Nature 423, 861–863.

Au, W.W.L., Herzing, D.L., 2003. Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*). Journal of the Acoustical Society of America 113, 598–604.

Au, W.W.L., Floyd, R.W., Penner, R.H., Murchison, A.E., 1974. Measurement of echolocation signals of the Atlantic bottlenose dolphin, *Tursiops truncatus* Montagu, in open waters. Journal of the Acoustical Society of America 56, 1280–1290.

Au, W.W.L., Floyd, R.W., Haun, J.E., 1978. Propagation of Atlantic bottlenose dolphin echolocation signals. Journal of the Acoustical Society of America 64, 411–422.

Au, W.W.L., Moore, P.W., Pawloski, D., 1986. Echolocation transmitting beam of the Atlantic bottlenose dolphin. Journal of the Acoustical Society of America 80, 688–691.

Au, W.W.L., Penner, R.H., Turl, C.W., 1987. Propagation of beluga echolocation signals. Journal of the Acoustical Society of America 82, 807–813.

Au, W.W.L., Pawloski, J.L., Nachtigall, P.E., Blonz, M., Gisiner, R.C., 1995. Echolocation signals and transmission beam pattern of a false killer whale (*Pseudorca crassidens*). Journal of the Acoustical Society of America 98 (1), 51–59.

Au, W.W.L., Kastelein, R.A., Rippe, T., Schooneman, N.M., 1999. Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*). Journal of the Acoustical Society of America 196, 3699–3705.

Au, W.W.L., Ford, J.K.B., Horne, J.K., Allman, K.A.N., 2004. Echolocation signals of free-ranging killer whales (*Orcinus orca*) and modelling of foraging for chinook salmon (*Oncorhyncus tswawytscha*). Journal of the Acoustical Society of America 115 (2), 901–909.

Aubauer, R., Lammers, M.O., Au, W.W.L., 2000. One-hydrophone method for estimating distance and depth of phonating dolphins in shallow water. Journal of the Acoustical Society of America 107 (5), 2744–2749.

Barlow, J., Taylor, B.L., 2005. Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. Marine Mammal Science 21 (3), 429–445.

Beedholm, K., Moehl, B., 2006. Directionality of sperm whale sonar clicks and its relation to piston radiation theory. Journal of the Acoustical Society of America 119 (2), EL14–EL19.

Brill, R., Pawloski, J.L., Helweg, D., Au, W.W.L., Moore, P.W., 1992. Target detection, shape discrimination and signal characteristics of an echolocating false killer whale (*Pseudorca crassidens*). Journal of the Acoustical Society of America 92 (3), 1324–1330.

Carlstroem, J., 2005. Diel variation in echolocation behaviour of wild harbour porpoises. Marine Mammal Science 21 (1), 1–12.

Diercks, K.J., Trochta, R.T., Evans, W.E., 1973. Dolphin sonar: measurements and analysis. Journal of the Acoustical Society of America 54 (1), 200–204.

Evans, W.E., 1973. Echolocation by marine delphinids and one species of freshwater dolphin. Journal of the Acoustical Society of America 54, 191–199.

Heerfordt, A., Moehl, B., Wahlberg, M., 2007. A wideband connection to sperm whales: a fiber-optic deep-sea hydrophone array. Deep Sea Research I 54, 428–436.

Johnson, M.P., Tyack, P.L., 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE Journal of Oceanic Engineering 28, 3–12.

Laplance, C., Adam, O., Lopatka, M., Motsch, J.-F., 2004. Depth/range localization of diving sperm whales using passive acoustics on a single hydrophone disregarding seafloor reflections. Journal of the Acoustical Society of America 119, 3403.

Leaper, R., Chappell, O., Gordon, J.C.D., 2002. The development of practical techniques for surveying sperm whale populations acoustically. Report of the International Whaling Commission 42, 549–560.

Levenson, C., 1974. Source level and bistatic target strength of the sperm whale (*Physeter catodon*) measured from an oceanographic aircraft. Journal of the Acoustical Society of America 55 (5), 1100–1103.

Levin, P., 1973. Calibration of hydrophones. B&K Technical review 1, 13–17.

Madsen, P.T., 2005. Marine mammals and noise: what is a safety level of 180 dB re. 1 μPa (rms) for transients? Journal of the Acoustical Society of America 117 (6), 3952–3957.

Madsen, P.T., Payne, R., Kristiansen, N.U., Kerr, I., Moehl, B., 2002a. Sperm whale sound production studied with ultrasound-time-depth-recording tags. Journal of Experimental Biology 205, 1899–1906.

Madsen, P.T., Wahlberg, M., Moehl, B., 2002b. Male sperm whale (*Physeter macrocephalus*) acoustics in a high latitude habitat: implications for echolocation and communication. Behavioral Ecology and Sociobiology 53, 31–41.

Madsen, P.T., Kerr, I., Payne, R., 2004a. Echolocation clicks of two free-ranging delphinids with different food preferences: false killer whales (*Pseudorca crassidens*) and Risso's dolphin (*Grampus griseus*). Journal of Experimental Biology 207, 1811–1823.

Madsen, P.T., Kerr, I., Payne, R., 2004b. Source parameter estimates of echolocation clicks from wild pygmy killer whales (*Feresa attenuata*) (L). Journal of the Acoustical Society of America 116, 1909–1912.

Madsen, P.T., Carder, D.A., Beedholm, K., Ridgway, S., 2005a. Porpoise clicks from a sperm whale nose: convergent evolution of toothed whale echolocation clicks? Bioacoustics 15, 195–206.

Madsen, P.T., Johnson, M., Aguilar de Soto, N., Zimmer, W.M.X., Tyack, P.L., 2005b. Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*). The Journal of Experimental Biology 208, 181–194.

Medwin, H., Clay, C.S., 1998. Acoustical Oceanography. Academic Press, Boston.

Miller, L.A., Pristed, J., Moehl, B., Surlykke, A., 1995. Click sounds from narhwals (*Monodon monoceros*) in Inglefield Bay, Northwest Greenland. Marine Mammal Science 11, 491–502.

Moehl, B., Surlykke, A., Miller, L.A., 1990. High intensity narwhal click. In: Thomas, J., Kastelein, R. (Eds.), Sensory Abilities of Cetaceans. Plenum Press, New York, pp. 295–304.

Moehl, B., Wahlberg, M., Madsen, P.T., Miller, L.A., Surlykke, A., 2000. Sperm whale clicks: directionality and source level revisited. Journal of the Acoustical Society of America 107, 638–648.

Moehl, B., Wahlberg, M., Heerfordt, A., 2001. A large-aperture array of nonlinked receivers for acoustic positioning of biological sound sources. Journal of the Acoustical Society of America 109 (1), 434–437.

Moehl, B., Wahlberg, M., Madsen, P.T., Heerfordt, Lund, A., 2003. The monopulsed nature of sperm whale clicks. Journal of the Acoustical Society of America 114, 1143–1154.

Murchison, A.E., 1980. Maximum detection range and range resolution in echolocating bottlenose porpoise (*Tursiops truncatus*). In: Busnel, R.G., Fish, J.F. (Eds.), Animal Sonar Systems. Plenum Press, NY, pp. 43–70.

Rasmussen, M.H., Miller, L.A., Au, W.W.L., 2002. Source levels of clicks from free-ranging white-beaked dolphins (*Lagenorhynchus albirostris* Gray 1846) recorded in Icelandic waters. Journal of the Acoustical Society of America 111, 1122–1125.

Rasmussen, M., Wahlberg, M., Miller, L.A., 2004. Estimating transmission beam pattern of clicks recorded from free-ranging white-beaked dolphins (*Lagenorhynchus albirostris*). Journal of the Acoustical Society of America 116 (3), 1826–1831.

Richardson, W.J., Greene Jr., C.R., Malme, C.I., Thomson, D.H., 1995. Marine Mammals and Noise. Academic Press, London.

Roitblat, H.L., Helweg, D., Harley, H.E., 1995. Echolocation and imagery. In: Kastelein, R., Thomas, J., Helweg, D. (Eds.), Sensory Systems of Aquatic Mammals. De Spil Publishers, Woerden, pp. 171–181.

Schotten, M., Au, W.W.L., Lammers, M.O., Aubauer, R., 2003. Echolocation recordings and localizations of wild spinner dolphins (*Stenella longirostris*) and pantropical spotted dolphins (*Stenella attenuata*) using a four hydrophone array. In: Thomas, J., Moss, C.F., Vater, M. (Eds.), Echolocation in Bats and Dolphins. University of Chicago Press, Chicago.

Spiesberger, J.L., 2001. Hyperbolic location errors due to insufficient numbers of receivers. Journal of the Acoustical Society of America 109 (6), 3076–3079.

Spiesberger, J.L., 2005. Probability distributions for locations of calling animals, receivers, sound speeds, winds, and data from travel time differences. Journal of the Acoustical Society of America 118 (3), 1790–1800.

Spiesberger, J.L., Fristrup, K.M., 1990. Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. American Naturalist 135, 107–153.

Spiesberger, J.L., Wahlberg, M., 2002. Probability density functions for hyperbolic and isodiachronic locations. Journal of the Acoustical Society of America 112 (6), 3046–3052.

Surlykke, A., Moss, C.F., 2000. Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory. Journal of the Acoustical Society of America 108, 2419–2429.

Thode, A., Mellinger, D.K., Stienessen, S., Martinez, A., Mullin, K., 2002. Depth-dependent acoustic features of diving sperm whales (*Physeter macrocephalus*) in the Gulf of Mexico. Journal of the Acoustical Society of America 112 (1), 308–321.

Urick, R.J., 1983. Principles of Underwater Sound. Peninsula Publishing, Los Altos.

Villadsgaard, A., Wahlberg, M., Tougaard, J., 2007. Echolocation signals of wild harbour porpoises, *Phocoena phocoena*. Journal of Experimental Biology 210, 56–64.

Wahlberg, M., 2002. The acoustic behaviour of diving sperm whales observed with a hydrophone array. Journal of Experimental Marine Biology and Ecology 281, 53–62.

Wahlberg, M., 2003. Comparing a linear with a non-linear technique for acoustic localization of right whales. Canadian Acoustics 32 (2), 125–131.

Wahlberg, M., Moehl, B., Madsen, P.T., 2001. Estimating source position accuracy of a larger-aperture hydrophone array for bioacoustics. Journal of the Acoustical Society of America 109 (1), 397–406.

Watkins, W.A., 1980. Click sounds from animals at sea. In: Busnel, R.G., Fish, J.F. (Eds.), Animal Sonar Systems, pp. 291–298.

Watkins, W.A., Daher, M.A., 1992. Underwater sound recording of animals. Bioacoustics 4, 195–209.

Watkins, W.A., Schevill, W.E., 1972. Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array. Deep-Sea Research 19, 691–706.

Weber, P.J., 1963. The tape recorder as an instrumentation device, Ampex Corporation.

Wenz, G., 1962. Acoustic ambient noise in the ocean: spectra and sources. Journal of the Acoustical Society of America 34 (12), 1936–1956.

Young, R.W., 1970. On the energy transported with a sound pulse. Journal of the Acoustical Society of America 47, 441–442.

Zimmer, W.M.X., Tyack, P.L., Johnson, M.P., Madsen, P.T., 2005a. Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis. Journal of the Acoustical Society of America 117, 1473–1485.

Zimmer, W.M.X., Johnson, M., Madsen, P.T., Tyack, P., 2005b. Echolocation clicks of free-ranging Cuviers beaked whales (*Ziphius cavirostris*). Journal of the Acoustical Society of America 117, 3919–3927.