Hearing Research 331 (2016) 83–91

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

High-frequency hearing in seals and sea lions

Kane A. Cunningham [a, *], Colleen Reichmuth [b]

[a] Department of Ocean Sciences, Long Marine Laboratory, University of California, Santa Cruz, 100 Shaffer Road, Santa Cruz, CA 95060, USA
[b] Institute of Marine Sciences, Long Marine Laboratory, University of California, Santa Cruz, Santa Cruz, CA 95060, USA

ARTICLE INFO

Article history:
Received 30 May 2015
Received in revised form
24 September 2015
Accepted 5 October 2015
Available online 28 October 2015

Keywords:
Ultrasound
Bone conduction
High-frequency hearing
Underwater hearing

ABSTRACT

Existing evidence suggests that some pinnipeds (seals, sea lions, and walruses) can detect underwater sound at frequencies well above the traditional high-frequency hearing limits for their species. This phenomenon, however, is not well studied: Sensitivity patterns at frequencies beyond traditional high-frequency limits are poorly resolved, and the nature of the auditory mechanism mediating hearing at these frequencies is unknown. In the first portion of this study, auditory sensitivity patterns in the 50–180 kHz range were measured for one California sea lion (*Zalophus californianus*), one harbor seal (*Phoca vitulina*), and one spotted seal (*Phoca largha*). Results show the presence of two distinct slope-regions at the high-frequency ends of the audiograms of all three subjects. The first region is characterized by a rapid decrease in sensitivity with increasing frequency—i.e. a steep slope—followed by a region of much less rapid sensitivity decrease—i.e. a shallower slope. In the second portion of this study, a masking experiment was conducted to investigate how the basilar membrane of a harbor seal subject responded to acoustic energy from a narrowband masking noise centered at 140 kHz. The measured masking pattern suggests that the initial, rapid decrease in sensitivity on the high-frequency end of the subject's audiogram is not due to cochlear constraints, as has been previously hypothesized, but rather to constraints on the conductive mechanism.

1. Introduction

The ears of pinnipeds are amphibious, meaning that they operate effectively both in air and under water (Reichmuth et al. 2013). This is remarkable given the different obstacles presented by the two media. In air, the major difficulty is the large impedance mismatch between the gaseous media surrounding the head and the cochlear fluids, which makes the efficient transfer of acoustic energy from one to the other challenging. In mammals, this problem is solved by the middle-ear ossicles, which act as an amplifier and impedance matcher. When an animal is submerged in water, however, the air-filled middle-ear cavity that contains the ossicles creates the inverse impedance-matching problem, effectively blocking the traditional air-conduction energy-transmission pathway to the cochlea. Because of this, it has long been assumed that sound must travel along an alternate pathway (or pathways) to the inner ear in order for pinnipeds to hear effectively under water, and that different auditory mechanisms must therefore be responsible for energy transmission in the two different media (Møhl, 1968a). In other words, pinniped underwater hearing is thought to be bone conducted, meaning that sound energy travels to the cochlea via direct coupling of the surrounding fluid to the animal's flesh and bone, allowing the air-filled middle-ear cavity to be bypassed (Møhl, 1968a; Repenning, 1972).

Psychophysical studies show that pinniped hearing capabilities are markedly different in air versus under water, supporting the hypothesis that different auditory mechanisms operate in the two media. Phocid seals,[1] in particular, have an expanded frequency range of hearing in water relative to their frequency range of hearing in air—a difference of over an octave (Reichmuth et al., 2013; Sills et al., 2014). Hemilä et al. (2006) hypothesized that this difference is primarily due to the massive middle-ear ossicles that are characteristic of phocids. These authors argued that ossicular inertia would impede the transmission of high-frequency energy in air, but could aid the transmission of bone-conducted underwater sound via amplification of the relative motion that

* Corresponding author.
E-mail address: kaacunni@ucsc.edu (K.A. Cunningham).

[1] The three extant families of pinnipeds are Phocidae (true seals), Otariidae (sea lions and fur seals), and Odobenidae (walruses).

http://dx.doi.org/10.1016/j.heares.2015.10.002
0378-5955/

occurs between the ossicles and the cochlea when the head vibrates. They further hypothesized that, while middle-ear inertia acts to constrain aerial high-frequency hearing in phocids, cochlear constraints—such as the termination of the tonotopic map on the basal end of the basilar membrane (BM), or the upper frequency limit of the cochlear amplifier—are the primary limits on underwater high-frequency hearing in both phocid and otariid pinnipeds.

The hypothesis that high-frequency underwater hearing is limited by cochlear constraints has recently been used to interpret auditory data for several pinniped species (e.g., Reichmuth et al., 2013; Mulsow et al., 2012). The apparent shape of pinniped audiograms at ultrahigh frequencies, however, challenges this influential idea. Studies of both California sea lions (*Zalophus californianus*) and harbor seals (*Phoca vitulina*) suggest that the portion of the audiogram near the high frequency hearing limit—around 40 kHz for California sea lions and 80 kHz for harbor seals—may be characterized by two distinct regions with different slopes (Møhl, 1968b; Schusterman et al., 1972; Kastelein et al., 2009; Cunningham et al., 2014). In both species, the rate at which sensitivity decreases with increasing frequency is initially very rapid, but appears to decrease once a certain frequency is reached. In other words, the slope of the audiogram near the high-frequency hearing limit starts off very steep, but then becomes shallower at even higher frequencies. However, audiometric data at these ultrahigh frequencies (frequencies in the second, shallow slope region) are relatively scarce and the two-slope phenomenon has yet to be shown conclusively. The confirmed existence of a second, shallower slope-region in the high-frequency portion of the audiogram would indicate that the cochlea is capable of processing sounds at ultrahigh frequencies, suggesting that the steep-slope portion of the audiogram is the result of constraints on the bone-conducted energy transmission mechanism, rather than constraints on the inner ear.

Overall, underwater high-frequency hearing in pinnipeds is poorly understood—both in terms of the underlying mechanism, and in terms of the shape of the audiogram. This is partially due to the fact that auditory sensitivity at high and ultrahigh frequencies is relatively low. Low sensitivity means that few naturally occurring ultrahigh-frequency sounds will be audible to pinnipeds, which can lead to the assumption that hearing at these frequencies is not biologically relevant. However, there are an increasing number of human-made marine technologies that generate sounds in the 80–180 kHz range at source levels as high as 230 dB re 1 μPa @ 1 m,[2] including commercial sonar and recreational fish finders (Hildebrand, 2004, 2009). Such noises may indeed be audible to certain pinnipeds. The relative lack of understanding as to how pinnipeds perceive sounds at ultrahigh frequencies has therefore become problematic, and has led to confusion as to which underwater sounds pinnipeds may be able to detect (Cunningham et al., 2014).

The primary goal of this study was to extend the existing underwater audiograms of one California sea lion, one harbor seal, and one spotted seal (*Phoca largha*) to 180 kHz by measuring behavioral auditory detection thresholds in quiet conditions. The secondary goal was to conduct an auditory masking experiment with a single pinniped subject (the harbor seal) to determine how ultrahigh frequency sounds are encoded on an individual's BM, thereby providing insight into the physiological factors contributing to the shape of the audiogram at ultrahigh frequencies.

[2] Note that other factors in addition to source level, such as beam width and transmission loss, must be taken into account when predicting the effects of noise on wild animals.

2. Materials and methods

Two experiments were conducted as part of this study. Both involved the measurement of underwater auditory detection thresholds for experienced animal subjects using cooperative behavioral methods. To ensure that threshold measurements were comparable to earlier measurements made for the same individuals, the methods used in this study and described below are the same as in Reichmuth et al. (2013) and Sills et al. (2014).

All experimental testing was conducted in an outdoor, semi-inground, circular, concrete test pool (7.6 m diameter, 1.8 m depth) filled with natural seawater, with the subject at a depth of 1 m and the acoustic projector approximately 60 cm in front of the subject. During testing, the presentation of acoustic stimuli was controlled by an experimenter in a remote location with visual access to the subject via an underwater camera. Fish rewards were delivered to the animal by a trainer positioned on the pool deck who was blind to the experimental condition. A go/no-go paradigm was used for the detection task. The animal subject would begin each trial at the surface and dive to a polyvinyl chloride (PVC) chin cup when cued by the trainer. Once the subject was calm and relatively motionless at the station, the experimenter would turn on an underwater light positioned in front of the subject to indicate the start of a 4 s listening interval. The subject had been previously trained to leave the station and touch its nose to a PVC target positioned 20 cm to its left when it detected a signal (go), and remain on the station for the full 4 s interval if it did not (no-go). Correct responses (detections and rejections) were reinforced with a single piece of fish. False alarms and missed detections were never reinforced and the subject was allowed to continue the task following a 2–3 s period at the surface. A test sequence of signal-absent and signal present trials was generated pseudo-randomly prior to the start of each experimental session, with the constraint that the same trial type could not occur more than four times consecutively. All experimental sessions began and ended with a series of up to ten test trials that were clearly audible to the subject to ensure adequate stimulus control.

Two psychophysical methods were used in succession for all hearing threshold measurements. A staircase (up/down) procedure was first used to obtain a preliminary threshold estimate for a subject at a given test frequency. Individual staircase sessions used a 2 dB step size and consisted of 40–60 trials depending on how long it took for the subject's response to stabilize, as indicated by five consecutive descending misses within ±3 dB of one another. Staircase sessions were run until the subject's performance stabilized to the point that the within-session thresholds across three separate sessions fell within a 3 dB range. For this study, this occurred within five staircase sessions for all subjects at all frequencies.

Staircase testing was followed by additional testing using a method of constant stimuli (MCS). Each MCS session comprised five test blocks, and each block included signal trials at +4 dB, +2 dB, +0 dB, -2 dB, and −4 dB relative to the threshold estimate obtained from staircase testing. MCS sessions were run until at least two sessions were obtained that met the following criteria: 1) session thresholds within 3 dB of one another across sessions, 2) 95% confidence intervals no wider than 4 dB, and 3) the probability of detection as a function of stimulus level was determined to follow a normal distribution as measured by a Chi-squared goodness of fit test (see below). In the rare event that MCS thresholds were not consistent with the staircase estimate, the staircase testing process was resumed until within-session thresholds again stabilized, at which point a new MCS testing block would begin.

Final 50% detection thresholds, standard error, and 95%

confidence intervals were calculated using probit analysis on the MCS data only (Finney, 1947). Probit analysis is based on maximum likelihood estimation under the assumptions that the response probability as a function of stimulus level is normally distributed, and that the number of correct detections for a given number of signal presentations at a given stimulus level is binomially distributed (i.e., a series of Bernoulli trials with a constant probability of detection). These assumptions are appropriate when care is taken to ensure that the subject's response is stable (i.e. systematic temporal effects are minimized), and the quality of fit of the estimated normal distribution is quantified. A stable subject response was obtained by extending testing until the previously mentioned stability criteria were met, and by utilizing the across-session threshold criteria outlined above. The appropriateness of the normal distribution model was quantified by requiring a *p*-value greater than 0.1 for a Chi-squared goodness-of-fit test applied to the probit model.

In this study, false alarm rates were controlled throughout testing by adjusting the ratio of signal trials to catch trials in between experimental sessions as needed (minimum 50% signal trials, maximum 70%). False alarm rates were calculated using only the trials in the portion of the session where the stimulus levels were near threshold, excluding supra-threshold trials in the warm-up and cool-down portions of the session. In order for a set of sessions to be used for preliminary threshold estimation (staircase) or final threshold calculations (MCS), the pooled false alarm rate was required to be above 0% and less than 30%. By restricting the false alarm rate in such a manner, differences in subject response bias across treatments were minimized.

The equipment chain used to project, receive, and calibrate signals was the same in both experiments. Outgoing signals were generated using custom software on a PC (HTP, Finneran, 2003). From the PC, signals were sent to a NI USB-6259 data acquisition system for digital-to-analog conversion (500 kHz sampling rate, National Instruments Corp., Austin, TX), attenuated using a TDT PA5 programmable attenuator (Tucker–Davis Technologies, Alachua, FL), and amplified using a Hafler P1000 amplifier (Hafler Professional, Tempe, AZ). Depending on frequency, signals were projected either from an ITC 1042 (International Transducer Corp., Santa Barbara, CA), or a Reson 4013 projector (Teledyne-RESON A/S, Slangerup, Denmark) and received by either a Reson 4032 or a Reson 4014 receiver.

Signals were calibrated immediately prior to each session. The calibration routine fit a linear model to received SPL as a function of outgoing voltage (with the receiving hydrophone positioned in the same location that was to be occupied by the center of the subject's head during testing). The model was then tested across the range of SPLs to be employed during testing. The RMS error for this test was required to be less than 1 dB for testing to commence. In addition, the spatial variability for each test stimulus was quantified prior to its use in any experimental sessions. This was done by measuring received SPL at 24 locations on a 7 cm × 7 cm x 7 cm cubic grid centered on the calibration position (as in Sills et al., 2014). SPLs at all locations were required to be within ±3 dB of each other in order for testing to commence.

Because this study deals with sounds at relatively high SPLs projected in an enclosed pool, extra care was taken to ensure signal fidelity in both time and frequency domains. For each test condition, the received signal (i.e. the signal as measured by a low-noise hydrophone positioned in the eventual location of the test subject's head) was carefully inspected for transients, artifacts, and distortion. This included visual assessment of digitized signals on a PC prior to each experimental session, as well as post-hoc analyses of temporal variability and additional inspection of recorded signals for distortion. The details of the post-hoc analyses are presented in Appendix B.

2.1. Sensitivity

The goal of the first experiment was to measure pinniped underwater auditory sensitivity in the 50–180 kHz range. To this end, 50% correct-detection thresholds were measured in quiet conditions for three subjects: One 6-year-old female California sea lion (*Ronan*, NOA0006602), one 26-year-old male harbor seal (*Sprouts*, NOA0001707), and one 4-year-old male spotted seal (*Tunu*, NOA0006674). The subjects were experienced in behavioral auditory testing and their hearing sensitivity had been previously measured under water at frequencies extending to 43 kHz for the sea lion (Reichmuth et al., 2013), and to 72.4 kHz for the harbor seal and spotted seal (Reichmuth et al., 2013; Sills et al., 2014).

The test signals were single narrowband FM sweeps (10% bandwidth) of 500 ms duration with 25 ms linear on/off ramps and were presented within a 4 s trial window. Narrowband FM sweeps were used in favor of pure tones to reduce the spatial variability of signal SPL in the reverberant test environment (Kastelein et al., 2002; Finneran and Schlundt, 2007). The hearing sensitivity of each subject was initially tested at signal frequencies 50, 80, 100, 140, and 180 kHz in a mixed order. The sea lion subject was tested at two additional frequencies—36 and 40 kHz—in order to fully capture the steep-slope portion of the audiogram. Additional hearing thresholds were obtained using the same methods in quiet conditions for the harbor seal subject at 40, 50, 60, 70, 80, 90, 100, 120, and 140 kHz within 4 months of initial testing as part of the masking portion of this study (see section 2.2). Therefore, at frequencies that were tested twice for the harbor seal subject, final thresholds were determined by taking the average of the two measurements. Testing for the sensitivity experiment took place between April 2014 and January 2015.

2.2. Masking

An additional masking experiment was conducted with a single subject—the harbor seal—in order to determine BM excitation patterns for a pinniped exposed to an ultrahigh frequency noise. For this experiment, 50% detection thresholds were measured at a total of nine signal frequencies in a mixed order (40, 50, 60, 70, 80, 90, 100, 120, 140 kHz), first in quiet conditions (see Section 2.1), and then in the presence of narrowband, 140 kHz masking noise. At each frequency, the subject fully completed testing in quiet conditions first, and then was tested at the same frequency in masked conditions.

The masking noise used for this experiment was narrowband, flat-spectrum noise centered at 140 kHz and extending to ±5% of the center frequency (133–147 kHz). This noise was generated in MATLAB (MathWorks, Natick, MA), sent to a NI PCI-6251 DAQ data acquisition system (500 kHz sampling rate, National Instruments Corp., Austin, TX), and then to a Behringer XENYX 502 mixer (MUSIC Group Co., Las Vegas, NV), where the signal and noise were mixed prior to amplification and projection from a single Reson 4013 projector. The SPL of the masking noise was kept constant across all treatments at 143 dB re 1 μPa, equivalent to a sensation level (SL) of 13 dB[3] for this subject. A narrowband masking noise was chosen over a broadband or tonal noise in order to stimulate a narrow region of the BM while minimizing the presence of acoustic beats (Egan and Hake, 1950), thereby producing a masking pattern representative of BM excitation in response to 140 kHz sound. Similar methods have been used to study BM excitation patterns for human subjects at

[3] The original intent was to have the masker at 15 dB SL, but subsequent threshold measurements at 140 kHz indicated that the actual sensation level was slightly less. Note that the masker level of 143 dB re 1 μPa is well below the estimate for the onset of temporary threshold shift in phocids (Southall et al., 2007; NOAA, 2013).

frequencies within the normal human hearing range (Moore, 1993), and to study the excitation patterns generated by an ultrasonic masker delivered via a bone vibrator (Nishimura et al., 2003).

To ensure that SPL measurements taken at the calibration position were representative of the sound field encountered by the subject, the spatial variability of the masker was determined by measuring the SPL at 24 locations on a 7 cm × 7 cm x 7 cm cubic grid centered around the location that was to be occupied by the subject's head during testing (as in Sills et al., 2014). The SPL of the masking noise was within ± 3 dB of the calibration position at all 24 locations. The gradual onset of the masking noise (0.5 s ramp) began approximately 2 s before the start of the 4 s listening interval of the trial, and the offset occurred after the end of the listening interval. Testing for the masking experiment took place between September 2014 and December 2014.

Following auditory testing, a masking model was built to test the hypothesis that the observed peak masking frequency (i.e., the frequency at which the threshold was most elevated by the presence of the masking noise) was determined by the end of the tonotopic map of the subject's BM (i.e., the highest frequency encoded by the BM). This model is described in full detail in Appendix A.

3. Results

3.1. Sensitivity

Absolute detection thresholds for signals from 36 to 180 kHz are given for each subject in Table 1. Thresholds ranged from a minimum of 64 dB re 1 μPa (spotted seal, 50 kHz) to a maximum of 148 dB re 1 μPa (sea lion, 180 kHz). The 95% confidence intervals for all thresholds were less than 3 dB at each frequency, and all standard errors were less than 0.6 dB. The false alarm rates (rate of response on signal-absent trials) ranged from 10 to 26% across all frequencies and subjects. For each subject, the high-frequency sensitivity curve exhibited two distinct slope-regions: a steep slope-region followed by a shallow slope-region. The average rate of sensitivity decrease in the steep slope-region across all three subjects was 320 dB/decade, compared to 70 dB/decade in the shallow slope-region. Fig. 1 depicts the high-frequency thresholds obtained for each subject, along with the most recent audiogram data available for the same individual, and the published psychophysical audiograms for other individuals of the same species. For comparison purposes, Fig. 2 depicts the extended audiograms (data from this study merged with previous data for the same individuals) of the three pinniped subjects in a single plot.

Table 1
Fifty-percent auditory detection thresholds, measured under water in quiet conditions, for one California sea lion, one harbor seal, and one spotted seal. Corresponding false alarm rates, expressed as the percentage of signal absent trials for which the subject responded, are provided in parentheses. Standard errors for all measurements were less than 0.6 dB.

Frequency kHz	Threshold, dB re 1 μPa (FA rate, %)		
	California sea lion	Harbor seal	Spotted seal
36	85 (10)		
40	110 (12)	75 (13)	
50	116 (10)	97 (20)	64 (13)
60		115 (12)	
70		117 (24)	
80	128 (13)	116 (13)	110 (26)
90		119 (24)	
100	132 (17)	120 (13)	113 (10)
120		125 (20)	
140	136 (10)	130 (20)	122 (10)
180	148 (16)	140 (10)	136 (17)

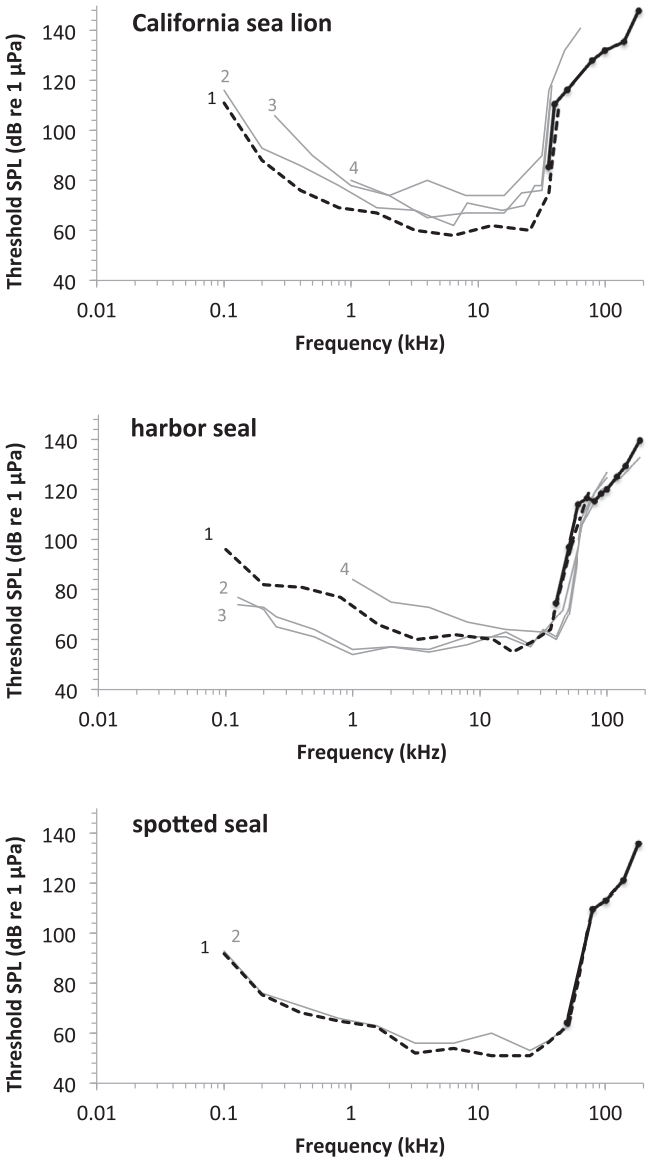

Fig. 1. High-frequency auditory detection thresholds measured in this study for three pinniped subjects (solid lines), along with previously published audiogram data for the same individuals (dashed lines), and data from other individuals of the same species (light gray lines). Note the presence of two distinct slope-regions on the high-frequency end of the audiogram for all three subjects. California sea lion: 1 (Reichmuth et al., 2013), 2 (Schusterman et al., 1972), with 50% thresholds calculated in Reichmuth and Southall (2012), 3 (Reichmuth and Southall, 2012), 4 (Mulsow et al., 2012). Harbor seal: 1 (Reichmuth et al., 2013), 2 & 3 (Kastelein et al., 2009), 4 (Møhl, B., 1968a). Spotted seal: 1 & 2 (Sills et al., 2014).

3.2. Masking

For the harbor seal subject, the amount of masking induced by a 140 kHz, 13 dB SL masking noise at signal frequencies from 40 to 140 kHz is depicted in Fig. 3 (upper panel). Masking amounts ranged from 0.6 to 4.4 dB, with peak masking occurring at 90 kHz. Masking increased monotonically with frequency from 40 to 90 kHz before decreasing sharply at 100 kHz and remaining low up to 140 kHz. Fig. 3 (lower panel) also shows the masking pattern predicted by the masking model developed for this study—the details of which are presented in Appendix A—alongside the empirical measurements from the harbor seal. This model was built on the assumptions that (1) the tonotopic map of the harbor seal

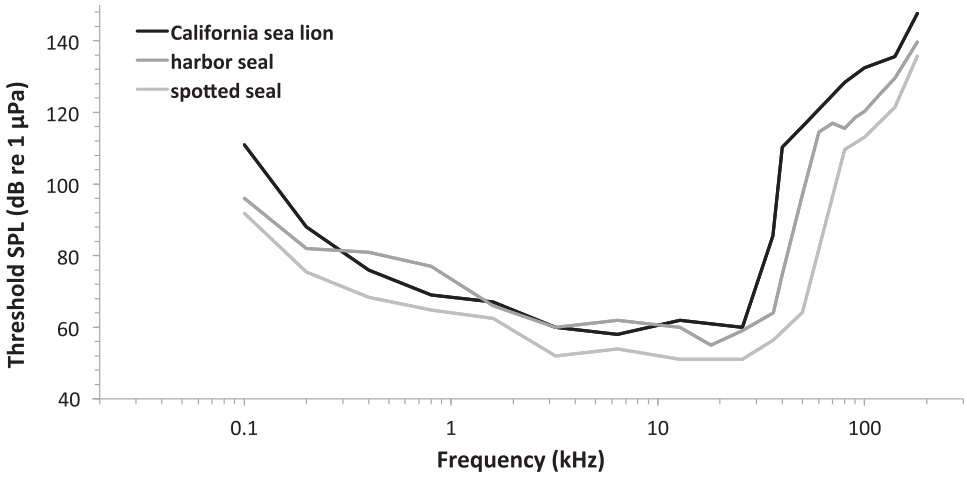

Fig. 2. Extended audiograms for three pinniped subjects. Audiograms are composites of previously published data for the same animals (California sea lion and harbor seal: Reichmuth et al., 2013; spotted seal: Sills et al., 2014) and high-frequency data from this study. Note that, while the different animals have different hearing ranges, all three exhibit two distinct slope-regions on the high-frequency end.

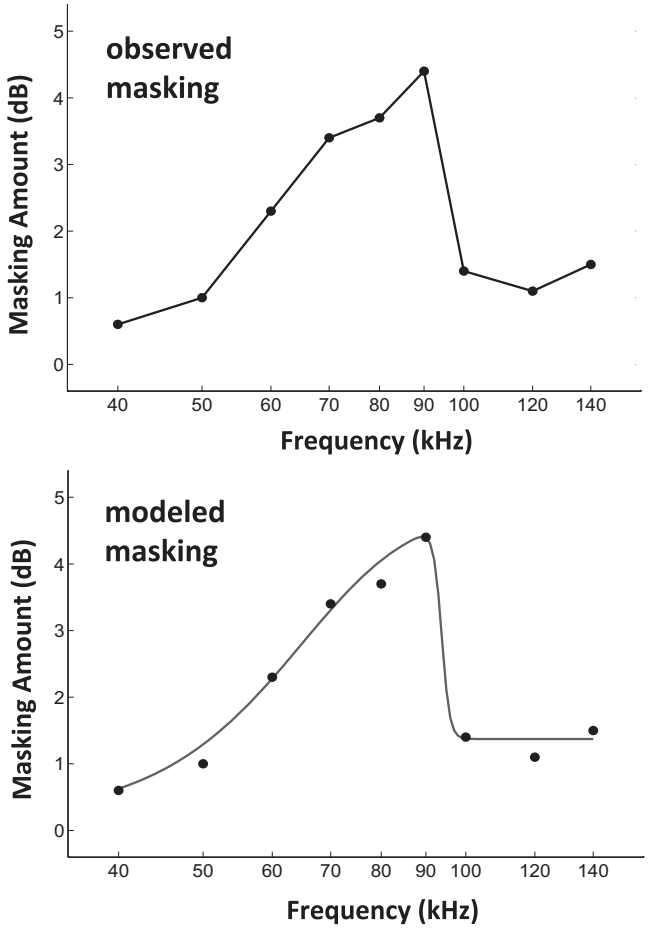

Fig. 3. Masking amounts for a harbor seal subject exposed to a narrowband masking noise centered at 140 kHz at a sensation level of 13 dB. The upper panel depicts the amount of masking for tonal signals in the 40–140 kHz range. Amount of masking is calculated as the difference between the detection threshold with and without the masker present. The lower panel depicts the masking amount predicted by a model based on BM displacement (solid line) plotted over the experimental data (points). The model assumed that the tonotopic map of the harbor seal's BM ended at 95 kHz and that signals at higher frequencies stimulated the very basal tip of the BM (see Appendix A).

subject's BM ends at 95 kHz, and (2) sound energy at frequencies beyond 95 kHz stimulates the very basal tip of the BM. The root-mean-square (RMS) error between the model and the experimental data was 0.4 dB; the coefficient of determination (R^2) was 0.98.

4. Discussion

This study confirms that pinniped underwater audiograms exhibit two distinct slope-regions on the high-frequency end: An initial region of steep slope, followed by a region of much shallower slope. This was true for all subjects tested, representing three species from two pinniped families (Otariidae and Phocidae). This two-slope phenomenon explains the ability of all subjects to detect signals at 180 kHz, well above their presumed individual high-frequency hearing limits (Kastelein et al., 2009; Møhl, 1968b; Mulsow et al., 2012; Reichmuth et al., 2013; Sills et al., 2014). When considered together, the three extended audiograms show that the transition point between the steeper and shallower slope-regions occurs at higher frequencies for subjects with better high-frequency hearing. The general shapes of the curves, however, are strikingly similar, as illustrated in Fig. 2. This suggests that similar physical mechanisms are determining sensitivity at high and ultrahigh frequencies across all three species, though different mechanisms may be active in the two different slope-regions.

The maximum amount of masking measured for the harbor seal subject exposed to 140 kHz noise occurred at 90 kHz. In typical masking scenarios—i.e. those not using an ultrahigh-frequency masker—maximum masking occurs at the frequency where signal and noise energy overlap the most (Fletcher, 1940). This would have resulted in peak masking at 140 kHz, more than ½-octave higher than the observed peak at 90 kHz. A possible explanation for this discrepancy is that the tonotopic map of the subject's BM ends somewhere between 90 and 100 kHz. If this were the case, the sudden decrease in masking above 90 kHz could be explained by the fact that no region of the BM is tuned to these ultrahigh frequencies, and hence the cochlear amplifier is inactive (Robles and Ruggero, 2001). The model used in this study was built specifically to test this hypothesis. The good agreement between the modeled masking patterns and the experimental data suggests that the active region of the cochlear amplifier—and presumably the tonotopic map of the BM—ends somewhere between 90 and 100 kHz for the harbor seal subject.

It would be improper to draw broad conclusions regarding the mechanisms mediating pinniped ultrahigh-frequency hearing from this single-subject masking study. However, it is important to note that the hypothesis that the cochlear amplifier is active up to 90 kHz in the harbor seal subject runs contrary to the hypothesis that cochlear constraints limit underwater high-frequency hearing in pinnipeds (Hemilä et al., 2006). If cochlear constraints were limiting high-frequency hearing, and if the cochlear amplifier were active up to 90 kHz, one would expect the rapid decrease in hearing sensitivity to begin around this frequency. Instead, the harbor seal audiogram shows that sensitivity begins to decrease rapidly approximately one octave below 90 kHz. Further, if cochlear constraints were the cause of the steep-slope region of the audiogram, it is unclear why there would be a second, shallower slope region at even higher frequencies. A better explanation of these results is that the efficiency of the primary bone conduction mechanism mediating the transmission of acoustic energy to the cochlea rapidly decreases beyond a certain frequency limit, resulting in a sharp decrease in sensitivity and therefore a steep slope. The two-slope phenomenon can then be explained by the presence of a second bone conduction mechanism that is able to operate more effectively at higher frequencies. This hypothesis is supported by studies of bone conduction in humans and other terrestrial mammals showing the existence of different bone conduction mechanisms (or modes) that operate more or less effectively in different frequency ranges (Stenfelt, 2011). While some hypotheses exist as to how bone conduction may be enhanced by certain derived characteristics in the auditory peripheries of pinnipeds (Repenning, 1972), the precise bone conduction mechanisms involved in pinniped underwater hearing are presently unknown.

The idea that pinniped underwater high-frequency hearing is constrained by the limits of a conductive mechanism is analogous to the hypothesis that the middle ear constrains aerial high-frequency hearing in humans. Human aerial audiograms exhibit a characteristic steep slope around 14–20 kHz. The presence of a single steep-slope region means that sound quickly becomes inaudible beyond a certain frequency, and that the high-frequency hearing limits of individuals tend to be well defined. This has led to the designation of sounds above 20 kHz as ultrasonic, i.e. inaudible to humans. The rapid loss of sensitivity at high frequencies is often attributed to the frequency response characteristics of the middle ear (Pumphrey, 1950; Hemilä et al., 1995; but see Ruggero and Temchin, 2002). This hypothesis is supported by the fact that humans *can* detect ultrasound in conditions where the stimulus is directly coupled to the head, such as via a bone vibrator (Corso, 1963), or via immersion in liquid media—as when listening underwater (Deatherage et al., 1954). Presumably, ultrasound is detectable in these situations because a bone-conducted pathway that bypasses the middle ear is available.

Not only can humans detect ultrasound when it is bone-conducted, human bone-conducted sensitivity patterns exhibit two slope-regions near the traditional high-frequency hearing limit (Corso, 1963), much like the pinniped underwater sensitivity patterns measured in this study. This leads to the hypothesis that similar mechanisms may be operating in both listening scenarios. This view is further supported by recent evidence that human underwater audiograms also exhibit two-slope regions on the high-frequency end (Qin et al., 2011; M. Qin 2015, pers. comm., 25 March). Similar patterns of high-frequency, bone-conducted hearing in humans and pinnipeds suggest that the ability to detect ultrahigh-frequency underwater sound is not specific to pinnipeds and did not evolve in response to the selective pressures associated with a semi-aquatic lifestyle. It also opens the possibility that such an epiphenomenon—if present across many mammalian species—may have been a precursor to the evolution of high-frequency biosonar in odontocetes.

Some researchers have suggested that the ability of human subjects to hear bone-conducted ultrasound is not due to sound energy bypassing the middle ear, but rather due to ultrasound being demodulated to lower frequencies by a non-linear process inherent to the bone-conducted energy pathway along which it travels (Lenhardt, 2003). Such a phenomenon could account for both the ability to detect ultrasound, and the fact that ultrasound is generally perceived as having a pitch in the high-audio range (Dieroff and Ertel, 1975; Kono et al., 1985). Hypotheses as to the precise bone-conducted pathway that could lead to demodulation are scarce. Lenhardt (2003) suggested that ultrasonic energy is demodulated as it travels from the brain tissue, to the cerebrospinal fluid, to the cochlear aqueduct, to the scala tympani. According to this idea, the frequency to which ultrasound is demodulated is determined by the size of the brain, and hence should vary inversely with head size. Therefore, the peak masking frequency resulting from an ultrasonic masker should also vary inversely with head size. By Lenhardt's (2003) calculations, the resonant frequency for a cat brain is around 40 kHz. Given that the brain of a harbor seal is larger than that of a cat, it would be reasonable to predict that, if a brain-demodulation mechanism were active for the harbor seal subject, the peak masking frequency would be less than 40 kHz. The observed peak masking frequency of 90 kHz, therefore, contradicts this demodulation hypothesis for the harbor seal.

On the whole, more research is necessary to resolve the physiological mechanisms governing the shape of the pinniped audiogram at high frequencies. The measurement of psychophysical tuning curves near the high-frequency hearing limit could be particularly useful in determining whether the sharp tuning thought to be a product of the cochlear amplifier is present in these frequency regions. Until further studies are conducted, however, it can not be assumed that pinniped underwater high-frequency hearing is limited by cochlear constraints.

In summary, while uncertainty remains as to the precise mechanisms supporting pinniped underwater hearing, the sensitivity curves from this study, combined with measurements from earlier studies (Møhl, 1968b; Schusterman et al., 1972; Kastelein et al., 2009), demonstrate that pinniped audiograms are characterized by two distinct slope-regions on the high-frequency end. This explains the surprising ability of seals and sea lions to detect sounds that are well above their nominal underwater high-frequency hearing limits. For example, the sea lion subject in this study could detect sounds at 180 kHz, two full octaves above the traditional California sea lion high-frequency hearing limit of 36–40 kHz (Mulsow et al., 2012; Reichmuth et al., 2013). Because of this, the two-slope phenomenon should be taken into account when predicting whether specific high-frequency, high-energy marine technologies will be audible to free-ranging pinnipeds.

Acknowledgments

Animal research was conducted with the approval of the Institutional Animal Care and Use Committee at UC Santa Cruz, and was authorized by the National Marine Fisheries Service under Marine Mammal Permit 14535 to C. R. No animals were harmed in this study. The project was supported by the Packard Endowment for Ocean Science and Technology at UC Santa Cruz and by NOAA/NMFS. We thank Sean Hayes at NOAA for encouraging this work given its relevance to fisheries acoustics, and we acknowledge Viginia Richards, Whitlow Au, Chris Edwards, and two anonymous reviewers for helpful comments on an earlier version of this manuscript. We thank the entire team at the Pinniped Cognition and Sensory Systems Laboratory at UC Santa Cruz for their participation in this research.

Appendix A

A masking model based on BM displacement was created for this study to test the hypotheses that (1) the tonotopic map of the harbor seal subject's BM ends between 90 and 100 kHz, and (2) sound energy beyond the range of the tonotopic map stimulates the basal tip of the BM without activating the cochlear amplifier. The model compares estimated BM displacement for tonal signals and a 140 kHz masking noise in order to predict masking amounts. For signals, BM displacement was modeled as a function of signal SPL and frequency:

$$D_s(p,f) = p \cdot \frac{m(f)}{2} + D_0, \tag{A.1}$$

where D_s is BM displacement in dB re d_Δ, and $m(f)$ is the rate of growth of displacement per unit pressure (in dB/dB) as a function of signal frequency. The factor of ½ associated with the growth function is included so that linear growth occurs at $m(f)$=1 dB/dB, given that pressure is modeled in terms of SPL (dB re 1 μPa), which is determined by the square of the RMS pressure. The reference displacement, d_Δ, is the smallest detectable displacement such that when the signal SPL equals the pressure corresponding to the detection threshold ($p = p_T$), the BM displacement is 0 dB. Using this fact, Eq. A.1 can be solved for D_0 at $p = p_T$:

$$D_0 = -p_T \cdot \frac{m(f)}{2}. \tag{A.2}$$

The rate of growth of BM displacement—($m(f)$)—depends on whether the center frequency of the BM location being considered is the same as the frequency of the signal. If the signal is stimulating a region of the BM tuned to the signal frequency, BM displacement will grow at a slower-than-linear rate with increasing pressure due to a well-known compressive nonlinearity within the BM response thought to be the result of an active mechanism known as the cochlear amplifier. When the center frequency of the BM location being considered is distant from the frequency of the stimulating signal, the cochlear amplifier is considered to be inactive and displacement grows linearly with pressure (Robles and Ruggero, 2001).

For each signal considered in the masking model, the BM region of interest is where the maximum displacement occurs. For signals at frequencies within the range represented on the tonotopic map, maximum displacement occurs at the BM location tuned to the signal center frequency. Because the center frequency of the BM region of interest (i.e., the region where displacement is being modeled) equals the signal frequency, displacement will grow at a less-than-linear rate due to the cochlear amplifier. However, signals at frequencies higher than the highest frequency encoded by the subject's BM are assumed to cause maximum displacement at the very basal tip of the BM (where the highest frequencies are encoded). As a result, the BM region of interest for such signals is tuned to a lower center frequency than that of the stimulating signal, and displacement is assumed to grow at a linear rate. Hence, the growth-rate function—($m(f)$)—should be less than linear for signals at frequencies within the range of the tonotopic map, and should transition to a linear rate once the end of the tonotopic map has been reached. This was achieved by modeling the growth function as a function of signal frequency using Eq. A.3:

$$m(f) = m_{NL} + \frac{m_L - m_{NL}}{1 + e^{-(f - f_{end})}}, \tag{A.3}$$

where m_{NL} is the non-linear rate of displacement growth for a BM location tuned to the frequency of the signal, and m_L is the linear growth rate for a BM location not tuned to the frequency of the stimulating signal. The function is a simple logistic curve that rapidly transitions from m_{NL} to m_L at the frequency corresponding to the end of the tonotopic map, f_{end}. Based on studies of the compressive non-linearity in multiple mammalian species (compiled in Robles and Ruggero, 2001), the value for m_{NL} was set to 0.3 dB/dB, and m_L was set to 1 dB/dB (linear growth with pressure). To test the hypothesis that the sharp drop-off in masking amount observed in this study between 90 and 100 kHz is caused by the end of the tonotopic map, the value of f_{end} was set to 95 kHz.

Because the energy in the 140 kHz masking noise was at frequencies above the assumed highest BM center frequency of 95 kHz, BM displacement due to the masker was modeled using the assumption that maximum displacement would occur at the very basal tip of the BM. As the displacement spreads apically, it was assumed to gradually decrease until it reached a magnitude that would not result in masking, i.e. d_Δ. Hence, BM displacement as a function of BM center frequency was modeled as a logistic function of the form:

$$d_m(F) = \frac{A \cdot d_\Delta}{1 + e^{-k(F - F_0)}} + d_\Delta, \tag{A.4}$$

where F is constrained to values less than or equal to 95 kHz, and $d_m(F)$ is in linear units of length.

In order to compare Eq. A.4 to Eq. A.1 (i.e., in order to compare the displacement caused by the signal alone to the displacement caused by the masker alone), Eq. A.4 must be expressed in dB re d_Δ, which gives Eq. 5:

$$D_m(F) = 10 \log \left(\frac{d_m(F)}{d_\Delta} \right). \tag{A.5}$$

Then, comparing the difference in signal displacement and noise displacement to a threshold value gives:

$$T < D_s(p,f) - D_m(F). \tag{A.6}$$

The threshold T was set to 0 dB, i.e., the point at which displacement due to the signal alone equals the displacement due to the masker alone.[4] Masked thresholds were subsequently modeled by combining Eqs. A.1, A.5, and A.6, and solving for pressure, p. The model was then fit to the experimental data by choosing the values for the three parameters determining the shape of the BM displacement—A, k, and, F_0—that minimize the RMS error between the model and the experimental data. The calculated values were $A = 0.2$, k=0.1, and, $F_0 = 62$ Hz. Fig. A.1 shows the resultant BM displacement. Using these parameters, the RMS error was 0.4 dB and the coefficient of determination (R^2) was 0.98.

Appendix B

When conducting underwater hearing experiments in an enclosed environment such as a test pool, acoustic reflections from the sides of the enclosure and from the air/water interface can create unwanted spatial and temporal variability in received signals. Because of this, special care was taken in this study to ensure signal fidelity. As previously mentioned, signal spatial variability was controlled by requiring that the received SPL for all projected sounds was within 3 dB for 24 positions in a cubic grid surrounding the calibration position before testing could begin. Temporal

[4] It is important to note that this model estimates the effects of the signal and noise separately. How the basilar membrane responds when both are presented simultaneously is not known.

variability in received SPL was also accounted for. For timescales greater than the signal duration (trial-to-trial timescales), temporal variability was quantified by measuring received SPL at the calibration position for 25 projections of the same signal over a 25–30 s period, and then calculating the standard deviation. For signals in the 50–180 kHz range, standard deviations increased slightly with increasing frequency, starting at 0.1 dB re 1 μPa at 50 kHz, and reaching a maximum of 0.5 dB re 1 μPa at 180 kHz. These measurements indicate that for all signals in this frequency range, received SPLs for different signal trials should be within ±1 dB of the desired level 95% of the time.

Temporal variability was also quantified on timescales shorter than the duration of a test signal (500 ms). This was achieved using custom MATLAB functions to measure the SPL of one thousand random samples, each 50 ms long, from recorded signals. The maximum SPL across all samples was then compared to the SPL over the full duration of the signal (the calibration level). For signals in the 50–180 kHz range, maximum SPLs for 50 ms samples were 1–2 dB higher than the full-duration SPLs. These measurements indicate that within-signal temporal variability was not a major factor in determining audibility thresholds.

The relatively high SPLs employed in this study further complicated the acoustic situation and made it necessary to carefully inspect signals for low-frequency distortion. In addition to visual inspection of signals prior to each experimental session, signals were recorded for a post-hoc analysis using a low-noise, battery powered Fostex FR2 digital field recorder connected to the Reson 4014 located in the daily calibration position. Signals were projected at 80, 100, 140, and 180 kHz and outgoing voltages were adjusted to obtain a received SPL of 140 dB re 1 μPa. The maximum sampling rate of the FR2 was 192 kHz, limiting recordings to frequencies below 96 kHz, a range sufficient to capture any low-frequency signal distortion that may have been detectable by the subjects (at frequencies greater than 96 kHz, the subjects' detection thresholds were high enough that any audible distortion would have been clearly apparent during visual inspection).

For comparative purposes, recordings of ambient noise in the test environment (received sound at the calibration position in the absence of a projected signal) were taken using the same procedure. Custom MATLAB functions were then used to measure the average power spectral density (PSD) across 1/3-octave bands centered at frequencies ranging from 0.18 to 80 kHz. For each signal, 20 samples were measured. One hundred 500 ms samples of ambient noise were also measured (a duration of 500 ms was chosen to match the signal duration). The results of this analysis (Table B.I) show no evidence of any low-frequency distortion, indicating that subjects were responding to energy at the signal frequency and not distortion or artifacts.

Table B.1
Average PSD levels for 1/3-octave bands ranging from 0.18 to 80 kHz for four signals and ambient noise in the testing environment. Note that the PSD measurements are similar for treatments with signals and for ambient noise, indicating that no low-frequency distortion was present in the signals. The one exception is the 80 kHz 1/3-octave bin in the 80 kHz signal condition, which is expected.

1/3-Octave Bin Freq. (kHz)	Power spectral density (dB re 1 μPa^2/Hz)				
	Ambient (n = 100)	80 kHz signal (n = 20)	100 kHz signal (n = 20)	140 kHz signal (n = 20)	180 kHz signal (n = 20)
0.18	65	64	64	65	64
0.20	62	61	61	63	61
0.25	59	58	58	58	58
0.32	57	56	56	54	56
0.40	54	55	54	54	53
0.50	53	53	53	53	53
0.63	52	51	51	52	51
0.80	51	51	51	52	51
1.25	50	50	51	50	50
1.58	50	50	50	50	50
2.00	50	50	50	50	50
2.50	50	50	49	50	49
3.15	49	50	49	50	49
4.00	49	49	49	50	49
5.00	50	50	50	50	50
6.30	49	49	49	49	49
8.00	49	49	49	49	49
10.0	48	48	48	48	48
12.5	48	48	48	48	48
16.0	48	48	48	48	48
20.0	48	48	48	48	48
25.0	48	48	48	48	48
32.0	48	48	48	48	48
40.0	48	48	48	48	48
50.0	48	48	48	48	48
63.0	48	48	48	48	48
80.0	48	97	50	48	48

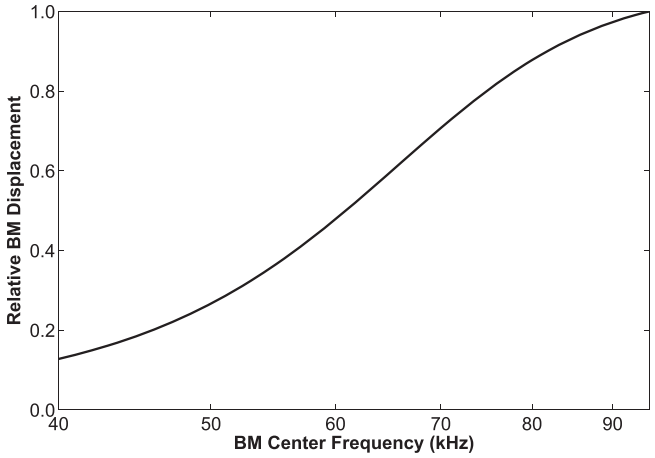

Fig. A.1. Model of BM displacement caused by a 140 kHz narrowband masker. Displacement as a function of BM center frequency was modeled as a logistic function, the parameters of which were determined by fitting the entire masking model to the empirical data using the minimum RMS error. Note that the end of the tonotopic map is assumed to occur at 95 kHz in the model used here.

Appendix C. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.heares.2015.10.002.

References

Corso, J.F., 1963. Bone-conduction thresholds for sonic and ultrasonic frequencies. J. Acoust. Soc. Am. 35, 1738–1743.

Cunningham, K.A., Hayes, S.A., Rub, A.M.W., Reichmuth, C., 2014. Auditory detection of ultrasonic coded transmitters by seals and sea lions. J. Acoust. Soc. Am. 135, 1978–1985.

Deatherage, B.H., Jeffress, L.A., Blodgett, H.C., 1954. A note on the audibility of intense ultrasonic sound. J. Acoust. Soc. Am. 26, 582–582.

Dieroff, H.G., Ertel, H., 1975. Some thoughts on the perception of ultrasonics by man. Arch. Oto-Rhino-Laryngol. 209, 277–290.

Egan, J.P., Hake, H.W., 1950. On the masking pattern of a simple auditory stimulus. J. Acoust. Soc. Am. 22, 622–630.

Finneran, J.J., 2003. An Integrated Computer-controlled System for Marine Mammal Auditory Testing. SSC San Diego, San Diego, CA, pp. 1–107.

Finneran, J.J., Schlundt, C.E., 2007. Underwater sound pressure variation and bottlenose dolphin (Tursiops truncatus) hearing thresholds in a small pool. J. Acoust. Soc. Am. 122, 606–614.

Finney, D.J., 1947. Probit Analysis. University Printing House, Cambridge, England, pp. 1–272.

Fletcher, H., 1940. Auditory patterns. Rev. Mod. Phys. 12, 47.

Hemilä, S., Nummela, S., Reuter, T., 1995. What middle ear parameters tell about impedance matching and high frequency hearing. Hear. Res. 85, 31–44.

Hemilä, S., Nummela, S., Berta, A., Reuter, T., 2006. High-frequency hearing in phocid and otariid pinnipeds: an interpretation based on inertial and cochlear constraints. J. Acoust. Soc. Am. 120, 3463–3466.

Hildebrand, J., 2004. Sources of anthropogenic sound in the marine environment. In: Report to the Policy on Sound and Marine Mammals: an International Workshop. US Marine Mammal Commission and Joint Nature Conservation Committee, London.

Hildebrand, J.A., 2009. Anthropogenic and natural sources of ambient noise in the ocean. Mar. Ecol. Prog. Ser. 395, 5–20.

Kastelein, R.A., Wensveen, P.J., Hoek, L., Verboom, W.C., Terhune, J.M., 2009. Underwater detection of tonal signals between 0.125 and 100kHz by harbor seals (*Phoca vitulina*). J. Acoust. Soc. Am. 125, 1222–1229.

Kastelein, R.A., 2002. Underwater audiogram of a Pacific walrus (Odobenus rosmarus divergens) measured with narrow-band frequency-modulated signals. J. Acoust. Soc. Am. 112, 2173.

Kono, S., Suzuki, Y., Sone, T., 1985. Some consideration on the auditory perception of ultrasound and its effects on hearing. J. Acoust. Soc. Jap. (E) 6, 3–8.

Lenhardt, M.L., 2003. Ultrasonic hearing in humans: applications for tinnitus treatment. Int. Tinnitus J. 9, 69–75.

Møhl, B., 1968a. Hearing in seals. In: Harrison, R.J., Hubbard, R.C., Peterson, R.S., Rice, C.E., Schusterman, R.J. (Eds.), The Behaviour and Physiology of Pinnipeds. Appleton-Century-Crofts, New York, pp. 172–195.

Møhl, B., 1968b. Auditory sensitivity of the common seal in air and water. J. Aud. Res. 8, 27–38.

Moore, B.C., 1993. Frequency analysis and pitch perception. In: Human Psychophysics. Springer, New York, NY, pp. 56–115.

Mulsow, J., Houser, D.S., Finneran, J.J., 2012. Underwater psychophysical audiogram of a young male California sea lion (*Zalophus californianus*). J. Acoust. Soc. Am. 131, 4182–4187.

Nishimura, T., Nakagawa, S., Sakaguchi, T., Hosoi, H., 2003. Ultrasonic masker clarifies ultrasonic perception in man. Hear. Res. 175, 171–177.

NOAA, 2013. Draft Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammals. Retrieved from. http://www.nmfs.noaa.gov/pr/acoustics/guidelines.htm. Last accessed: 8/13/2015.

Pumphrey, R.J., 1950. Upper limit of frequency for human hearing. Nature 167, 438–439.

Qin, M.K., Schwaller, D., Babina, M., Cudahy, E., 2011. Human underwater and bone conduction hearing in the sonic and ultrasonic range. J. Acoust. Soc. Am. 129, 2485–2485.

Reichmuth, C., Southall, B.L., 2012. Underwater hearing in California sea lions (*Zalophus californianus*): expansion and interpretation of existing data. Mar. Mammal. Sci. 28, 358–363.

Reichmuth, C., Holt, M.M., Mulsow, J., Sills, J.M., Southall, B.L., 2013. Comparative assessment of amphibious hearing in pinnipeds. J. Comp. Physiol. A 199, 491–507.

Repenning, C.A., 1972. Underwater hearing in seals: functional morphology. In: Harrison, R.J. (Ed.), Functional Anatomy of Marine Mammals, vol. 1. Academic Press, New York, pp. 307–331.

Robles, L., Ruggero, M.A., 2001. Mechanics of the mammalian cochlea. Physiol. Rev. 81, 1305–1352.

Ruggero, M.A., Temchin, A.N., 2002. The roles of the external, middle, and inner ears in determining the bandwidth of hearing. P. Natl. Acad. Sci. 99, 13206–13210.

Schusterman, R.J., Balliet, R.F., Nixon, J., 1972. Underwater audiogram of the California sea lion by the conditioned vocalization technique. J. Exp. Anal. Behav. 17, 339–350.

Sills, J.M., Southall, B.L., Reichmuth, C., 2014. Amphibious hearing in spotted seals (*Phoca largha*): underwater audiograms, aerial audiograms and critical ratio measurements. J. Exp. Biol. 217, 726–734.

Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene Jr., C.R., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, J.W., Thomas, J.A., Tyack, P.L., 2007. Marine mammal noise-exposure criteria: initial scientific recommendations. Aquat. Mamm. 33, 411–521.

Stenfelt, S., 2011. Acoustic and physiologic aspects of bone conduction hearing. Adv. Oto-Rhino-Laryngol. 71, 10–21.