Animal behaviour

The long-range echo scene of the sperm whale biosonar

Pernille Tønnesen1, Cláudia Oliveira2, Mark Johnson3,1 and Peter Teglberg Madsen1

1Zoophysiology, Department of Biology, Aarhus University, 8000 Aarhus, Denmark
2Okeanos R&D Centre and IMAR - Institute of Marine Research, University of the Azores, 9901-862 Horta, Portugal
3Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St. Andrews, Fife KY16 8LB, UK

Sperm whales use their gigantic nose to produce the most powerful sounds in the animal kingdom, presumably to echolocate deep-sea prey at long ranges and possibly to debilitate prey. To test these hypotheses, we deployed sound recording tags (DTAG-4) on the tip of the nose of three sperm whales. One of these recordings yielded over 6000 echo streams from organisms detected up to 144 m ahead of the whale, supporting a long-range prey detection function of the sperm whale biosonar. The whale navigated this complex acoustic scene by maintaining a stable, long-range acoustic gaze suggesting continual resource evaluation. Less than 10% of the echoic organisms recorded by the tag were targeted for capture and only 18% of the buzzes were emitted within the 50 m depth interval of maximum organism encounter rate, demonstrating echo-guided prey selection. Buzzes were initiated more than 20 m from the prey, showing that sperm whales do not debilitate their prey with sound, but trade echo levels for reduced forward masking and rapid updates on prey location in keeping with the lower manoeuvrability of these large predators. We conclude that the powerful biosonar of sperm whales enables long-range echolocation and selection of prey, but not acoustic debilitation.

1. Introduction

The world’s largest tooth-bearing predator, the sperm whale, pushes the ecological boundaries for how large predators can be as they have to catch more of proportionally smaller prey than do smaller toothed whales [1]. Sperm whales are presumed to solve this problem by exploiting rich food niches such as the deep mesopelagic zone with a powerful biosonar system to obtain long-range echo information, thereby making foraging very efficient [1,2].

The characteristic hypertrophied nose of sperm whales that makes up more than 1/4 of the total body length [3] enables production of the most powerful sounds in the animal kingdom [4]. Mounting circumstantial evidence suggests that sperm whales indeed use their powerful clicks to echolocate prey over long ranges at great depths [5–8], but direct tests of this notion are lacking. Possession of such long-range biosonar would confer sperm whales the benefit of enhancing prey selection and thus increase their foraging efficiency. However, owing to their deep-diving behaviour, it is still unknown how sperm whales with their large bulky heads in fact capture prey [9–12]. One particular persistent theory posits that sperm whales use their high intensity clicks to acoustically debilitate their prey [13].

The acoustic behaviour of some smaller odontocetes has been related to characteristics of their prey fields by recording returning prey echoes [14–17].
Similar recordings have, however, likely been hampered for sperm whales owing to shadowing effects of their gigantic nose. Here, by deploying sound recording tags on the very tip of their nose, we attempted to record echoes from ensonified organisms to test whether sperm whales emit high intensity clicks to (i) perform long-range echolocation and (ii) acoustically debilitate prey [13].

2. Material and methods

Three sperm whales were instrumented with suction-cup-attached DTAG sound and movement recording tags far forward on the nose in the Azores archipelago (38°N, 26°W). The DTAG-4 (www.soundtags.org) recorded single-channel audio at 192 kHz with 16-bit resolution, providing a flat (±2 dB) frequency response between 0.4 and 60 kHz and a clip level of 184 dB re 1 µPa. Pressure and acceleration data were sampled at 50 and 250 Hz, 16 bit, respectively, and decimated to 25 Hz for analysis.

All analyses were conducted in Matlab R2019b (Mathworks) for the one whale for which the recording yielded echoes from ensonified organisms. Clicks were automatically detected and verified by visual inspection of spectrograms of sound recorded during deep dives (defined as greater than 100 m). Clicks from the tagged whale were easily distinguishable owing to clipping. Buzzes, indicating prey capture attempts [6], were distinguished by inter-click intervals below 0.154 s (sensu [7]). Echoes from ensonified organisms were identified using echograms (sensu [18]) made by stacking the envelope of filtered (Butterworth, 4 pole, band-pass: 5-40 kHz) sound segments following subsequent clicks (sensu [18]). Each recording segment was normalized according to the signal-to-noise ratio calculated by comparing the magnitude of each pixel to the 5th percentile magnitude of all pixels and a threshold of -3 dB above the normalization was applied. Echo streams from individual organisms were identified when echoes from three or more consecutive clicks formed a line with a slope (range/second) consistent with the likely forward speed of sperm whales (i.e. 0.5-2.5 m s⁻¹) (sensu [19]).

The ranges to ensonified organisms were calculated from the time delay to the echo after the emitted click assuming a sound speed of 1500 m s⁻¹. The observed encounter rate of organisms was defined as the number of detected echo streams divided by the number of minutes spent in 50 m depth intervals.

The range at which the sperm whale initiated the capture phase was extrapolated from the closing speed of the nearest echo stream within 20 s before the buzzes (69 buzzes, similar results were found using 5 and 30 s windows) and from echo streams within buzzes (2 buzzes).

To test if the whale adjusted its acoustic gaze according to the degree of acoustic complexity, we used a generalized linear mixed effect model (GLMM, Gaussian distribution, link function=identity, using the fitglme function in Matlab) to test for a relationship between median inter-click interval (dependent variable) and the number of ensonified organisms (explanatory variable) both sampled from ten second intervals during the bottom phase of the dives (1505 intervals), excluding periods from one minute before to one minute after buzzes. We included dive number as a random intercept and number of echoes as a random slope to account for the dependent nature of data coming from the same dive.

3. Results

This study shows that tag placement is crucial for recording returning echoes from sperm whale echolocation. Of three tags deployed on the front half of the nose of sperm whales, only the tag placed laterally right next to the blowhole (figure 1d,e) recorded returning echoes from ensonified organisms. This tag recorded for 22 h spanning 27 deep dives with 108 buzzes (median 4 buzzes per dive, range 0–8, figure 1a). The tagged sperm whale was, based on its size, categorized as a juvenile. The whale performed overall normal acoustic and diving behaviour, but made shallower and shorter dives (median 608 m, IQR: 448–652 m and median 36 min, IQR: 27–40 min, with dive depth likely limited by bathymetry) with fewer buzzes compared to adult sperm whales tagged in the Azores (C. Oliveira, M. Wahlberg, M.A. Silva, M. Johnson, R. Prieto, D.M. Wisniewska, A. Fais, J. Gonçalves, P.T. Madsen 2020, unpublished data). The tag changed recording position from pointing slightly downward to pointing slightly upward after the fifth dive and pointing horizontally after the eleventh dive (figure 1b,c).

Echoes from organisms ahead of the whale were detected in 24 of the deep dives (figure 1c). A total of 6395 echo streams from single organisms were detected with a median of 264 echo streams per dive (range: 4 to 882, IQR: 100–343; figure 1c). Buzzes were initiated a median of 24 m (IQR: 17–30, n = 69 buzzes, figure 2a) from the prey and no echo streams from escaping prey were detected after buzzes.

Echoes from organisms were detected up to a median range of 41 m ahead of the whale, while the maximum and minimum detection ranges were 144 and 9 m, respectively. The whale encountered three presumed schools of organisms (figure 2e); these were detected at maximum ranges of 245, 163 and 192 m.

During deep dives, we detected one or two orders of magnitude more ensonified organisms than prey capture attempts (i.e. buzzes) in 22 of the 23 dives in which both echo streams and buzzes were recorded. Only 18% of the buzzes occurred within the 50 m depth interval in which the maximum encounter rate of organisms was observed (figure 1c). The density of echoic organisms did not appear to move vertically over the 22 h tag attachment in a way consistent with diel migration.

The results of the GLMM show a significant, but very minute decrease by 0.009 s of inter-click interval for each added echo in the echoic scene (slope = -0.009, SE = 0.001, p-value = <0.001, R² = 0.07). When testing time intervals of increasing duration (steps of 10 s), we obtained similar p-values and R²-values but decreasing slope angles. Adding echoes as random slope to account for the dependent nature of echoes from the same dive did not have any significant effect (see electronic supplementary material S1).

4. Discussion

By placing a sensitive sound recording tag at the very tip of the gigantic sound producing nose of a young sperm whale, we have recorded echoes from ensonified organisms, providing the first direct evidence to support the hypothesis that the world’s biggest tooth-bearing predator, the sperm whale, hunts by echolocation. Owing to the substantial challenge of placing the tag correctly, this study is based on data from a single whale, fortunately offering a long recording period (22 h).

We were able to detect echoes in the tag sound recording at ranges of up to 144 m for single organisms and 245 m for...
schools, showing that the powerful clicks emitted by sperm whales enable long-range echolocation. Following the echogram method, echoes have been detected up to 20 m, 30 m and 3 m in front of beaked whales [16,18,20], bottlenose dolphins (large target [21]) and harbour porpoises [17]. Furthermore, owing to the hearing directionality and likely better hearing sensitivity of the whale compared to the tag, the detection range of the whale is likely 2–3 times larger than our reporting (see electronic supplementary material S2), in keeping with theoretical estimates of some 300 m prey detection range for sperm whales [22]. This is consistent with the substantially longer detection range of bottlenose dolphins obtained through psychophysical experiments [23] compared to estimated echo detection ranges visually gauged from echograms [21]. In combination, this highlights the advantage that sperm whales have gained by developing a large sound-generating nose to probe further into the deep sea. This long-range biosonar enables sperm whales to assess resources over large swathes of the deep-sea environment, thereby enhancing their foraging efficiency by allowing them to direct deep foraging dives to areas of high quality prey patches. Sperm whales therefore function as extreme deliberative sensory-motor predators with very large sensory volumes compared to their motor volumes [24]. The ever-present sea floor echoes from the powerful echolocation clicks (figure 1a and 2c) provide bathymetric landmarks to navigate prey fields, further highlighting the sensory superiority of sperm whales. In comparison, the substantially shorter detection ranges of other deep-diving toothed whales likely make these species much more dependent on stable, reliable prey distributions and hence likely restricts suitable habitats. Ultimately, this well-informed and therefore efficient resource utilization of sperm whales, facilitated by large investments in the biosonar apparatus, has likely played a key role in the evolution of unusually large body size of sperm whales compared to other odontocetes [1].

The potential challenge of such a long-range sensory capability is that sperm whales must contend with a multi-targeted auditory scene further complicated by ever-present strong reverberation from the seafloor whenever they echolocate and interfering echolocation from conspecifics, as often recorded by the tag (figure 2c,d,e). Despite the statistically significant result of the GLMM, indicating a minute and biologically irrelevant change (−0.009 s) of inter-click interval...
for each added echo within the echoic scene, we show that the whale navigated this acoustic environment by maintaining stable, high inter-click intervals allowing for a long acoustic gaze without range ambiguity, while continuously evaluating large water volumes.

In this multi-targeted prey field, the whale targeted less than 10% of ensonified organisms for capture. Prey tended to be targeted for captures at depths away from the largest echo densities that presumably coincide with the deep scattering layer (DSL), demonstrating echo-guided prey selection similar to that shown for Blainville’s beaked whales, suggesting that this is a common strategy of deep-diving odontocetes [14,20,25]. Echo-based foraging in the outskirts of the DSL may target larger prey (e.g. that are themselves predators on DSL organisms) and will reduce the confusion effect of prey aggregations [26]. A simpler auditory scene with less clutter may also facilitate auditory streaming and high sampling rates in the capture phase while avoiding range ambiguity from other organisms.

Having singled out a desired prey this approximately 9 m long [27] juvenile whale started the capture, as indicated by the start of the buzz, at a median of 24 m from the prey. This is substantially further than the one body length at which beaked whales switch to buzzes [28] and four times longer in body lengths compared to small captive odontocetes, which initiate the buzz phase at about half a body length from the target [29,30]. These differences might relate to the size-related manoeuvrability of these predators relative to their prey [31,32], such that the slow-turning sperm whale needs to know the detailed movement of the prey, obtained by fast clicking in the buzz, at a greater range to plan its own movements. Alternatively, it might be an effect of the high source levels of sperm whales’ usual clicks [4], allowing for the reduction of source level in the buzz (greater than 20 dB [33]) at a greater range while maintaining audible prey echoes and reducing forward masking. As sperm whales maintain a stable high output level during the search and approach phase and markedly reduce output levels in the buzzes, the large distance from the prey at which the buzzes are initiated (24 m) leads to a transmission loss of more than 25 dB from geometric spreading. Hence, the maximum received level experienced by prey is well below experimental sound levels shown to have no effect on squid swimming ability [34]. We therefore join Fais et al. [35] in rejecting the acoustic debilitation hypothesis.

In conclusion, by deploying a sound recording tag on the very tip of the gigantic nose of a sperm whale, we have for the first time recorded prey echoes to show that the powerful sound-generating nasal complex of sperm whales enables long-range echolocation and prey selection, but not acoustic

---

**Figure 2.** Echograms from the first deep dive. (a) Prey echo prior to and within a buzz with normalized jerk signal for the same time interval in (b). (c) Echogram with medium echo density and a seafloor echo. (d) Echogram with high echo density. (e) Echo stream of a school of organisms. Isolated red spots in (c,d,e) represent echolocation clicks from other sperm whales.
References


33. Madsen PT, Wahlberg M, Mahi B. 2002 Male sperm whale (Physeter macrocephalus) acoustics in a high-

