JASA ARTICLE

Latencies of click-evoked auditory responses in a harbor porpoise exceed the time interval between subsequent echolocation clicks

K. Beedholm,[1] M. Ladegaard,[1] P. T. Madsen,[1] and P. L. Tyack[2,a)]
[1]*Zoophysiology, Department of Biology, Aarhus University, Aarhus C 8000, Denmark*
[2]*School of Biology, University of St Andrews, St Andrews KY16 9ST, Scotland*

ABSTRACT:
Most auditory evoked potential (AEP) studies in echolocating toothed whales measure neural responses to outgoing clicks and returning echoes using short-latency auditory brainstem responses (ABRs) arising a few ms after acoustic stimuli. However, little is known about longer-latency cortical AEPs despite their relevance for understanding echo processing and auditory stream segregation. Here, we used a non-invasive AEP setup with low click repetition rates on a trained harbor porpoise to test the long-standing hypothesis that echo information from distant targets is completely processed before the next click is emitted. We reject this hypothesis by finding reliable click-related AEP peaks with latencies of 90 and 160 ms, which are longer than 99% of click intervals used by echolocating porpoises, demonstrating that some higher-order echo processing continues well after the next click emission even during slow clicking. We propose that some of the echo information, such as range to evasive prey, is used to guide vocal-motor responses within 50–100 ms, but that information used for discrimination and auditory scene analysis is processed more slowly, integrating information over many click-echo pairs. We conclude by showing theoretically that the identified long-latency AEPs may enable hearing sensitivity measurements at frequencies ten times lower than current ABR methods. © *2023 Acoustical Society of America*. https://doi.org/10.1121/10.0017163

(Received 14 July 2022; revised 6 December 2022; accepted 17 January 2023; published online 7 February 2023)
[Editor: Colleen Reichmuth] Pages: 952–960

I. INTRODUCTION

Echolocating toothed whales hunt and navigate by acoustically probing their surroundings using short, predominantly ultrasonic, clicks and analyzing the temporal and spectral properties of the returning echoes (e.g., Au, 1993). Harbor porpoises (*Phocoena phocoena*) click at rates between 10 and 500 Hz, corresponding to inter-click intervals (ICIs) of 2–100 ms (Ladegaard and Madsen, 2019). Their echolocation clicks consist of directional (MacAulay *et al.*, 2020) narrowband high-frequency (NBHF) clicks centered at 129–145 kHz and a half power bandwidth of 6–26 kHz (Villadsgaard *et al.*, 2007). When toothed whales, including porpoises, echolocate on a distant target, the time interval between consecutive clicks (ICI) is usually several tens of ms longer than the two-way travel time (TWTT) delay between the time of click production and the time of the returning echo from the target (Morozov *et al.*, 1972; Au *et al.*, 1974; Ladegaard and Madsen, 2019). One simple interpretation of this so-called lag time suggests that it provides enough time to process echo information before toothed whales produce the next click (Au, 1993). Verfuß *et al.* (2009) describe this process as a sequential pulse mode of echolocation, and they suggest that as porpoises approach a target, lag times of 20 ms or more provide enough time for pulse mode echolocation.

An echolocating predator can adjust its clicking rate to update information about prey location and movement (Vance *et al.*, 2021). However, if two clicks are emitted before the echo from the first click has returned, there may be ambiguity about which click the echo came from and, therefore, about target range (Madsen and Surlykke, 2013). Most toothed whales mitigate this potential range ambiguity problem by waiting to click until echoes of interest from the previous click have arrived plus an additional lag time (e.g., Morozov *et al.*, 1972). Some lag time after the echo is received is required to prevent the outgoing next click from masking echoes from the previous click, so-called backward masking. Reducing the risk of range ambiguity provides a different explanation for the observation that echolocating odontocetes wait for echoes of interest before producing the next click.

During the final stages of target approach, echolocating odontocetes accelerate their click rate, with porpoises producing clicks at intervals around 1.5 ms (Verfuß *et al.*, 2009). Au (1993) suggests that at short ranges with such short ICIs, odontocetes may process several echoes at a time. Following Nordmark (1960), Verfuß *et al.* (2009) suggest that short delays between click and echo might be perceived as a pitch proportional to the inverse of the delay. As an echolocating porpoise approaches a close target, the decreasing click-echo

a)Also at: Biology Department, Woods Hole Oceanographic Institution, Falmouth, MA 02543-1050, USA. Electronic mail: ptyack@whoi.edu

https://doi.org/10.1121/10.0017163

delays could be perceived as increasing pitch, allowing the porpoise to estimate decreasing range to the target. Ridgway (2011) mentions the possibility that rapid processing in the odontocete brainstem/midbrain, which is adapted for rapid transmission of acoustic information, might enable them to estimate target range and time the production of the next click within lag times of 20 ms or more. However, Ridgway (2011) argues that 20 ms is perhaps an order of magnitude too short for the auditory system to transduce the echo, transmit the information to relevant brain networks, decide timing for the next click, and generate the motor commands to produce the click. He suggests that slower cortical processes could change pulse rates based on changes in the echolocation scene provided by series of echo returns, and he suggests that studying auditory evoked potentials (AEPs) with latencies of 60–200 ms would help test these hypotheses.

In this study, we use AEPs resulting from passive click stimulation to explore the neurophysiological consequences of these possible timing and processing scenarios. In humans, AEPs that occur within 10 ms of onset of a stimulus are often called short-latency auditory evoked potentials (SLAEPs) (Legatt *et al.*, 1986). These originate in the auditory nerve and brainstem in mammals and so are also known as auditory brainstem responses (ABRs). Response peaks occurring with latencies from 10 to 50 ms are defined as mid-latency auditory evoked potentials (MLAEPs), and AEPs with latencies >50 ms are dubbed long-latency auditory evoked potentials (LLAEPs) (Luck, 2014). Studies of these longer-latency evoked potentials have added to our knowledge that different regions of the brain process auditory inputs over different time scales to yield different outputs. Processes that require precise timing of arrivals at each ear take place in brainstem circuits that are specialized for rapid temporal processing (Yin *et al.*, 2019). Cortical neurons respond more slowly, integrating information across frequency and longer time intervals to perform downstream processing of features such as the recognition of auditory patterns and the formation of auditory objects that change more slowly (Asokan *et al.*, 2021).

Recent studies of neural processing of both hearing and echolocation in toothed whales rely on non-invasive AEPs measured with external electrodes embedded in suction cups (e.g., Nachtigall *et al.*, 2007). Most AEP studies in marine mammals focus on the SLAEPs with onsets within 1 ms and peaks at delays of 1.5–4.5 ms (Ridgway *et al.*, 1981; Supin and Popov, 1995; Popov and Supin, 2007), and there are relatively few studies of longer-latency evoked potentials. Popov *et al.* (1986) implanted electrodes into the auditory cortex of an awake, restrained harbor porpoise (*P. phocoena*) and found evoked potentials arising several tens of ms after the onset of noise bursts, with a negative peak at latencies of 15–20 ms. Woods *et al.* (1986) implanted wire electrodes over the skull of bottlenose dolphins (*Tursiops truncatus*) and detected AEPs to click stimuli with peaks at about 25, 200, and 450 ms. Also, in dolphins, a few recent papers have described long latency potentials recorded with suction cup electrodes. Hernandez *et al.* (2007) report peaks in *Tursiops* responses at latencies of 50, 75, and 150 ms and Schalles *et al.* (2021) report peaks around 25, 50, and 75 ms. However, LLAEPs have not previously been described for harbor porpoises.

Here, we explore click-evoked MLAEPs and LLAEPs using non-invasive recordings in a trained harbor porpoise to gain insight into the time course of processing information from clicks. It is difficult to study mid- and long-latency responses to clicks during active echolocation because porpoises will almost always click again before the response is over (which is the topic of this study). Hence, in this study, we used an approach with passively presented transient stimuli. Specifically, we test the hypothesis of Au (1993) and Verfuß *et al.* (2009) that toothed whales have a sequential pulse mode for distant targets, where the time lag between detection of an echo and emission of the next click is long enough to process received echoes and generate the motor response that emits the next click.

A primary use for SLAEPs in toothed whales is to obtain audiograms (e.g., Popov and Supin, 1990; Mooney *et al.*, 2012; Yuen *et al.*, 2005). AEP studies require averaging over hundreds of responses synched to similar stimuli, so it is less time consuming to study responses to stimuli presented at short latencies than long. SLAEP responses of the auditory nerve and brain stem also vary less across individuals than cortical responses that vary due to a variety of factors (Picton *et al.*, 1977). Audiograms derived from SLAEPs match relatively well with behavioral audiograms except at frequencies below some 10 kHz, where a loss in phase synchrony in neural responses precludes reliable estimation of hearing sensitivity (Finneran *et al.*, 2016; Houser and Finneran, 2006). Recognizing the limitations of longer-latency AEPs for audiometry (Supin *et al.*, 2001), we therefore also explore and discuss their potential for measuring hearing thresholds at frequencies below those accessible using SLAEP.

II. MATERIALS AND METHODS

A. Setup and procedure

The experimental subject was one 21-year-old captive female harbor porpoise (*P. phocoena*) named Freja. The subject was housed and trained in a net pen at Fjord and Bælt (Kerteminde, Denmark) under permits SN 343/FY-0014 and 1996–3446–0021. During recordings, Freja was trained to rest for 60 s on a bite plate 60 cm below surface while fitted with two gold-plated electrodes [Grass (West Warwick, RI), Ø = 10 mm] in custom-made silicone suction cups: one behind the blowhole and the reference close to the dorsal fin [following Beedholm *et al.* (2006) and Popov and Supin (1990)]. An identical gold-plated electrode was placed off the animal close to the bite plate to serve as common ground. Electrode inputs were amplified by 60 dB and bandpass filtered between 0.1 and 10 kHz (two poles) using a Grass P55 differential amplifier with 50 Hz rejection. Stimuli were delivered passively (not triggered by echolocation clicks) at intervals of 1013 ms to allow for reception of

long-latency potentials before the next stimulus and to avoid averaging on time epochs divisible by 50 Hz noise (20 ms period) from line voltage sources in the vicinity. Each trial lasted about 50–60 s, so responses from about 50–60 presentations were collected during each trial. The click stimuli were produced by a Brüel & Kjær (Nærum, Denmark) 8105 hydrophone placed 15 cm in front of the subject's jaw. The stimuli were delivered to the transmitting hydrophone from the digital-to-audio converter (DAC) channel of a USB-6356 multifunction box (National Instruments, Austin, TX) with a buffer amplifier (based on an LM386 chip, LC Tech, Shenzen, China) to avoid the influence of nonlinear effects close to the resonance frequency of the piezoceramic element of the hydrophone. A Reson (Slangerup, Denmark) TC4013 hydrophone, mounted on the bite plate, served to monitor received levels of AEP stimuli and click emission of the stationed porpoise. The output from this hydrophone was amplified by 30 dB and bandpass filtered (three poles) between 10 and 250 kHz using a Reson VP 2000 conditioning box. Data collection from the three electrodes and the hydrophone was performed with four 16-bit ADC channels of the same National Instruments multifunction box, running off a common clock at 500 kHz sampling rate (16 bit). All parts of the setup ran on battery with no connections to mains power; the common (wet) ground for all instruments was the reference electrode for the AEP recording.

Inputs and outputs from the multifunction box were controlled by custom programs written in LabVIEW. The animal performed in 3–5 trials with the same electrode placement in a session and participated in 2–3 sessions per day after an initial training period of 4 weeks.

The duration of the stimulus waveform influences the timing and phase of peaks in the response waveform. Since AEP responses roughly follow the stimulus envelope (Dolphin *et al.*, 1995), if the stimulus is long compared to the click-evoked AEP, the response is smeared out in time, and the effective result becomes the neural onset and offset responses elicited by the stimulus (Supin and Popov, 1995). The present study focuses on responses that could have been elicited by echoes from echolocation clicks. We, therefore, chose stimuli relatively similar to the echo produced when a porpoise echolocation click reflects off a target. Porpoise clicks have a variable duration of 50–100 μs and most energy between 110 and 150 kHz (Møhl and Andersen, 1973). When porpoise clicks ensonify most targets, the echoes have a complex structure involving reflections from multiple features (e.g., Au *et al.*, 2009). Our echo-like stimulus was made from two consecutive 50 μs chirps spanning 100–150 kHz each with a received peak-to-peak amplitude of 124 dB re 1 μPa (Fig. 1). The stimulus was spectrally 10 kHz lower than a typical echo from a target to make it stand out somewhat from natural echoes occurring from echolocation (spectrum part of Fig. 1). The received level is similar to levels that were accepted by a porpoise as an echo from a simulated echo generator (Beedholm *et al.*, 2006) and is also near the steepest part of the amplitude-response curve for the porpoise SLAEP signal to ensure a good

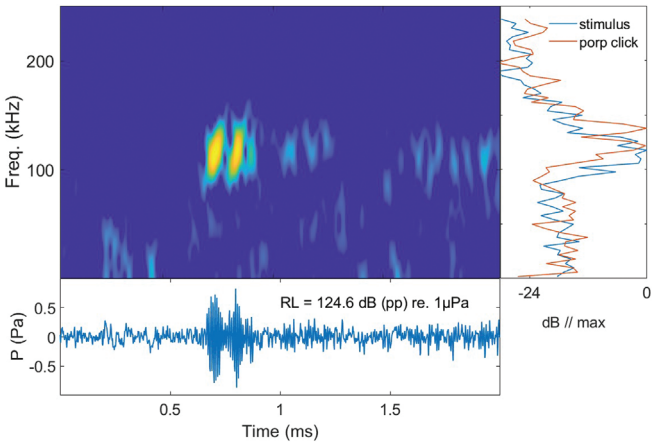

FIG. 1. AEP stimulus presented passively to the porpoise during stationing. Shown are spectrogram, waveform, and spectrum (in blue) of the two-click stimulus used in this study as recorded on the observation platform during experiments with a TC4013 Reson hydrophone. This recording is high-pass filtered at 15 kHz. The actual peak of the unfiltered recording is at a lower frequency. Also included in the spectrum is a click emitted by the porpoise during data collection (in red).

signal-to-noise ratio (SNR) without being close to saturation (Smith *et al.*, 2021). The level would correspond to echoes received 1.5 m from a 50.8 mm solid spherical target, when ensonified with clicks with short latencies (SLs) of 156 dB peak-to-peak (p.p.) re 1 μPa, close to Freja's mean source level at such a range in Ladegaard and Madsen (2019). For comparison between AEP response latencies and ICIs, we used ICIs from another study (Ladegaard and Madsen, 2019) where the same subject approached a standard target in the same pool. The clicks were detected on a DTAG4 (www.animaltags.org/) during 50 trials to include search, approach, and buzz phase clicks. Figure 2 shows the placement of the tag during these recording sessions.

B. Data processing

All off-line data processing was done in MATLAB (version 2016a, MathWorks, Natick, MA). The electrophysiological signals were filtered off-line in the range of 10–2000 Hz, followed by notch filtering at integer multiples

FIG. 2. Illustration of D-Tag placement during collection of echolocation parameter data in a target approach experiment (photo taken during reward phase). See Ladegaard and Madsen (2019) for details. Photograph courtesy of Monika Dyndo (Fjord & Bælt).

https://doi.org/10.1121/10.0017163

of 50 Hz using a zero phase fast Fourier transform (FFT) filter on the complete trial, which allows for very narrow reject bandwidths of noise from 50 Hz line voltage. FFT-based filtering of this kind results in onset and offset artifacts [Fig. 3(F)]. Therefore, the first and last seconds of the filtered recording were not processed further. Setting the high pass filter to 10 Hz may attenuate some of the long-latency AEPs we measured but was required to limit low frequency noise in the recording.

C. Electrocardiographic (ECG) transient removal

ECG signals superimposed on the AEP data constituted a dominant source of interference (Schalles *et al.*, 2021) that deserves special attention in this context. These signals do not normally interfere with the extraction of SLAEP (ABR), because they are lower in frequency than the ABR signals, but with the long latency AEPs at low frequencies that we sought to study here, responses are difficult to detect without a dedicated effort to remove ECG signals. Since this is essential to AEP studies using low frequency components in small cetaceans, we describe the method in detail here. First, to reliably detect the ECGs, we filtered another copy of the ca. 1-min-long raw electrophysiological trace with a filter that would emphasize the ECG signals (here fourth-order low-pass filter with –3 dB cut-off at 20 Hz). We then applied peak detection to locate all ECG events in the trace.

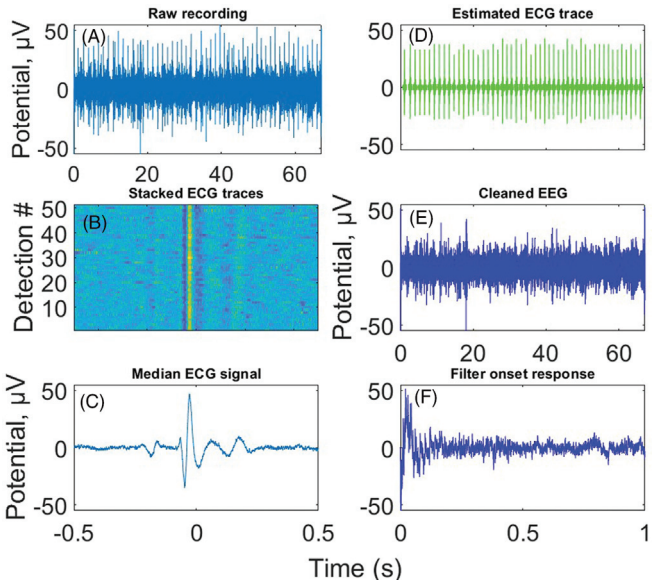

FIG. 3. Removal of ECG artifacts from bioelectric traces. (A) The raw EEG trace from a single trial lasting about 65 s. (B) Matrix of aligned, isolated ECG signals from the trace depicted in (A). (C) Median of the ECG signals in (B). This almost noise-free signal serves as the template for the construction of the heart rate trace for the whole trial. (D) The estimated ECG artifact part of the trace formed by suitably scaling copies of the template at each detection position so as to arrive at a minimum residual energy value when subtracted. This panel uses the time scale of the panel below. (E) The difference between the raw waveform (A) and the estimated ECG trace (D). At the start and end of this trace, onset and offset filtering artifacts from the FFT-based comb filter (50, 100, 150 Hz…) are evident. (F) The first second from the trace in (E), showing onset artifact of filtering. Stimulus presentations began more than 1 s after the trace started.

The average heart rate was usually about 40–45 beats/min, so each 1-min-long trial contained 40–45 ECG signals. Using the detected events, we created a matrix consisting of unfiltered ECG waveforms, windowed around the sample of detection and time-aligned so that the ECG signals were in phase [Fig. 3(C)]. From the stacked ECG signals, we took the median (Özdamar and Kalayci, 1999), which then served as the ECG waveform template for that particular trial in the next steps [Fig. 3(E)]. At each point in the original unfiltered waveform where an ECG detection was made, we determined the best delay (using cross correlation) and the scaling factor of the template that—when subtracted—resulted in the lowest residual root mean square (rms) amplitude in the waveform [Fig. 3(D)]. The now suitably scaled and time shifted template was subtracted from the waveform at each ECG location [Fig. 3(E)].

D. AEP extraction and statistical approach

From the simultaneous recording of the transmitted click stimuli and the AEP responses to these stimuli, we detected the stimuli and sorted the responses into matrices of time-aligned AEP following stimulus onset. We then computed the median of these responses (the median is a better estimate of central tendency than the mean as it reduces the influence of outliers) (Supin and Popov, 2007).

To determine whether a given peak in the averaged waveform was significant or just residual noise, a Monte Carlo permutation test was made on both raw and Hilbert transformed copies of the waveforms in the matrix of responses. Each run randomly drew half of the waveforms. The test counted the number of times (out of 10 000 trials) the observed response sample was either (1) larger than the absolute value of the median of samples drawn from anywhere in the matrix or (2) larger than the absolute value of the median of the randomly drawn Hilbert transformed data points from anywhere in the matrix of Hilbert transformed data. Since we are testing two-sided (both positive and negative excursions) in both raw data and in the Hilbert transform, the significance threshold becomes $\alpha/4$, so for the 5% significance level, the threshold for a data point being significant is that fewer than 1.25% of the averages of the randomly drawn values were higher than the average drawn from the traces that were time-aligned to the stimulus onset.

III. RESULTS AND DISCUSSION

A. Implications for neural processing of echoes

By averaging 2022 stimulation epochs, each following the presentation of a simulated echo of 124 dB re 1 μPa (p.p.) with 1013 ms stimulus intervals, we not only evoked standard ABR responses at latencies of 3–5 ms and another SLAEP at 7–10 ms, referred to as n8, but also recorded statistically highly significant click-evoked mid- to long-latency potentials in our porpoise subject [Fig. 4(A)]. Prominent peaks linked to stimulus presentation time appear at mid-latencies of 25–30 ms, referred to as n25, and long latencies of 70–100 ms, referred to as n70 and p90 [Fig. 4(A)]. Another

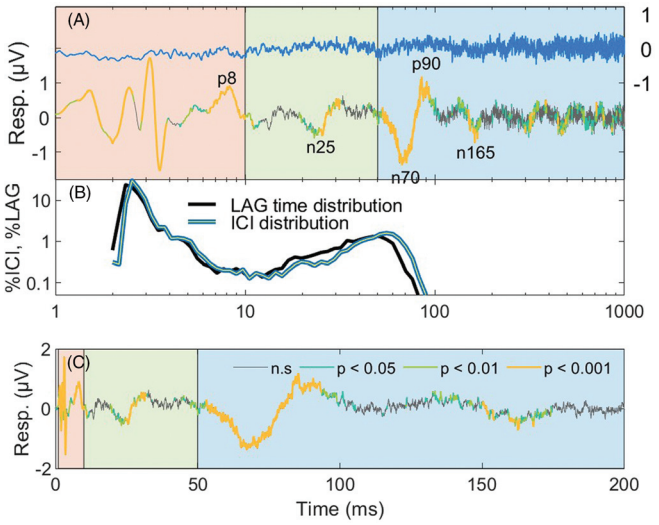

FIG. 4. AEP responses to echo stimuli, distribution of ICIs, and lag times from a porpoise. (A) AEP as a function of time (ms) following stimulus presentation onset. Stimulus interval was 1013 ms. Significance level was determined at each sample as a Monte Carlo permutation test for either the real or the imaginary part of the analytical AEP signals. The time segment assigned to SLAEP (ABR) is shown on a light red background, MLAEP (10–50 ms) in light green, and LLAEP (>50 ms) on powder blue. In blue is shown an average of the same data matrix, where each trace has been moved by a random amount (0–1013 ms), so that responses are no longer time-aligned to stimulus onset. The distribution against which the sample-by-sample significance was tested was built from a collection of 10 000 such scrambled data matrices. (B) Density distribution of ICIs produced by Freja during a target approach experiment and the lag time defined as ICI − TWTT, where TWTT is the two-way travel time from porpoise to target and back to the porpoise. (C) Same as (A) but on a linear time scale up to 200 ms to emphasize the relative amount of signal energy in the segment from 50 to 100 ms for comparison with the ABR segment (light red). n.s., not significant.

significant negative peak occurs at 165 ms, referred to as n165, and additional, weaker but still statistically significant peaks at latencies between 300 and 800 ms [Fig. 4(A)]. The blue trace on the top of Fig. 4(A) shows an example of the median of the data set, when the timing of each epoch is randomized with respect to the stimulus. From 800 to 1013 ms, the number of significant events detected in the AEP data is as expected for data with random timing with respect to the stimulus, 5%, 1%, and 0.1% of data points being significant at the 0.05, 0.01, and 0.001 significance levels, respectively, so none of these are considered statistically significant.

AEPs with latencies of 15 ms to peaks have been measured directly in the auditory cortex of porpoises (Popov *et al.*, 1986) and as short as 9–14 ms in bottlenose dolphins (Supin *et al.*, 1978), leading us to surmise that the p90, n165, and the later peaks are all generated in the porpoise cortex and not at the midbrain level.

Here, we test the sequential pulse mode hypothesis by comparing the latencies of AEP peaks from a porpoise to her ICIs and lag times recorded during echolocation. Figure 4(B) shows the probability density function of ICIs used by Freja during a target approach experiment (Ladegaard and Madsen, 2019). This includes search and approach phases with ICIs between 12 and 60 ms and buzz phases with ICIs down to 1.6 ms. An acoustic recording tag on the porpoise (Fig. 2) recorded each outgoing click and returning echoes, enabling measurement of the TWTT and the lag time = ICI – TWTT. Figure 4(B) shows that all lag times and 99.6% of ICIs are shorter than the 90 ms latency of the p90 AEP [Fig. 4(A)]. Thus, the p90 and longer latency potentials that are evoked by our click stimuli show that the porpoise is still processing information from the previous click or its echoes by the time it produces the next click. This leads us to reject the hypothesis that toothed whales complete processing echo information arising from each click before emitting the next click during the lag time given by the ICI − TWTT to the target.

Under the range processing hypothesis of Verfuß *et al.* (2009), perhaps a fast neural network, consistent with peaks of shorter latencies in the early section of Fig. 4(A), can process click-echo delays to estimate target range at early stages of auditory processing and generate the motor activity in time for controlled production of the next click. If (1) porpoises engage such peripheral rapid neural networks in early phases of auditory processing to estimate click-echo delays, (2) these early phases rapidly inform the motor system when to produce the next click, and (3) the motor system can produce the requested click within the lag time, then this could satisfy this alternative sequential pulse hypothesis despite our observation of later phases of processing of echo information. Regarding topics 1 and 2, in echolocating bats, neurons in the midbrain are tuned to the delay between pulse and echo (Dear and Suga, 1995). This information projects not only to the auditory cortex, but also to other midbrain pathways that project to vocal production nuclei. These latter pathways may support rapid adaptation of motor activities such as clicking in response to changes in target range (Sinha and Moss, 2007). However, although latencies for some motor responses stimulated by auditory stimuli through brainstem networks can be as short as 5–10 ms (e.g., acoustic startle response in rats; Koch and Schnitzler, 1997), similar measures for porpoises are closer to 60 ms (Elmegaard *et al.*, 2021). Vance *et al.* (2021) used the same porpoise, Freja, as was used in this study, to show that the latency between target movement and ICI adjustments is 50–100 ms. This latency is longer than the majority of ICIs from this porpoise, demonstrating that, as suspected by Ridgway (2011), porpoises cannot respond via auditory processing and vocal-motor control rapidly enough to changing click-echo delays for all three conditions of the range processing hypothesis to be met.

Work showing that motor regions interact with auditory areas to produce predictive timing for entraining to a rhythm (Merchant *et al.*, 2015) suggests a simple neural model where toothed whales generate a click series using interactions between motor and auditory cortical and subcortical areas operating at different time scales. Accordingly, building upon Ridgway (2011), we propose that rather than processing echo information from each click before deciding on the timing for the next click, toothed whales simultaneously process echo information from a past sequence of

https://doi.org/10.1121/10.0017163

echolocation clicks to plan a click rate. At the same time as echo information is being processed, the motor system must use auditory input and expectations of changes in target range and three-dimensional (3D) location to plan a future sequence of clicks along with the porpoise's 3D movement and orientation of the sonar beam. As has been shown for bats, this likely involves the formation of predictive models of expected sensory outcomes of the animal's own movement and that of the target, planning echolocation and movement based upon the predictive model and using feedback to update the model with new information as it arrives (Salles *et al.*, 2020). Au (1993) and Verfuß *et al.* (2009) propose two distinct modes of echolocation, the sequential pulse model for long ranges and ICIs, and a different model for close range, where several echoes must be processed at the same time ("pitch mode"). The data presented here [along with Vance *et al.* (2021)] would suggest that there is no need for two separate modes of operation, since in neither case will processing be complete in time for the next click to be emitted, so that both patterns of echolocation require parallel processing of sequences of click-echo pairs linked with planning of click rates.

B. Potential use for low frequency responses in audiometry

In addition to allowing us to test the sequential pulse mode hypothesis, we note that the existence of strong long-latency MLAEP (>50 ms) responses [Fig. 4(C)] may enable AEP-based estimation of low frequency hearing sensitivity. Three different kinds of stimuli are used for AEP tests of hearing: broadband clicks, narrowband tone bursts, and amplitude modulated pure tones. MLAEP responses to click stimuli are so strong and reliable that they were among the first auditory evoked potentials averaged from scalp electrodes in humans (Geisler *et al.*, 1958). More narrowband tone bursts enable studies of sensitivity at different tone frequencies, and Picton *et al.* (1977) (p. 99) suggested that "middle latency evoked components provide perhaps the best means of evaluating thresholds at a variety of frequencies." Galambos *et al.* (1981) showed that humans exposed to clicks generate LLAEPs similar to a sine wave with a period of about 25 ms. This yielded a peak response amplitude when clicks were presented at 40 Hz, corresponding to the period of the MLAEP in humans. When short bursts of tones were used to study auditory thresholds at specific frequencies, these 40 Hz AEPs yielded estimates that were close to behavioral thresholds.

Figure 4(C) shows the AEP responses we recorded from our porpoise subject using a linear time base. This plot reveals that the response from 60 to 100 ms also is similar to a sine wave with a period of about 40 ms. The Galambos *et al.* (1981) results suggest that repeating our short stimulus at 25 Hz would likely yield a maximal evoked response.

Early AEP research repeated presentations of short duration stimuli. Today many AEP audiometry studies make use of the observation that AEP responses of mammals follow the envelope of a stimulus (Dolphin and Mountain, 1992). Sinusoidally amplitude modulated (SAM) stimuli use this envelope-following response to estimate audiograms (e.g., Houser and Finneran, 2006). If a carrier frequency f_c is amplitude modulated with a modulation rate of f_m, the strength of the response evoked at f_c can be estimated by measuring the energy of the spectrum of the AEP at f_m. The strength of this envelope-following response varies depending upon f_c and f_m. SAM stimuli are designed to have a value of f_m that produces a strong AEP. f_m is then held constant, and the AEP is measured for different levels of each carrier frequency f_c of interest.

Initial studies using SAM to estimate hearing sensitivity in humans used these mid-latency responses to modulation rates near 40 Hz, similar to the optimal repetition rates of short stimuli. However, studies in infants use higher modulation frequencies of 75–110 Hz because they can reliably be recorded in infants and are less sensitive to behavioral context (Picton *et al.*, 1977). The auditory system of echolocating toothed whales shows strong envelope-following responses at modulation frequencies f_m with peaks close to the peak of the spectrum of the SLAEP (Dolphin *et al.*, 1995). The spectral peaks of odontocete SLAEPs are at higher frequencies than seen in other species, as their hearing system is equipped with large diameter axons and large synapses, which accelerate neural processing. ABR-based audiometric studies of odontocetes use SAM stimuli at the peak of the spectrum of the SLAEP (ABR) response to a click, in the range from 600 Hz to 1.4 kHz depending on the species (for porpoises, the range is from 1.1 to 1.4 kHz) (Supin *et al.*, 2001; Linnenschmidt *et al.*, 2013). However, this choice presents problems for studying hearing at frequencies below ~10 kHz. As illustrated in Fig. 5, if f_c is low enough that the critical band of hearing at f_c no longer includes the side bands, this means that the sidebands fall outside the auditory filters, thus, missing the modulation altogether (Kuwada *et al.*, 1986; Houser and Finneran, 2006). Auditory filter bandwidths have not, to our knowledge, been reported below 5 kHz for toothed whales. The equivalent rectangular bandwidths in humans at frequencies of 100, 200, 400, and 800 Hz are 36, 47, 87, and 147 Hz (Moore *et al.*, 1990). Since we do not know the critical band at low frequencies for toothed whales, it would be prudent to make sure that $f_c > 10\,f_m$. This would ensure that $2\,f_m$ would be a value lower than any of the low frequency human filter bandwidths. This suggests that $f_m = 1000$ Hz cannot support reliable estimation of hearing sensitivity below 10 kHz. Following the same logic, Houser and Finneran (2006) reduced f_m to 500 Hz to test $f_c = 5000$ Hz.

Understanding low frequency hearing is important for basic research and also has important applications for cetaceans, which, despite high-frequency specializations, can hear frequencies down to below 100 Hz. Most of the energy in vessel noise, air-gun pulses, LF sonar sounds, and pile-driving transients lies 1–2 orders of magnitude below 10 kHz. Measuring temporary changes in hearing thresholds after exposure to intense sound relies upon AEP methods for rapid measurement of thresholds (Finneran *et al.*, 2005).

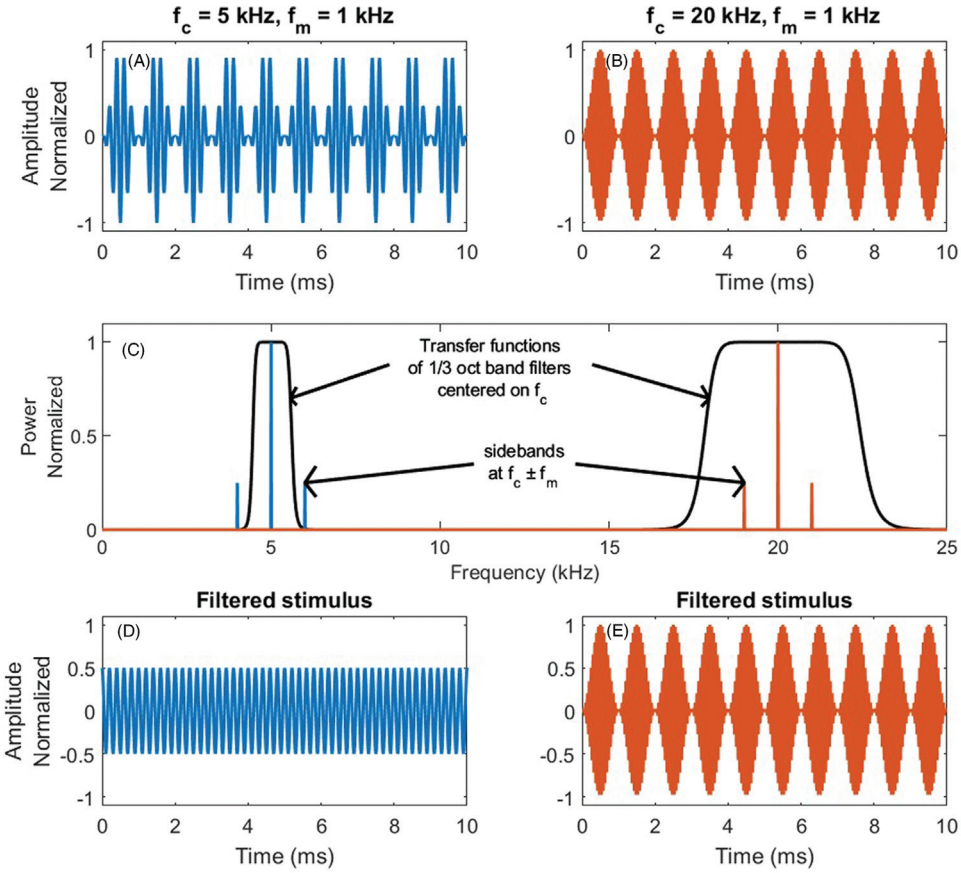

FIG. 5. The relationship between modulator, carrier, and auditory critical bands in a theoretical SAM-stimulus experiment. (A) and (B) show time domain representations of SAM stimuli with $f_c = 5$ kHz [(A), blue] and $f_c = 20$ kHz [(B), red]. Both are amplitude modulated by a 1 kHz raised cosine wave. (C) shows frequency representations (power spectra) of the same signals together with the transfer functions of 1/3 octave filters (black) centered on the carrier frequencies of each stimulus. The auditory filters are constant Q, so for low f_c values [(A) and (C), left], the auditory filter will not cover the full spectrum of the stimulus, and the input to the sensory system is then not expected to follow the amplitude changes of the modulated signal. (D) The center frequency at 5 kHz is low enough that the sidebands are outside of the auditory filter, so that the filtered output is not amplitude modulated. (E) Only when the center frequency is sufficiently high that the sidebands lie within the auditory filter [(C) (right) and (E)] does the output from the filter follow the amplitude modulation of stimulus.

Therefore, measuring temporary threshold shifts after exposure to these low frequency sounds requires AEP methods suitable for lower frequencies that may not be reliably assessable using methods based on early responses (ABR/SLAEP).

Here, we have demonstrated strong MLAEP/LLAEP responses to clicks in a porpoise with period close to 40 ms [Fig. 4(C)], corresponding to a SAM frequency of 25 Hz [see Fig. 6(B)]. Previous studies of AEP responses to amplitude modulation rate in cetaceans have not tested frequencies below 50 Hz (Supin and Popov, 1995; Dolphin *et al.*, 1995). The strong MLAEP response found here [Fig. 4(C)] suggests that repetition or modulation rates of 25 Hz (1/40 ms) may produce a strong signal (Galambos, 1981). To explore this idea, we studied the frequency-dependent SNR of the click-evoked AEP from our porpoise subject to map the best candidate(s) for a modulation frequency. Figure 6(A) shows in blue the spectrum of the coherently averaged AEP that was shown as a time series on a linear time scale in Fig. 4(C). This is compared to the average of the spectra of each AEP time series shown in orange. This latter orange spectrum is incoherently averaged by averaging the spectra of each time series rather than by averaging the times series coherently and then taking the spectrum. If there was no AEP signal, these spectra would be similar, except that the noise spectrum would be reduced by sqrt(2022) = 33 dB, because of the reduction in random noise with averaging. The AEP signal is not random and, therefore, does not reduce with averaging. We calculated the frequency-dependent SNR by subtracting the orange noise spectrum from the blue AEP spectrum. Figure 6(B) shows that the SNR at 20 and 40 Hz is greater than the SNR at 1.2 kHz, because the energy of the p90 complex is higher than the transient 1.2 kHz ABR. Note that the SNR at 50 Hz is close to 0, which is consistent with the lack of response at this rate reported for *Tursiops* by Supin and Popov (1995). As illustrated in Fig. 6, this longer, lower-frequency response peak can contain multiple cycles of much lower carrier frequencies, in principle down to a few hundred Hz. We, therefore, suggest a 20 or 40 Hz modulation frequency to extend by a decade downward the range of frequencies in the audiogram that can be tested using fast AEP techniques.

https://doi.org/10.1121/10.0017163

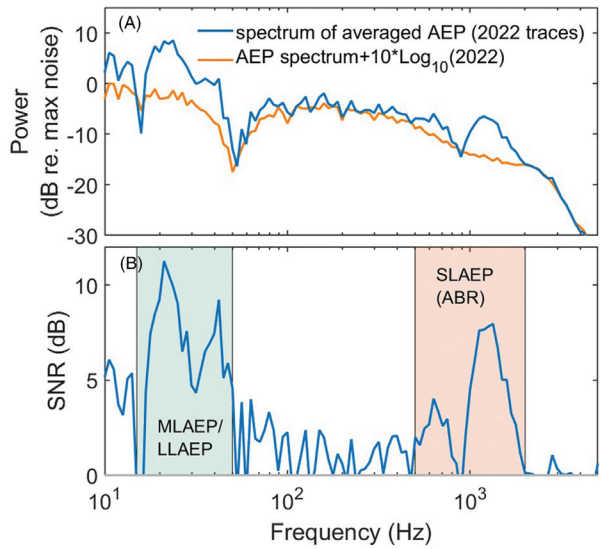

FIG. 6. Frequency-dependent SNR improvement with averaging. (A) Orange line, average 1/12 octave spectrum of all recorded responses with $10\log_{10}(2022)$ added to account for the expected decrease in power with incoherent averaging. Blue line, 1/12 octave spectrum of the coherently averaged response. Both lines are normalized to the maximum level of the incoherent trace. The 0.7–1.5 kHz range of the traditional ABR response has higher spectral level than expected from random noise alone. A region in the frequency range of 10–50 Hz also has the same good SNR, corresponding to the mid- to long-latency potentials documented in Figs. 4(A) and 4(C). The difference between the two spectra highlights the existence of two regions with high SNR that are suitable for AEP studies, the usual 0.5–2 kHz and 20 and 40 Hz, corresponding to the components at long latencies in the AEP traces. (B) Light red and green boxes demarcate the two regions with good SNR for AEP studies.

In summary, the presentation of passive stimuli for AEP data for latencies out to 1 s has demonstrated that processing of transient acoustic signals, such as echoes, takes place with a cascading series of stimulus-associated peaks over a time span out to 165 ms, long enough that it necessitates parallel processing of sequences of echolocation clicks to drive click rates that are appropriate for changes in the acoustic scene. We also found that the AEP trace obtained this way suggests that there could be a potential for utilizing slow cerebral responses in obtaining an AEP-based threshold to frequencies well below 10 kHz.

ACKNOWLEDGMENTS

We are grateful to trainers J. Larsson, F. Johansson, M. Kjølby, and J. H. Kristensen at Fjord & Bælt (Kerteminde, Denmark) for assistance in conducting the experiments. The authors declare no competing or financial interests. Author contributions were as follows: conceptualization: K.B., P.T.M., and P.L.T.; methodology: K.B.; software: K.B.; validation: K.B. and M.L.; formal analysis: K.B. and M.L.; investigation: K.B.; resources: P.T.M.; data curation: K.B. and P.T.M.; writing—original draft: K.B., P.T.M., and P.L.T.; writing—review and editing: K.B., M.L., P.T.M., and P.L.T.; visualization: K.B.; supervision: P.T.M.; project administration: P.T.M.; funding acquisition: P.L.T. This project was funded by U.S. Office of Naval Research Grant Nos. N00014-18-1-2062 and N00014-20-1-2709.

Asokan, M. M., Williamson, R. S., Hancock, K. E., and Polley, D. B. (**2021**). "Inverted central auditory hierarchies for encoding local intervals and global temporal patterns," Curr. Biol. **31**(8), 1762–1770.
Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer, New York).
Au, W. W. L., Branstetter, B. K., Benoit-Bird, K. J., and Kastelein, R. A. (**2009**). "Acoustic basis for fish prey discrimination by echolocating dolphins and porpoises," J. Acoust. Soc. Am. **126**, 460–467.
Au, W. W. L., Floyd, R. W., Penner, R. H., and Murchison, A. E. (**1974**). "Measurement of echolocation signals of the Atlantic bottlenose dolphin, *Tursiops truncatus* Montagu, in open waters," J. Acoust. Soc. Am. **56**, 1280–1290.
Beedholm, K., Miller, L. A., and Blanchet, M. A. (**2006**). "Auditory brainstem response in a harbor porpoise show lack of automatic gain control for simulated echoes," J. Acoust. Soc. Am. **119**, EL41–EL46.
Dear, S. P., and Suga, N. (**1995**). "Delay-tuned neurons in the midbrain of the big brown bat," J. Neurophysiol. **73**(3), 1084–1100.
Dolphin, W. F., Au, W. W. L., Nachtigall, P. E., and Pawloski, J. (**1995**). "Modulation rate transfer functions to low-frequency carriers in three species of cetaceans," J. Comp. Physiol. A **177**, 235–245.
Dolphin, W. F., and Mountain, D. C. (**1992**). "The envelope following response: Scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic signals," Hear. Res. **58**, 70–78.
Elmegaard, S. L., McDonald, B. I., Teilmann, J., and Madsen, P. T. (**2021**). "Heart rate and startle responses in diving, captive harbour porpoises (*Phocoena phocoena*) exposed to transient noise and sonar," Biol. Open **10**, 058679.
Finneran, J. J., Carder, D. A., Schlundt, C. E., and Ridgway, S. H. (**2005**). "Temporary threshold shift (TTS) in bottlenose dolphins (*Tursiops truncatus*) exposed to mid-frequency tones," J. Acoust. Soc. Am. **118**, 2696–2705.
Finneran, J. J., Mulsow, J., Houser, D. S., and Burkard, R. F. (**2016**). "Place specificity of the click-evoked auditory brainstem response in the bottlenose dolphin (*Tursiops truncatus*)," J. Acoust. Soc. Am. **140**, 2593–2602.
Galambos, R., Makeig, S., and Talmachoff, P. J. (**1981**). "A 40-Hz auditory potential recorded from the human scalp," Proc. Natl. Acad. Sci. U.S.A. **78**, 2643–2647.
Geisler, C., Frishkopf, L., and Rosenblith, W. (**1958**). "Extracranial responses to acoustic clicks in man," Science **128**, 1210–1211.
Hernandez, E. N., Kuczaj, S., Houser, D. S., and Finneran, J. J. (**2007**). "Middle- and long-latency auditory evoked potentials in the bottlenose dolphin (*Tursiops truncatus*) resulting from frequent and oddball stimuli," Aquat. Mamm. **33**, 34–42.
Houser, D. S., and Finneran, J. J. (**2006**). "A comparison of underwater hearing sensitivity in bottlenose dolphins (*Tursiops truncatus*) determined by electrophysiological and behavioral methods," J. Acoust. Soc. Am. **120**, 1713–1722.
Koch, M., and Schnitzler, H. U. (**1997**). "The acoustic startle response in rats—circuits mediating evocation, inhibition and potentiation," Behav. Brain. Res. **89**(1–2), 35–49.
Kuwada, S., Batra, R., and Maher, V. L. (**1986**). "Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones," Hear. Res. **21**(2), 179–192.
Ladegaard, M., and Madsen, P. T. (**2019**). "Context-dependent biosonar adjustments during active target approaches in echolocating harbour porpoises," J. Exp. Biol. **222**(16), jeb206169.
Legatt, A. D., Arezzo, J. C., and Vaughn, H. G., Jr. (**1986**). "Short-latency auditory evoked potentials in the monkey. II. Intracranial generators," Electroenceph. Clin. Neurophysiol. **64**, 53–73.
Linnenschmidt, M., Wahlberg, M., and Damsgaard Hansen, J. (**2013**). "The modulation rate transfer function of a harbour porpoise (*Phocoena phocoena*)," J. Comp. Physiol. A **199**(2), 115–126.
Luck, S. J. (**2014**). *An Introduction to the Event-related Potential Technique*, 2nd ed. (MIT, Cambridge, MA).
MacAulay, J. D. J., Malinka, C. E., Gillespie, D., and Madsen, P. T. (**2020**). "High resolution three-dimensional beam radiation pattern of harbour porpoise clicks with implications for passive acoustic monitoring," J. Acoust. Soc. Am. **147**(6), 4175–4188.

Madsen, P. T., and Surlykke, A. (**2013**). "Functional convergence in bat and toothed whale biosonars," Physiology **28**(5), 276–283.
Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., and Prado, L. (**2015**). "Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque," Eur. J. Neurosci. **41**(5), 586–602.
Møhl, B., and Andersen, S. (**1973**). "Echolocation: High-frequency component in the click of the Harbour Porpoise (*Phocoena ph. L.*)," J. Acoust. Soc. Am. **54**(5), 1368–1372.
Mooney, T. A., Yamato, M., and Branstetter, B. K. (**2012**). "Hearing in cetaceans: From natural history to experimental biology," Adv. Mar. Biol. **63**, 197–246.
Moore, B. C., Peters, R. W., and Glasberg, B. R. (**1990**). "Auditory filter shapes at low center frequencies," J. Acoust. Soc. Am. **88**(1), 132–140.
Morozov, V. P., Akopian, A. I., Burdin, V. I., Zaitseva, K. A., and Sokovykh, Y. A. (**1972**). "Tracking frequency of the location signals of dolphins as a function of distance to the target," Biofizika **17**, 139–145.
Nachtigall, P. E., Mooney, T. A., Taylor, K. A., and Yuen, M. M. (**2007**). "Hearing and auditory evoked potential methods applied to odontocete cetaceans," Aquat. Mamm. **33**(1), 6–13.
Nordmark, J. (**1960**). "Perception of distance in animal echo-location," Nature **188**, 1009–1010.
Özdamar, Ö., and Kalayci, T. (**1999**). "Median averaging of auditory brain stem responses," Ear Hear. **20**, 253–264.
Picton, T. W., Woods, D. L., Baribeau-Braun, J., and Healey, T. M. (**1977**). "Evoked potential audiometry," J. Otolaryngol. **6**, 90–119.
Popov, V. V., Ladygina, T. F., and Supin, A. Y. (**1986**). "Evoked potentials of the auditory cortex of the porpoise, *Phocoena phocoena*," J. Comp. Physiol. **158**, 705–711.
Popov, V. V., and Supin, A. Y. (**1990**). "Auditory brainstem responses in characterization of dolphin hearing," J. Comp. Physiol. A **166**, 385–393.
Popov, V. V., and Supin, A. Y. (**2007**). "Analysis of auditory information in the brains of cetaceans," Neurosci. Behav. Physiol. **37**, 285–291.
Ridgway, S. H. (**2011**). "Neural time and movement time in choice of whistle or pulse burst responses to different auditory stimuli by dolphins," J. Acoust. Soc. Am. **129**, 1073–1080.
Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., and Galambos, R. (**1981**). "Auditory brainstem response in dolphins," Proc. Natl. Acad. Sci. U.S.A. **78**, 1943–1947.
Salles, A., Diebold, C. A., and Moss, C. F. (**2020**). "Echolocating bats accumulate information from acoustic snapshots to predict auditory object motion," Proc. Natl. Acad. Sci. U.S.A. **117**(46), 29229–29238.
Schalles, M. D., Houser, D. S., Finneran, J. J., Tyack, P., Shinn-Cunningham, B., and Mulsow, J. (**2021**). "Measuring auditory cortical responses in *Tursiops truncatus*," J. Comp. Physiol. A **207**(5), 629–640.
Sinha, S. R., and Moss, C. F. (**2007**). "Vocal premotor activity in the superior colliculus," J. Neurosci. **27**(1), 98–110.
Smith, A. B., Madsen, P. T., Johnson, M., Tyack, P., and Wahlberg, M. (**2021**). "Toothed whale auditory brainstem responses measured with a non-invasive, on-animal tag," JASA Express Lett. **1**(9), 091201.
Supin, A. Y., Mukhametov, L. M., Ladygina, T. F., Popov, V. V., Mass, A. M., and Poliakova, E. C. (**1978**). *An Electrophysiological Investigation of the Dolphin Brain* (Nauka, Moscow).
Supin, A. Y., and Popov, V. V. (**1995**). "Envelope-following response and modulation transfer function in the dolphin's auditory system," Hear. Res. **92**(1–2), 38–46.
Supin, A. Y., and Popov, V. V. (**2007**). "Improved techniques of evoked-potential audiometry in odontocetes," Aquat. Mamm. **33**, 14–23.
Supin, A. Y., Popov, V. V., and Mass, A. M. (**2001**). *The Sensory Physiology of Aquatic Mammals* (Kluwer Academic, Boston).
Vance, H. M., Madsen, P. T., Aguilar de Soto, N., Wisniewska, D. M., Ladegaard, M., Hooker, S. K., and Johnson, M. (**2021**). "Echolocating toothed whales use ultra-fast echo-kinetic responses to track evasive prey," eLife **10**, e68825.
Verfuß, U. K., Miller, L. A., Pilz, P. K., and Schnitzler, H. U. (**2009**). "Echolocation by two foraging harbour porpoises (*Phocoena phocoena*)," J. Exp. Biol **212**(6), 823–834.
Villadsgaard, A., Wahlberg, M., and Tougaard, J. (**2007**). "Echolocation signals of wild harbour porpoises, *Phocoena phocoena*," J. Exp. Biol. **210**, 56–64.
Woods, D. L., Ridgway, S. H., Carder, D. A., and Bullock, T. H. (**1986**). "Middle and long-latency auditory event-related potentials in dolphins," in *Dolphin Cognition and Behavior: A Comparative Approach*, edited by R. J. Schusterman, J. A. Thomas, and F. G. Wood (Erlbaum, Hillsdale NJ), Vol. 3, pp. 61–77.
Yin, T. C., Smith, P. H., and Joris, P. X. (**2019**). "Neural mechanisms of binaural processing in the auditory brainstem," Comp. Physiol. **9**(4), 1503–1575.
Yuen, M. M., Nachtigall, P. E., Breese, M., and Supin, A. Y. (**2005**). "Behavioral and auditory evoked potential audiograms of a false killer whale (*Pseudorca crassidens*)," J. Acoust. Soc. Am. **118**(4), 2688–2695.