Masking release at 4 and 32 kHz in harbor seals associated with sinusoidal amplitude-modulated masking noise

Ronald A. Kastelein ; Lean Helder-Hoek ; Linde N. Defillet ; John M. Terhune ; Rainer Beutelmann ; Georg M. Klump

Check for updates

CrossMark

View Online

Export Citation

Masking release at 4 and 32 kHz in harbor seals associated with sinusoidal amplitude-modulated masking noise

Ronald A. Kastelein,[1,a] Lean Helder-Hoek,[1] Linde N. Defillet,[1] John M. Terhune,[2] Rainer Beutelmann,[3] and Georg M. Klump[3]

[1]*Sea Mammal Research Company (SEAMARCO), Julianalaan 46, 3843 CC Harderwijk, The Netherlands*

[2]*Department of Biological Sciences, University of New Brunswick, PO Box 5050, Saint John, New Brunswick, E2L 4L5, Canada*

[3]*Cluster of Excellence "Hearing4all" and Research Centre Neurosensory Science, Department of Neuroscience, School of Medicine and Health Science, Carl von Ossietzky University Oldenburg, 26111 Oldenburg, Germany*

ABSTRACT:

Masking can reduce the efficiency of communication and prey and predator detection. Most underwater sounds fluctuate in amplitude, which may influence the amount of masking experienced by marine mammals. The hearing thresholds of two harbor seals for tonal sweeps (centered at 4 and 32 kHz) masked by sinusoidal amplitude modulated (SAM) Gaussian one-third octave noise bands centered around the narrow-band test sweep frequencies, were studied with a psychoacoustic technique. Masking was assessed in relation to signal duration, (500, 1000, and 2000 ms) and masker level, at eight amplitude modulation rates (1–90 Hz). Masking release (MR) due to SAM compared thresholds in modulated and unmodulated maskers. Unmodulated maskers resulted in critical ratios of 21 dB at 4 kHz and 31 dB at 32 kHz. Masked thresholds were similarly affected by SAM rate with the lowest thresholds and the largest MR being observed for SAM rates of 1 and 2 Hz at higher masker levels. MR was higher for 32-kHz maskers than for 4-kHz maskers. Increasing signal duration from 500 ms to 2000 ms had minimal effect on MR. The results are discussed with respect to MR resulting from envelope variation and the impact of noise in the environment on target signal detection. © 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).
https://doi.org/10.1121/10.0019631

I. INTRODUCTION

The average background underwater noise level has increased during the last century due to global industrialization (McDonald *et al.*, 2006; McDonald *et al.*, 2008). Anthropogenic noise—such as that generated by shipping, oil and gas exploration and production, wind turbine installation and operation, and military sonar—can affect marine mammals by displacing them from biologically important areas (Duarte *et al.*, 2021; Southall *et al.*, 2021), masking relevant signals (Erbe *et al.*, 2016; Branstetter and Sills, 2022), or resulting in a physiological reduction in hearing sensitivity that may be temporary (temporary threshold shift, TTS) or permanent (permanent threshold shift, PTS; Southall *et al.*, 2019). Auditory masking can interfere with an animal's ability to detect or recognize prey, detect other animals such as predators, navigate, and communicate with conspecifics (Bain and Dahlheim, 1994; Erbe *et al.*, 2016).

Masking occurs when one sound (the noise) interferes with the detection of another sound (the signal). The degree of interference depends upon a range of factors, including the relative amplitudes of the two sounds, the degree of spectral overlap between signal and noise, and the temporal and directional properties of the masker and signal. Also, the listener's perception of the sounds, including auditory frequency filter bandwidths and temporal resolution abilities will influence the masking levels (Branstetter and Sills, 2022). The lowest signal-to-noise ratio at which a subject can detect a tonal signal in a broadband continuous Gaussian masking noise is defined as the critical-masking ratio (CR) (see Fletcher, 1940; Hawkins and Stevens, 1950). The lower an animal's CR, the better its ability to detect a signal in a broadband noise. The CRs can be used to calculate detection threshold levels of signals under certain continuous and flat-spectrum background noise conditions (Scharf, 1970). The common assumption for calculating the CR bandwidth is that, at the detection threshold for the signal, the amount of energy of the masker is equal to the amount of energy of the tone in the auditory filter processing both signals. Frequencies beyond this bandwidth do not contribute to the masking by broadband continuous Gaussian noise (Scharf, 1970). On the other hand, the properties of the temporal envelope of masker and signal will determine its detectability (Verhey *et al.*, 2003). Signal detection in amplitude-modulated broadband maskers with temporally coherent level fluctuations across frequency is considerably improved relative to incoherent fluctuations across frequency, a condition described as comodulation masking

a)Electronic mail: rk@seamarco.nl

release (CMR) (Hall *et al.*, 1984). Masking release (MR) often is defined as the level of masking (dB) in a certain condition relative to the level of masking caused by continuous Gaussian noise at the same amplitude and frequency. This difference has also been referred to as the "modulated unmodulated difference" (MUD) (see Bacon *et al.,* 1997).

In marine mammals such as seals, the hearing thresholds increase when the background noise level increases above a certain level (Terhune and Ronald, 1975; Turnbull and Terhune, 1990; Southall *et al.*, 2000; for overviews see Erbe *et al.*, 2016 and Branstetter and Sills, 2022). However, studies of masking in marine mammals have, for the most part, been limited to simple stimuli such as continuous, flat spectrum, random Gaussian noise maskers, and pure tone signals. This is not realistic, as in the seas and oceans, underwater noise is not Gaussian white noise and is generally not continuous and of constant amplitude. Branstetter and Finneran (2008) showed that in bottlenose dolphins (*Tursiops truncatus*) temporally fluctuating comodulated noise produces up to 17 dB lower masked thresholds compared to constant-amplitude continuous Gaussian noise of the same spectral density level. In the harbor porpoise (*Phocoena phocoena*), Kastelein *et al.* (2021) observed a release from masking by slow sinusoidal amplitude modulation of the masker, of up to 15 dB. The results by Branstetter and Finneran (2008) and Kastelein *et al.* (2021) suggest that conventional models of masking derived from experiments using random Gaussian noise may not generalize well to conditions with environmental noise that marine mammals encounter in real life situations, but these laboratory conditions can provide a reasonable conservative initial estimate of masking.

So far, mainly masking due to continuous random Gaussian white noise has been studied in harp seals (*Pagophilus groenlandicus;* Terhune and Ronald, 1971), ringed seals (*Pusa hispida;* Terhune and Ronald, 1975; Sills *et al.*, 2015), harbor seals (*Phoca vitulina;* Turnbull and Terhune 1990; Southall *et al.*, 2000), spotted seals (*Phoca largha;* Sills *et al.*, 2014), bearded seals (*Erignathus barbatus;* Sills *et al.*, 2020), Hawaiian monk seals (*Neomonachus schauinslandi;* Sills *et al.*, 2021) and northern elephant seals (*Mirounga angustirostris;* Southall *et al.*, 2000). The CRs derived in these studies, when used in environmental impact assessment models, can overestimate the masking effect of natural and anthropogenic sounds, as these are generally not constant in amplitude and are spectrally complex. Cunningham *et al.* (2014) studied masking in a harbor seal using complex signals and found that signals with frequency modulation, amplitude modulation and harmonic structure showed lower hearing thresholds in both Gaussian broad-band noise with time-varying amplitude and shipping noise than pure-tone signals masked by constant amplitude white noise. Thus, signal detection can be affected by the temporal structure of both signals and maskers.

The harbor seal has an extensive geographical range spanning the Baltic Sea as well as both the eastern and western coasts of the Atlantic and Pacific Oceans. It leads an amphibious life, resting and pupping on land, but migrating,

foraging, and performing courtship underwater (Burns, 2002). Harbor seals often swim in murky water, are active at night, and dive to depths where light hardly penetrates even during the day. Therefore, these animals often depend, apart from using their whiskers (Murphy *et al.*, 2015), on sound rather than vision for orientation and communication. Harbor seals can hear well both in air (Møhl, 1968, Terhune, 1991; Wolski *et al.*, 2003; Reichmuth *et al.*, 2013) and under water (Møhl, 1968; Terhune, 1988; Turnbull and Terhune, 1993; Kastak and Schusterman, 1998; Kastelein *et al.*, 2009a; Kastelein *et al.*, 2009b; Kastelein *et al.*, 2010; Reichmuth *et al.*, 2013). In the sea, noise is produced naturally by, for instance, wind, waves, geological (abiotic) events, and biological activities including calls from other animals such as non-conspecifics, conspecifics, prey, and predators (Urick, 1983). Therefore, seal hearing must have evolved to function in the presence of interfering noise from natural sources that typically have varying amplitudes and can have a wide bandwidth.

Here, we investigate the masking effect of noise with a time-varying amplitude on signal detection in two harbor seals. We quantified MR associated with sinusoidally amplitude-modulated (SAM) masking noise centered around 4 and 32 kHz in two harbor seals. The goal was to gain insight into the factors affecting the harbor seals' ability to detect tonal signals in SAM noise. We assessed the effect of the SAM rate, masking noise level, and the duration of the tonal test signal, on the amount of MR. Such information can aid in providing more realistic environmental masking impact assessments in relation to underwater noise.

II. METHODS

A. Study animals

The study animals were two female harbor seals (F01 and F02), which were captive-born in the same year, and were moved to the SEAMARCO Research Institute soon after they had been weaned. During the study they were healthy, aged from 9 to 13 years old, and of similar size; their body lengths were ~145 cm, and their body weight varied between ~40 kg (in summer) and ~65 kg (in winter). The two seals had sensitive hearing, very similar hearing thresholds and were very familiar with the signal detection task as they have participated in over 12 psychophysical hearing studies (Kastelein *et al.*, 2009a; Kastelein *et al.*, 2009b; Kastelein *et al.*, 2010; Kastelein *et al.*, 2015, Kastelein *et al.*, 2018a; Kastelein *et al.*, 2018b; Kastelein *et al.*, 2019c; Kastelein *et al.*, 2019a; Kastelein *et al.*, 2019b; Kastelein *et al.*, 2020a; Kastelein *et al.*, 2020b; Kastelein *et al.*, 2020c). The hearing abilities of these two seals are likely indicative of the abilities of healthy harbor seals in general, and perhaps also for several other phocid species (Southall *et al.*, 2019).

The seals consumed thawed fish divided into four equal meals per day, three or four of which were given during research sessions. Variation in the animals' performance was minimized by making weekly adjustments (usually on the order of 100 g) to their daily food ration, based on their weight, performance during the previous week, and the

expected change in water and air temperatures in the following week. Their diet adjustments did not result in reducing their weights below normal levels associated with the prevailing seasonal temperatures.

B. Study facility

The study was conducted at the SEAMARCO Research Institute, which is in a remote area that was specifically selected for acoustic research. The measurements were conducted in an outdoor pool (8 m × 7 m, 2 m deep) with an adjacent haul-out platform. The lower half of the pool was below the ground level (see Fig. 2). The pool walls were covered with aquatic vegetation which reduced sound reflections. The bottom of the pool was covered with approximately 20 cm of sloping sand. Skimmers kept the water level constant, and seawater was pumped in directly from the nearby Eastern Scheldt, a lagoon of the North Sea. Twenty percent of the water was replaced daily. Most of the water (80%) was re-circulated daily through a biological filter system to ensure year-round water clarity, so that the animals' behavior could be observed via underwater cameras during the test sessions.

To limit the amount of noise that the seals were exposed to constantly, the water circulation system and the aeration system for the bio-filter were designed to be as quiet as possible. There was no current in the pool during the experiments, as the water circulation pump and the air pump of the bio-filter were switched off at the beginning of the day at least 60 min before the first test session, and remained off for the remainder of the working day. This also reduced flow noise from the skimmers. The water temperature varied between 0 °C in February and 22 °C in August, and the salinity was around 3.4%.

C. Acoustics

1. Test signal

The animals were trained (using positive reinforcement conditioning) to report the detection of narrow band up-sweeps, by swimming away from the listening station. A linear frequency-modulated tone was presented instead of a pure tone because sweeps resulted in very stable and precise thresholds (Finneran and Schlundt, 2007). The test signal was generated digitally (Adobe Audition, version 3.0). Linear up-sweeps centered at 4 and 32 kHz started and ended at ±2.5% of the center frequency (i.e., 3.9–4.1 kHz; 31.2–32.8 kHz) and had a duration of 500, 1000, or 2000 ms, including a 50-ms linear rise and fall in amplitude. All tone durations were longer than the integration time without a masking noise (Kastelein et al., 2010). Four kHz was chosen at the low end of the harbor seal hearing range, which could also be heard by the operators so they could better monitor the responses of the seals (especially during training), and 32 kHz was chosen as a frequency at the high end of harbor seal hearing (Kastelein et al., 2009a).

Test signals were played on a laptop computer (Acer Aspire 5750 model P5WEO, Acer, New Taipei City, Taiwan)

using LabVIEW and an external AD/DA card (National Instruments USB-6251, National Instruments, Austin, TX) with the output level being controlled with the LabVIEW program in steps of 2 dB. The output of the card passed through a ground loop isolator, a custom-built low-pass filter/buffer, a custom-built passive low-pass filter together serving as an anti-aliasing filter, a custom built-mixer and an isolation transformer (Lubell AC202, Lubell, Fort Lauderdale, FL) being broadcast by a balanced tonpilz piezoelectric acoustic transducer (Lubell LL916; 4 kHz), or (but without the isolation transformer) a high frequency transducer (EDO Western 337, 32 kHz, EDO Western, New York).

2. Masking noise

The masking noise was 1/3-octave bandwidth Gaussian white noise centered on 4 or 32 kHz (Filter: 6th order Butterworth). The reasons for choosing a 1/3 octave noise band instead of 1/1 octave noise bands were because it is much more difficult to produce a flat plateau for 1/1 octave noise bands than for 1/3 octave noise bands, and we believe that a 1/3 octave noise band is wider than the critical bandwidths of mammals. The rates of SAM were 1, 2, 5, 10, 20, 40, 80, and 90 Hz. The noise levels, measured over 10 s, did not vary more than 0.2 dB between the SAM rates (Table I). The constant-amplitude masking noise used for measuring the critical masking ratio had a 4 dB higher amplitude than the SAM noise. In each testing session, the noise level was controllable with an accuracy of 1 dB. The spectrum density level (SDL) (dB re 1 μPa^2/Hz; Ainslie et al., 2022) did not vary by more than 2 dB across the 1/3-octave band.

The digitally generated 18-s masking noise (Adobe Audition 3, and Audacity 2.2.0 (Audacity, 2017; WAV file; sample rate, 768 kHz) was played back by a second laptop computer (Acer Aspire V5 series model ZRI) using LabVIEW and an external AD/DA card (National Instruments USB-6251). The output of the card passed through a ground loop isolator and custom-built low-pass buffer/filter (anti-aliasing filter) to the custom-built mixer where the masking noise was mixed with the above-mentioned hearing test signal and was broadcast by the above-mentioned transducers. The linearity of the transmitter system of the masking noise was checked during each calibration and was found to deviate at most by 1 dB from the expected value within a 40 dB range.

The masker conditions were as follows: (1) No masker [i.e., ambient noise only, see Figs. 1(a) and 1(b)]; (2) SAM noise (AM depth of the electric signal going to the transducer 100%, AM frequency 1, 2, 5, 10, 20, 40, 80, or 90 Hz); (3) constant-amplitude noise (0% AM depth of the electric signal going to the transducer).

3. Calibration procedures and background noise

The amplitudes of the background noise, the received masking noise, and received hearing test signals were measured once every two months during the study period by an external company (TNO). The sound-measurement equipment consisted of two hydrophones [Brüel and Kjaer (B&K) – 8106;

J. Acoust. Soc. Am. **154** (1), July 2023

Kastelein et al.   83

TABLE I. Depth of modulation (MD, %), power-spectral density (SDL; dB re 1 μPa^2/Hz), and peak-to-trough differences (PTD) in sound pressure level (dB SPL) for 4-kHz and 32-kHz maskers in relation to the modulation frequency. Data show averages measured with two hydrophones in the setup at the positions of the seal's left and right ear. Constant level refers to a masker without amplitude modulation. Overall masker levels in the recordings for this analysis were 94 dB and 101 dB (re 1 μPa) for 4 kHz and 32 kHz maskers, respectively.

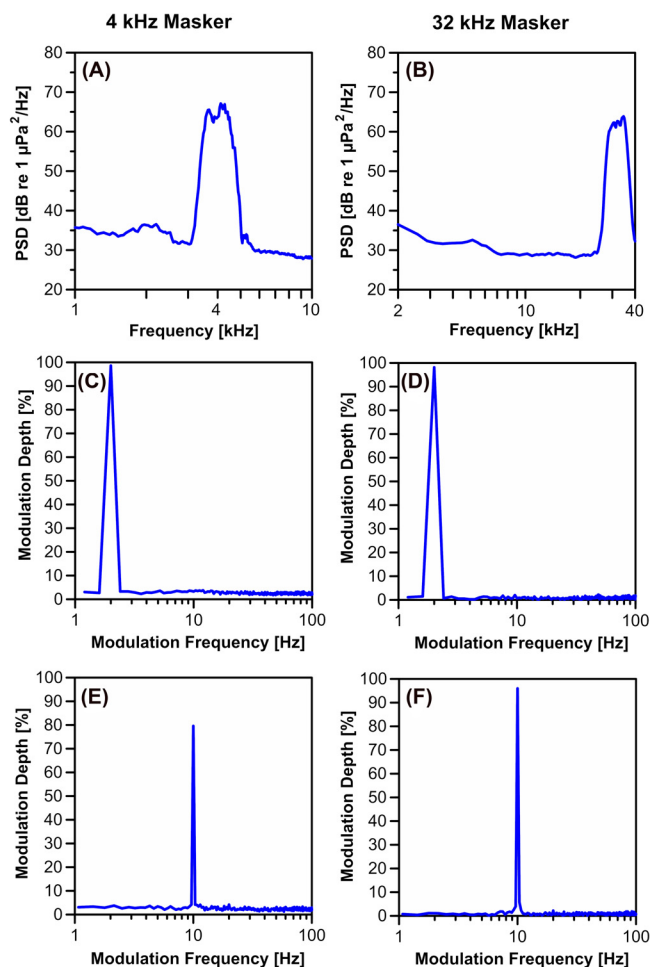| Modulation Frequency (Hz)	4 kHz MD	32 kHz MD	4 kHz SDL (dB re 1 μPa^2/Hz)	32 kHz SDL (dB re 1 μPa^2/Hz)	4 kHz PTD (dB)	32 kHz PTD (dB)
Constant level	0.5%	0.3%	69.1	66.5	0.0	0.0
1	97.8%	99.2%	64.8	62.1	33.2	41.9
2	98.1%	98.6%	64.7	62.3	34.4	37.1
5	90.8%	98.1%	64.7	62.1	20.7	34.4
10	73.4%	96.0%	64.7	62.1	11.5	28.0
20	43.8%	89.3%	64.6	62.2	5.0	19.4
40	29.7%	79.3%	64.6	62.1	3.1	13.7
80	20.9%	60.6%	64.7	62.1	2.0	8.1
90	20.1%	59.8%	64.7	62.1	1.9	7.9

FIG. 1. (Color online) Analysis results for exemplary hydrophone recordings of 4-kHz and 32-kHz SAM maskers broadcast in the pool. (A), (B) Masker power spectra (power spectral density PSD dB re 1 μPa^2/Hz) for 2-Hz SAM maskers. (C), (D) masker modulation spectra at 2 Hz SAM. (E), (F) masker modulation spectra at 10 Hz SAM. Masker power spectra represent both the masker frequencies (third-octave bandwidth peaks) and the ambient noise floor (PSD below and above the peaks). Graphs showing the full analysis for modulation frequencies between 1 Hz and 90 Hz are provided as supplementary material.[1] Depth of modulation and masker levels for all masker conditions are presented in Table I and in the supplementary material.[1]

Brüel and Kjaer, Naerum, Denmark] with a multichannel high frequency analyzer (B&K PULSE 3560D) and a laptop computer with B&K PULSE software (Labshop, version 12.1). The system was calibrated with a pistonphone (B&K 4223). The broadband sound pressure level (SPL, dB re 1 μPa) of the hearing test signal was derived from the received 90% energy flux density and the corresponding 90% time duration (t_{90}) (Madsen, 2005). The recording time over which the SPL of the masking noise was determined was 10 s.

The level of background noise in the pool was very low. At 4 kHz, the 1/3-octave band ambient noise level was ~62 dB re 1μPa [SDL: 32 (dB re 1 μPa^2/Hz)]. At 32 kHz, the 1/3-octave band ambient noise level was ~72 dB re 1μPa [SDL: 33 (dB re 1 μPa^2/Hz)].

For calibration of both the hearing test signals and the masking noise, the received SPL of each sound was measured with a hydrophone at the position where the left and right auditory meatus of the harbor seal would be while the seal was at the listening station. The received SPLs were calibrated at levels of approximately 10–40 dB above the threshold levels found in the present study. The linearity of the transmitter system was checked during calibration and deviated at most by 1 dB from the expected value within the 30 dB range. The SPL at the two locations varied by up to 2 dB on different calibration-measurement days. The average SPL of the two hydrophones was used to calculate the received stimulus SPL during hearing threshold tests.

Since the actual acoustic AM depths could differ from the 100% AM depth in the electrical signal due to ambient noise and acoustic reflections in the pool, we analyzed the masking noise recordings made through the two hydrophones positioned at the precise position of the seals' auditory meatus in the experiment (during testing the animals did not move their heads more than 2 cm in each direction in the horizontal plane, and within 1 cm in each direction in the vertical plane). Exemplary measurement results for power spectra and modulation spectra of 4 and 32 kHz maskers are shown in Fig. 1 (for a complete set of power spectra and modulation spectra see the supplementary material[1]). In all

conditions, the maskers were presented with a spectrum level (SDL) that was at least 20 dB above the spectrum level of the background noise. For the 32-kHz masker, the acoustically determined depth of modulation was above 95% at modulation frequencies up to 10 Hz and then decreased to 60% as the modulation frequencies were increased to 90 Hz. For the 4-kHz masker, the acoustically determined depth of modulation was above 95% at modulation frequencies below 5 Hz and it decreased to 20% with modulation frequencies increasing to 90 Hz. Table I shows the data for all acoustically determined depths of modulation.

Before each test session, the voltage output of the emitting system to the transducer was checked with a voltmeter (Agilent 34401 A; Agilent, Santa Clara, CA). The acoustic underwater signal was checked with a B&K 8101 hydrophone at the listening station, a pre-amplifier (B&K 2365) and a spectrum analyzer (Velleman PCSU1000; Velleman, Fort Worth, TX). If the values were similar (i.e., within 2 dB) as those obtained during the SPL measurements by the external company, the SPLs were assumed to be correct, and a hearing test could be performed. Additional care was taken to make the harbor seal's listening environment as quiet as possible. Only researchers involved in the hearing tests were allowed within 15 m of the pool during hearing test sessions, and they were required to be quiet and stand still during the tests. Furthermore, before each session, the background noise level was checked with the hydrophone and spectrum analyzer to make sure it did not deviate too much from the general low background noise level.

D. Experimental procedure

A behavioral technique was used for audiometric testing. A trial began with the test animal at the start/response buoy and the other animal near the haul-out platform with the other trainer (Fig. 2). During the ~12-min hearing test
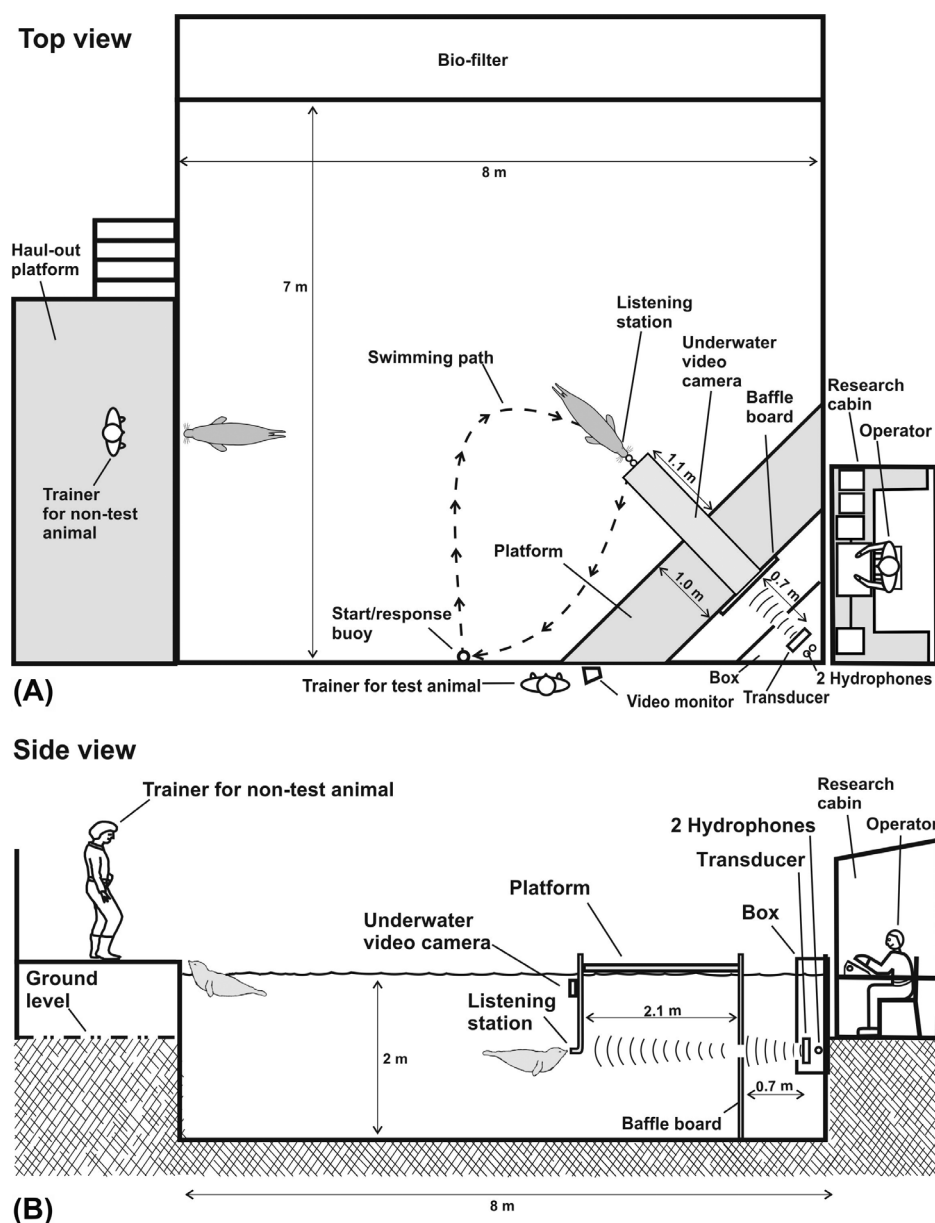

FIG. 2. The study area, showing the test harbor seal in position at the underwater listening station, and the non-test animal with the other trainer; (A) top view and (B) quasi side view, both to scale.

sessions, the animal not being tested was trained to keep very still in the water, as any waves might increase the hearing threshold of the test animal. Fish rewards were given to the non-test animal at the same time when the test animal was rewarded, to prevent distraction during the trials. The signal operator and the equipment used to produce the stimuli and listen to underwater sounds were in a research cabin next to the pool, out of sight of the animals. The level of the hearing test sound used in the first trial of the session was approximately 6 dB above the hearing threshold determined during the previous sessions with the same hearing test signal and noise parameters. The harbor seal was trained to swim to the listening station in response to a hand signal from the trainer. The methods were as described in detail by Kastelein *et al.* (2019b).

At the same moment when the operator hand-signaled to the trainer to send the animal to the listening station, the masking noise was switched on and remained on for 18 s. It took the seal about 2 s to reach the listening station. When the trainer gave a hand signal, the seal being tested swam to the listening station (Fig. 2). When at the listening station, the seal could not see the trainer, who was not aware of the trial type. Signals were produced at a random time 4–12 s after the seal stationed at the listening station. The signal level was varied according to the one-up one-down adaptive staircase method (Cornsweet, 1962), and 2 dB steps were used. This conventional psychometric technique (Robinson and Watson, 1972) produces a 50% correct detection threshold (Levitt, 1971). A switch from a test signal level that a harbor seal responded to (a hit), to a level that she did not respond to (a miss), and vice versa, was called a reversal. Each complete hearing session consisted of ~25 trials and lasted for ~12 min. Sessions consisted of 2/3 signal-present and 1/3 signal-absent trials (also called "catch trials") offered in quasi-random order; there were never more than three consecutive signal-present or signal-absent trials.

Generally, four experimental sessions per day were conducted seven days per week (around 0900, 1130, 1400, and 1600 h). Hearing tests were conducted under three general conditions: the quiet, no added noise condition [see Figs. 1(a) and 1(b)], one of the eight SAM noise conditions, and the constant amplitude noise condition. These conditions were tested in random order (one condition per session). Each noise condition was tested until a total of 20 reversals were obtained in two sessions (ten reversals per session). All collected data were used in the analysis. Data were collected between July 2015 and December 2016.

E. Experimental conditions

Detection thresholds were determined under an ambient noise condition and with constant-amplitude or SAM masking noise using signals with a duration of 500, 1000, or 2000 ms. The modulation depth of the computed SAM masker was always 100%; the SPL of the electronic signal

going to the transducer varied by ≥ 40 dB. Reverberation within the testing area resulted in lower actual modulation depths, especially at the higher AM rates and at 4 kHz (Table I). When constant amplitude Gaussian noise was amplitude modulated, the long-term RMS SPL was reduced due to the modulation (see also Table I). To account for such differences, masked thresholds are reported as the signal-to-noise ratio relative to the spectral density level of the masker. In the results, the threshold values obtained in different sessions with a unique set of parameters were combined for further statistical analysis. The masking release (MR) due to the masker amplitude modulation was calculated by subtracting the threshold for tones in the SAM masker from the threshold for the same tones in the constant-amplitude masker (i.e., the critical ratio).

The pre-stimulus response rates were recorded for each staircase threshold measurement (20 reversals across two test sessions). The relationships between the pre-stimulus response rates of each individual seal and the masking noise levels, masking noise bandwidths, and signal durations were examined.

1. Effect of the SAM rate of the masking noise

At low SAM rates, the duration of low-amplitude dips within masking noise is longer than at high SAM rates. If harbor seals are employing dip listening as a mechanism to reduce masking (Buus, 1985), longer dip durations should facilitate signal detection. In each staircase measurement of the threshold, the SAM rate and the other parameters (frequency, masker level, signal duration) were kept constant while the signal level was adapted in the up/down procedure. Masked thresholds and MR were examined using AM rates of 1, 2, 5, 10, 20, 40, 80, and 90 Hz and unmodulated noise.

2. Effect of the duration of the test signal

When a listener attempts to detect a tonal signal being masked by a broadband SAM noise, the signal is most likely detected during dips in the masking noise's amplitude. The longer the duration of the signal, the higher the probability of detection, because longer signals are potentially audible during a larger number of dips than shorter signals (Buus, 1999; Cooke, 2006). Masked thresholds and MR were examined using signal durations of 500, 1000, or 2000 ms. In the experiments investigating the effect of signal duration, the masker spectral density level was 74 dB re 1 μPa^2/Hz.

3. Effect of the level of the masking noise

Typically, anthropogenic noise, in addition to natural background noise, impacts hearing thresholds. The SAM noise masker that was used to simulate temporally varying anthropogenic noise, will not interact with the more constant-amplitude ambient noise if presented well above the level of the ambient noise. However, when the SAM

noise masker is closer in level to the ambient noise, a listener's perception may be affected by the combination of both the SAM masker and the ambient noise. This is especially likely to occur during the low-level dips in the SAM noise when ambient noise levels can become effective. Masker spectral density levels were varied between 54 dB re 1 μPa^2/Hz and 74 dB re 1 μPa^2/Hz. Thus, the MR is predicted to be less for the low-amplitude SAM masking noise than for the high-amplitude SAM masking noise. To investigate this possibility, the seals were presented with several AM masking noise levels that ranged from just above the ambient noise level to clearly above the ambient noise level.

III. RESULTS

The seals' responses were under good stimulus control. The mean pre-stimulus response rate at 4 kHz was 5 ± 4.3% (range 0%–20%) in seal F01, and 4 ± 3.9% (range 0%–18%) in seal F02. At 32 kHz, the mean pre-stimulus response rate was 6 ± 3.7% (range 0%–16%) in seal F01, and 5 ± 3.7% (range 0%–16%) in seal F02. There was no significant difference between the pre-stimulus response rates of the two individuals indicating similar criteria for reporting the tonal signals. Pre-stimulus response rates did not change throughout the study and it follows that the signal-detection criteria were stable throughout the threshold measurements. Given the similar pre-stimulus response rates for different hearing test signal and masking noise combinations, varying false-alarm rates (i.e., a change in criterion) are unlikely to have affected the results.

The seals' unmasked mean 50% detection hearing thresholds for 500 ms signal durations were 55 (F01) and 51 (F02) dB re 1 μPa at 4 kHz and 65 (F01) and 60 (F02) dB re 1 μPa at 32 kHz. The seal's unmasked mean 50% detection hearing thresholds for 1000 ms signal durations were 56 (F01) and 55 (F02) dB re 1 μPa at 4 kHz and 65 (F01) and 65 (F02) dB re 1 μPa at 32 kHz. The seal's unmasked mean 50% detection hearing thresholds for 2000 ms signal

durations were 52 (F01) and 48 (F02) dB re 1 μPa at 4 kHz and 65 (F01) and 59 (F02) dB re 1 μPa at 32 kHz.

A. Effects of SAM rate and signal duration on masked thresholds and masking release

The seals' CR, i.e., the signal-to-noise ratio for detecting a tone in a constant-amplitude masker with a third-octave bandwidth, was affected by the signal frequency and signal duration (mixed-model analysis of variance, ANOVA, main effect signal frequency $F[1,6] = 343.4$, $p < 0.0005$, main effect signal duration $F[2,6] = 19.31$, $p < 0.003$, two-way interaction $F[2,6] = 6.29$, $p < 0.05$). The average CRs were 22.7 and 30.7 dB at 4 and 32 kHz, respectively (Fig. 3). At 4 kHz, the CR decreased from 24.8 dB for a 500-ms signal, to 23.5 dB for a 1000-ms signal, to 19.9 dB for a 2000-ms signal. At 32 kHz, the CR remained relatively constant for signals of different duration. If the average CR is taken as an estimate of the rectangular auditory filter bandwidth, the auditory filter bandwidths were 186 Hz (0.07 oct) and 1175 Hz (0.05 oct) for 4 and 32 kHz, respectively.

Compared to thresholds in constant-amplitude maskers, masked thresholds were considerably reduced in one-third octave bandwidth SAM noise with slow (i.e., $\leq$ 5 Hz) amplitude modulations (Fig. 3). Masked thresholds in SAM noise were affected by signal frequency, signal duration and SAM rate (mixed-model ANOVA, dependent variable signal-to-noise ratio at threshold, main effect signal frequency $F[1,48] = 53.3$, $p < 0.0005$, main effect signal duration $F[2,48] = 24.3$, $p < 0.0005$, main effect SAM rate $F[7,48] = 103.0$, $p < 0.0005$). Similar to the thresholds in constant-amplitude noise, thresholds in SAM noise increased with increasing signal frequency and decreased with increasing signal duration. This decrease was smaller at 32 than at 4 kHz indicated by a significant interaction between signal frequency and signal duration ($F[2,48] = 16.7$, $p < 0.0005$). The lowest masked thresholds were observed at SAM rates of 1–2 Hz. The masked threshold increased with increasing SAM frequency above the modulation frequency of 2 Hz, reaching values that were close to the
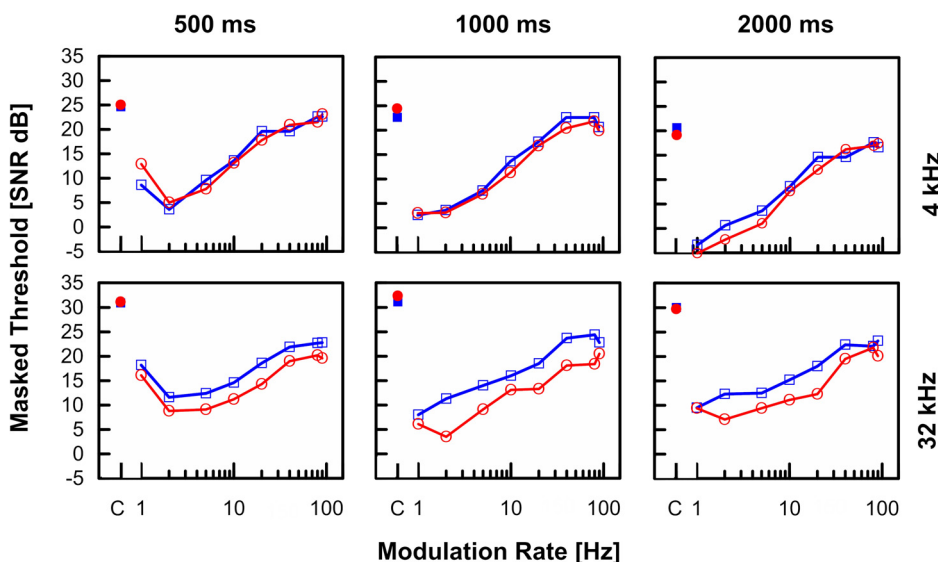

FIG. 3. (Color online) Detection thresholds for 4 kHz and 32 kHz tone signals embedded in a third-octave bandwidth masker in relation to the SAM rate of the masker (the C on the X-axis indicates a constant-amplitude masker). Masker spectral-density level was 74 dB re 1 μPa^2/Hz. Signal duration indicated above the columns was an additional parameter. Open symbols show data obtained with SAM maskers, and filled symbols data obtained with constant-amplitude maskers. The colors represent the two different seals.

CR for 4 kHz signals at modulation rates of 90 Hz. For 32-kHz signals, the masked thresholds did not increase as much with increasing modulation frequency resulting in masked thresholds that were below the CR at a modulation rate of 90 Hz (the mixed-model ANOVA revealed a significant interaction between signal frequency and SAM rate, $F[7,48] = 5.97$, $p < 0.0005$). The mixed-model ANOVA also revealed a significant interaction between the effects of signal duration and SAM rate ($F[14,48] = 2.83$, $p < 0.05$) on the masked detection threshold. No three-way interaction was observed.

Since the depth of modulation co-varies with the modulation rate (Table I), the question arises if the effects of both factors can be separated. The results of partial correlation analyses indicate that at 32 kHz both modulation rate and depth of modulation affect the masked thresholds, whereas at 4 kHz, the effect of both parameters is not separable. For 32 kHz, threshold was highly correlated with the log-transformed modulation rate (zero order correlation $r = 0.831$, $p < 0.001$). The corresponding partial correlation in which we controlled for the depth of modulation still showed a significant, although smaller, correlation with $r = 0.390$ ($p = 0.007$). This indicates that, in the present study, modulation rate alone can affect the masked threshold at 32 kHz. For 4 kHz, however, the high correlation between threshold and the log-transformed modulation rate (zero order correlation $r = 0.889$, $p < 0.001$) vanished (i.e., the correlation was much smaller and was no longer significant) if we controlled for the depth of modulation in a partial correlation analysis.

The amount of masking release (Fig. 4) was affected by signal frequency and SAM rate, but not by signal duration (mixed-model ANOVA, main effect signal frequency $F[1,48] = 94.3$, $p < 0.0005$, main effect SAM rate $F[7,48] = 79.4$, $p < 0.0005$, main effect signal duration $F[2,48] = 1.94$, n.s.). There were significant interactions between the effects of signal frequency and SAM rate ($F[7,48] = 4.60$, $p = 0.0005$), signal duration and SAM rate ($F[14,48] = 2.17$, $p < 0.05$) and the effects of signal frequency and signal duration ($F[2,48] = 3.84$, $p < 0.05$; a three-way

interaction was not observed). The highest masking release was observed for modulation frequencies of 1 or 2 Hz, and the masking release generally decreased with increasing SAM rate above a SAM rate of 2 Hz. The decline of the masking release differed between the two signal frequencies (see significant interaction), which may be related to differences in the acoustic depth of modulation observed with hydrophone recordings (see Table I and supplementary material[1]). At 32 kHz, there was less reduction in the acoustically determined depth of modulation with increasing modulation frequency (99.2% and 98.6% at 1 and 2 Hz SAM rate, respectively, to 59.8% at 90 Hz SAM rate) compared to 4 kHz for which the corresponding depth of modulation decreased considerably with increasing modulation rate (from 97.8% and 98.1% at 1 and 2 Hz SAM rate respectively to 20.1% at 90 Hz SAM rate). Thus, the higher decrease in the amount of masking release at 4 kHz can be attributed to both the effects of SAM rate and related differences in the acoustical depth of modulation, whereas at 32 kHz the decrease can be primarily attributed to the SAM rate up to a modulation frequency of 20 Hz. The differences in the depth of modulation may also contribute to the main effect of frequency and to the other interactions.

B. Effects of masker level on masked thresholds and masking release

If the masker level is reduced, the amount of energetic masking is decreased, thus reducing masked thresholds. At the same time, ambient noise potentially exerts a stronger effect on the modulation depth of the masker, decreasing the actual depth of modulation with decreasing masker levels which may affect thresholds, especially at low SAM rates (Fig. 5). In these measurements, signal duration was constant at 1000 ms. A mixed-model ANOVA with the signal-to-noise ratio at the masked threshold as the dependent variable revealed significant main effects of signal frequency ($F[1,48] = 107.7$, $p < 0.0005$), noise spectral density level ($F[2,48] = 62.2$, $p < 0.0005$) and of SAM rate
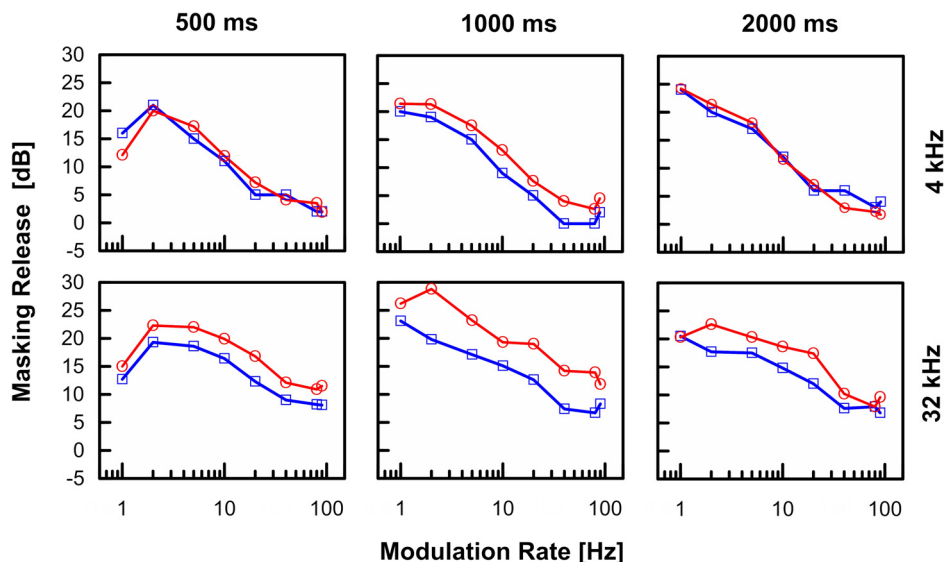

FIG. 4. (Color online) Masking release for 4 kHz and 32 kHz tone signals embedded in a third-octave bandwidth masker in relation to the SAM rate of the masker. Masker spectral-density level was 74 dB re 1 μPa^2/Hz. Signal duration indicated above the columns was an additional parameter. The colors represent the two different seals.
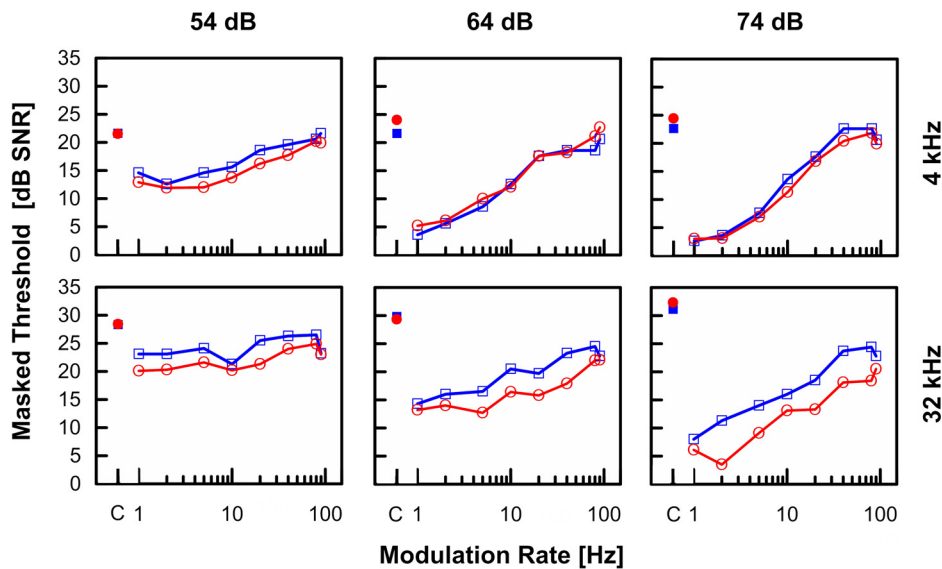
FIG. 5. (Color online) Signal-to-noise ratios (dB SNR) for 4 kHz and 32 kHz tone signals embedded in a third-octave bandwidth masker in relation to the SAM rate of the masker (The C on the X-axis indicates a constant-amplitude masker). The masker spectral density level (dB re 1 μPa^2/Hz) indicated above the columns was an additional parameter. Signal duration was 1000 ms. Open symbols show data obtained with SAM maskers, and filled symbols data obtained with constant-amplitude maskers. The colors represent the two different seals.

$(F[7,48] = 67.0, \ p < 0.0005)$. The first two main effects, as well as a significant interaction of signal frequency and noise spectral density level $(F[2,48] = 12.2, \ p < 0.0005)$ can be explained by differences in auditory energetic masking caused by differences in masker level at 4 and 32 kHz. The effect of SAM rate is similar to that in the previous analysis (see Sec. III A above). There were significant interactions between the effects of SAM rate and signal frequency $(F[7,48] = 4.58, \ p < 0.001)$ and of SAM rate and noise spectral density level $(F[14,48] = 6.12, \ p < 0.0005)$ on the masked threshold. There was no significant three-way interaction.

A mixed model ANOVA with the CR, i.e., the signal-to-noise ratio for detecting the 1000-ms tone in a constant amplitude masker, as the dependent variable and signal frequency and noise spectral density level as factors, revealed a significantly larger CR at 32 kHz than at 4 kHz $(F[1,6] = 75.3, \ p < 0.0005)$ as observed before (Sec. III A above), but no significant effect of noise spectral density level on the CR $(F[4,6] = 4.32, \text{n.s.})$.

As in the previous analysis, the amount of masking release (Fig. 6) was affected by signal frequency and SAM rate, and it was affected by masker spectral density level (mixed-model ANOVA, main effect signal frequency $F[1,48] = 38.9, \ p < 0.0005$, main effect SAM rate $F[7,48] = 49.4, \ p < 0.0005$, main effect of masker spectral density level $F[2,48] = 98.1, \ p < 0.0005$). As before, there was a significant interaction of signal frequency and SAM rate $(F[1,48] = 3.38, \ p < 0.01)$. Significant interactions were observed involving masker spectral density level and SAM rate $(F[14,48] = 4.52, \ p < 0.00005)$ and masker spectral density level and signal frequency $(F[2,48] = 15.5, \ p < 0.0005)$. At the signal frequency of 4 kHz, the masking release on average dropped by 5 dB when the masker spectral density level was reduced from 74 dB re 1 μPa^2/Hz to 54 dB re 1 μPa^2/Hz. At the signal frequency of 32 kHz, the masking release on average dropped by 11 dB when the masker spectral density level was reduced from 74 dB re 1 μPa^2/Hz to 54 dB re 1 μPa^2/Hz. Even at the

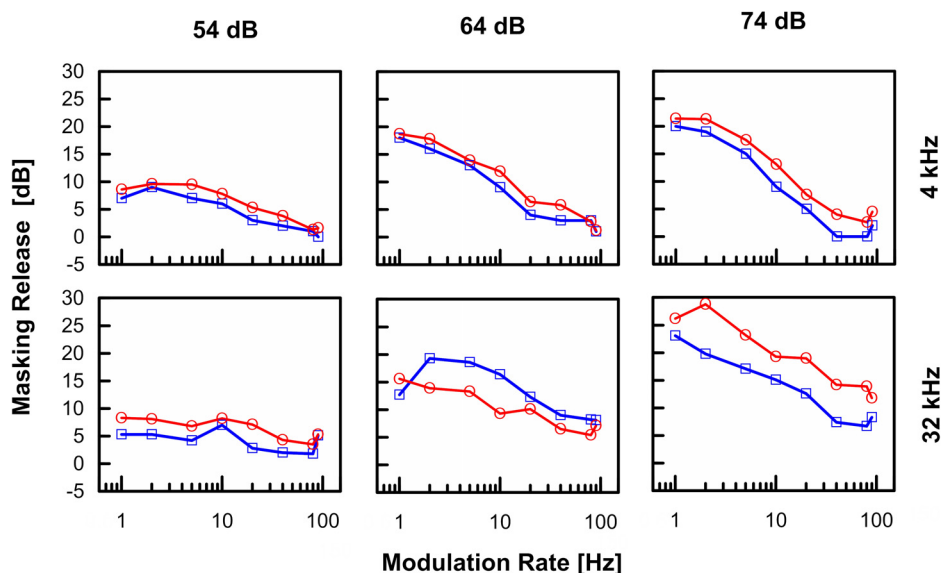FIG. 6. (Color online) Masking release for 4 kHz and 32 kHz tone signals embedded in a third-octave bandwidth masker in relation to the SAM rate of the masker. The masker spectral density level (dB re 1 μPa^2/Hz) indicated above the columns was an additional parameter. Signal duration was 1000 ms. The colors represent the two different seals.

J. Acoust. Soc. Am. **154** (1), July 2023

Kastelein *et al.* 89

lowest masker level of 54 dB re 1 μPa^2/Hz, the average masking release was significantly different from 0 dB as judged from the 95% confidence intervals. There was no significant three-way interaction.

IV. DISCUSSION

In the present study, we investigated in two harbor seals how sinusoidal level fluctuations of a masker with a bandwidth that is larger than the critical-ratio bandwidth can be exploited for improving tone-signal detection at 4 and 32 kHz. The results of the two seals were quite similar. A considerable masking release was observed in low-frequency SAM masker rates compared to constant level maskers with the same bandwidth. The various factors determining the masking release and the relevance of the masking release for evaluating the effect of temporally structured environmental noise on perception will be discussed in the following.

A. Thresholds in quiet and unmodulated noise maskers

The unmasked thresholds of the seals in this study are within a few dB of measures obtained from the same seals in previous studies (Kastelein et al., 2009a; Kastelein et al., 2009b; Kastelein et al., 2010) and are similar to the thresholds of other seal species (e.g., Reichmuth et al., 2013; Sills et al., 2014; Sills et al., 2015). The CR values are within a few dB of mean values obtained from harbor seals and from other seal species in the range of frequencies tested in the present study (Turnbull and Terhune, 1990; Terhune, 1991; Southall et al., 2000; Sills et al., 2014; Sills et al., 2015).

B. Masking release experienced in SAM noise

1. Effect of the SAM rate of the masking noise

In the present study, the masking release being calculated relative to the masked thresholds for the unmodulated (i.e., constant-amplitude) Gaussian noise masker, revealed considerable improvement in signal detection due to the modulation. The largest improvement of between 15 dB and 29 dB was observed for SAM frequencies of 1 or 2 Hz, i.e., low rates of modulation. At these SAM frequencies, the depth of modulation was close to 100% for both 4- and 32-kHz maskers (Table I). Above a SAM rate of 2 Hz, the masking release decreased with an increasing SAM rate. The observation of a more pronounced decrease at 4 kHz than at 32 kHz can likely be explained by the differences in the acoustics in the experimental setup. For 32-kHz maskers, the depth of modulation was still high (89%) at a SAM frequency of 20 Hz, and the depth of modulation decreased to only about 60% for the highest SAM rates. For 4-kHz maskers, a high depth of modulation (>90%) was observed for SAM rates of 5 Hz and below and the depth of modulation decreased to 20% for the highest SAM rates (for all data see Table I and the supplemental material[1]). Thus, the drop in masking release with SAM rates increasing

above 5 Hz for the 4-kHz masker and above 40 Hz for the 32 kHz masker may be partially explained by the decrease in the depth of modulation. Below these SAM rates, however, the drop in masking release with increasing SAM rate must be mainly due to limitations in the temporal processing of the masker envelope rather than limitations due to a change in the depth of modulation. The partial correlation analysis demonstrated that at 32 kHz the effects of SAM rate and acoustic depth of modulation were separable. The sizes of the masking release experienced by a harbor porpoise presented with identical 4-kHz signals and maskers (Kastelein et al., 2021) as were used in the present study, and in macaque monkeys (Macaca sylvanus) with SAM noise maskers tested with modulation rates ranging from 5 to 20 Hz (Dylla et al., 2013) were similar to those observed in the harbor seals in the present study.

The observation of a high masking release for SAM maskers with slowly fluctuating envelopes corresponds to the observations made in the context of CMR (e.g., Hall et al., 1984) and fluctuation strength (e.g., Fastl, 1982). CMR has been observed in a range of vertebrate species (e.g., Klump and Langemann, 1995; Branstetter and Finneran, 2008; Trickey et al., 2010; Fay, 2011; Branstetter et al., 2013; Vélez and Bee, 2013; Kastelein et al., 2021) including humans (e.g., Hall et al., 1984; Buus, 1985; Schooneveldt and Moore, 1989; Moore and Schooneveldt, 1990; see also a review by Verhey et al., 2003). Thus, the masking release observed must be mainly due to the temporal processing within an auditory frequency filter of the harbor seal.

Overall, the relation between SAM rate and MR resembles the perception of the strength of masker envelope fluctuations (Fastl, 1982; Zwicker and Fastl, 2007). The perceived fluctuation strength can be viewed as an indicator of how well the temporal variations in the envelope can be perceived by the auditory system. The fluctuation strength of broadband SAM noise in human subjects is high for SAM rates between 1 to 8 Hz and decreases below and above that frequency, i.e., shows a bandpass characteristic (Fastl, 1982). Above SAM rates of 8 Hz, perception of fluctuation strength decreases with increasing modulation rate. A model proposed by Zwicker and Fastl (2007) suggests that the representation of the temporal structure of the SAM masker in the pattern of excitation of the inner ear, especially of the low-amplitude dips, deteriorates with an increasing rate of modulation. The reduction in the depth of modulation of the response may have a similar effect as a reduction of the depth of modulation of an acoustic masker. A smaller depth of modulation will decrease the ability of exploiting the higher signal-to-masker ratios in the dips for improving signal detection within an auditory filter (dip-listening, see Buus, 1985). This may also explain the observed relation between masking release and SAM rate in the harbor seal.

2. Effect of test signal duration

In general, the CRs determined with third-octave bandwidth, constant amplitude maskers with a spectral density of

74 dB re 1 μPa^2/Hz were higher at 32 than at 4 kHz, and for the 4 kHz data, CRs were higher when the signal duration was shorter. The former may be explained by an increase in the auditory filter bandwidth from 4 to 32 kHz and by the increase in the masker bandwidth. The effect of duration on the CR with the lowest CR being observed for a 2000 ms signal suggests that the CR was not only determined by the peripheral filters in the inner ear but also by central processes with a much longer integration time. All three signal durations used in the present experiment exceeded the integration time of the seals determined for tonal signal detection in low levels of ambient noise (218 ms at 4 kHz and 14 ms at 32 kHz; Kastelein *et al.*, 2010). Also, in the SAM maskers, the masked threshold decreased with increasing signal duration. This observation suggests that the seals can benefit from having more opportunities to detect longer signals in the dips of the SAM masker (i.e., benefitting from "multiple looks"; Buus, 1985).

Signal duration did not affect the average masking release. Since the masking release was computed as the difference between the signal threshold in SAM maskers and the signal threshold in an unmodulated masker (CR), this indicates that the long-duration integration processes may be similar for SAM noise maskers and constant amplitude noise maskers. The interactions of the effects of duration and SAM rate and of duration and signal frequency on the masking release may be due to the differences observed in the change of acoustic depth of modulation with increasing SAM rate that also differed between 4 and 32 kHz.

3. Effect of the level of the masking noise

For unmodulated maskers, signal to-noise ratios at threshold were affected little by the masker level. This can be expected based on the common observation that the CR changes little with masker level if ambient noise does not contribute to masking (e.g., Hawkins and Stevens, 1950; Johnson, 1968). Even at the lowest masking noise SDL of 54 dB re 1 μPa^2/Hz, the masking noise levels were 21–22 dB above the SDL of the ambient noise level. Thus, it is unlikely that the ambient noise levels were interfering with the influence of the continuous masking noise. At the lowest masker level with an SDL of 54 dB re 1 μPa^2/Hz, a masked threshold in constant maskers of on average 79 and 87 dB SPL was observed for 4- and 32-kHz maskers, respectively. This is more than 20 dB higher than the average threshold without a masker (56 and 63 dB SPL at 4 and 32 kHz, respectively). This further indicates that the ambient noise has no substantial effect on the results regarding the CR.

The effects of the ambient noise will be stronger for the SAM noise maskers. However, for SAM noise maskers, average signal-to-noise ratios at the threshold were increased with decreasing masker level differing by 5.5 dB between a masker SDL of 54 and 74 dB re 1 μPa^2/Hz. This increase may reflect an increasing effect of the ambient background noise on the SAM masker. Especially in case of the lowest noise-masker SDL (54 dB re 1 μPa^2/Hz), the

ambient noise will fill in the dips in the SAM masker, effectively reducing the depth of modulation and thus reducing a gain in sensitivity by dip-listening (Buus, 1985). Filling in the dips reduces the envelope fluctuations of the masker, which can explain both the reduced variation of the masked threshold in relation to SAM rate and the observation that the lowest masked thresholds for the tone detection in the SAM noise masker become more similar to the CR.

The filling in of the dips by ambient noise will have an effect on the masking release in that it effectively reduces the depth of modulation and decreases on average with decreasing masker level. In addition, the effect of SAM rate on the masking release becomes smaller with decreasing masker level. At a masker SDL of 54 dB re 1 μPa^2/Hz and an ambient noise SDL of 32 or 33 dB re 1 μPa^2/Hz, the largest modulation depths would be 79% or 78%, respectively. The MD percentages calculated in Table I were measured at much higher masking noise levels than were presented to the seals. At the lowest masking noise level presented, the ambient noise level would severely limit the modulation depth and thus only a small masking release could be observed and the variation with a change in SAM rate was smaller than at higher levels. In experiments that were conducted on human subjects with headphones in a quiet background, Bacon *et al.* (1997) observed a reduction of the masking release with decreasing SAM masker level. This suggests that not only ambient noise but also neural noise (e.g., being due to the neurons' spontaneous activity) may contribute to the effect of AM masker levels on masking release. To ensure that a signal is masked, the masker level must be set sufficiently high relative to threshold.

Sills *et al.* (2017) measured masked threshold levels of two seal species to seismic airguns during and just after individual sound pulses. They found that "When noise amplitude varied significantly in time, the results suggested that detection was driven by higher signal-to-noise ratios within time windows shorter than the full signal duration." The thresholds measured in the short duration at the height of the pulse were higher than those measured just after (when the amplitude was lower). Thus, as the amplitude of the masking noise rose and fell, the detection thresholds followed suit over short time periods (Sills *et al.*, 2017). This supports our finding that seals are able to respond to short duration dips in the AM masker amplitude.

C. Ecological significance

The hearing abilities of harbor seals and related behaviors have evolved in the presence of sounds from geological, meteorological and biological sources. Anthropogenic underwater noise prevalence and levels have increased over the past century and this has resulted in numerous negative impacts on a wide variety of marine animals (Duarte *et al.*, 2021). Underwater anthropogenic noise can result in phocids avoiding ensonified areas, changing their behavioral patterns, undergoing physiological changes, and having their sound detection abilities reduced by noise masking (Southall

et al., 2019; Duarte *et al.*, 2021). Brief (impulsive) noise sources are unlikely to mask signals that are important to harbor seals, although the signal's duration may increase due to reverberations, which can increase the likelihood of masking (Sills *et al.*, 2017). Non-impulsive (i.e., continuous) noise will mask calls or other important signals when frequency and temporal overlap occurs and when any amplitude modulation does not result in significant dips in the noise level. Other factors being equal, if the anthropogenic noise source has no or only a low degree of amplitude variation, the masking effect can be considered to be related to the SDL and the CR. In addition, other masking reduction strategies such as spatial masking release, vocal repetition, adopting the Lombard effect, etc., may affect overall masking. Such possible effects are yet to be fully examined. It will be important to understand the ways in which man-made noises interfere with sound perception, including masking, by seals so that adverse impacts can be assessed and appropriate mitigation measures undertaken, when necessary.

Harbor seals have relatively good low-frequency hearing sensitivity and thus are likely to be disturbed by the low-frequency sounds produced by large ships (Kastelein *et al.*, 2009a; Kastelein *et al.*, 2009b). Ship-generated noise can result in auditory masking of sounds that are important to harbor seals. If a single ship is close, the propeller noise will be temporally structured producing temporally modulated masker envelopes with high depth of modulation in which temporal masking release becomes important. If ships are cruising at a distance or if multiple ships are passing by, the masking noise has smaller envelope variations and masking effects may better be evaluated based on the critical masking ratio. In addition to shipping noise, underwater acoustic communication systems used to convey data from remotely operated vehicles, underwater sensor networks, and various navigation and military operations will produce high-frequency sound (Pranitha and Anjaneyulu, 2020) that could mask high-frequency signals relevant for harbor seals. One consequence of this masking is that their acoustic communication space will be reduced when noise levels are higher (Clark *et al.*, 2009; Erbe *et al.*, 2016). Communication space calculations include the noise levels, such that as the noise levels increase, the communication space decreases. The potential for masking release due to the temporal structure of the maskers will also have to be considered in such calculations. There is a need for gathering more information on envelope fluctuations in environmental noise and their temporal coherence across frequencies before we can make better quantitative predictions of communication distances in noise. Studies by Branstetter and Finneran (2008), Kastelein *et al.* (2021), and the present study indicate that bottlenose dolphins, harbor porpoises, and harbor seals can exploit noise with temporally fluctuating amplitude for improving communication and detection distances.

ACKNOWLEDGMENTS

[1]See supplementary material at https://doi.org/10.1121/10.0019631 for graphs showing the full analysis for modulation frequencies between 1 Hz and 90 Hz.

Ainslie, M. A., Halvorsen, M. B., and Robinson, S. P. (**2022**). "A terminology standard for underwater acoustics and the benefits of international standardization," IEEE J. Oceanic Eng. **47**, 179–200.

Audacity (**2017**). "Audacity 2.2.0, The Audacity Team," http://www.audacityteam.org/ (Last viewed February 2017).

Bacon, S., Lee, J., Peterson, D. N., and Rainey, D. (**1997**). "Masking by modulated and unmodulated noise: Effect of bandwidth, modulation rate, signal frequency, and masker level," J. Acoust. Soc. Am. **101**, 1600–1610.

Bain, D., and Dahlheim, M. (**1994**). "Effects of masking noise on detection thresholds of killer whales," in *Marine Mammals and the Exxon Valdez*, edited by T. Loughlin (Academic Press, San Diego, CA), pp. 243–256.

Branstetter, B. K., and Finneran, J. J. (**2008**). "Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*)," J. Acoust. Soc. Am. **124**, 625–633.

Branstetter, B. K., and Sills, J. M. (**2022**). "Mechanisms of auditory masking in marine mammals," Anim. Cogn. **25**, 1029–1047.

Branstetter, B. K., Trickey, J. S., Black, A., Aihara, H., and Finneran, J. J. (**2013**). "Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise," J. Acoust. Soc. Am. **133**, 1811–1818.

Burns, J. J., Perrin, W. F., Würsig, B., and Thewissen, J. G. M. (**2002**). "Harbor seal and spotted seal," in *Encyclopedia of Marine Mammals*, edited by W. F. Perrin, B. Würsig, and J. G. M. Thewissen (Academic Press, San Diego, CA), pp. 552–560.

Buus, S. R. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Buus, S. R. (**1999**). "Temporal integration and multiple looks revisited: Weights as a function of time," J. Acoust. Soc. Am. **105**, 2466–2475.

Clark, C. W., Ellison, W. T., Southall, B. L., Hatch, L., Van Parijs, S. M., Frankel, A., and Ponirakis, D. (**2009**). "Acoustic masking in marine ecosystems: Intuitions, analysis, and implication," Mar. Ecol. Prog. Ser. **395**, 201–222.

Cooke, M. (**2006**). "A glimpsing model of speech perception in noise," J. Acoust. Soc. Am. **119**, 1562–1573.

Kastelein *et al.*

Cornsweet, T. N. (**1962**). "The staircase-method in psychophysics," Am. J. Psychol. **75**, 485–491.

Cunningham, K. A., Southall, B. L., and Reichmuth, C. (**2014**). "Auditory sensitivity of seals and sea lions in complex listening scenarios," J. Acoust. Soc. Am. **136**, 3410–3421.

Duarte, C. M., Chapuis, L., Collin, S. P., Costa, D. P., Devassy, P., Eguiluz, M. E., Erbe, C., Gordon, T. A. C., Halpern, B. S., Harding, H. R., Havlik, M. N., Meekan, M., Merchant, N. D., Miksis-Olds, J. L., Parsons, M., Predragovic, M., Radford, A. N., Radford, C. A., Simpson, S. D., Slabbekoorn, H., Staaterman, E., Van Opzeeland, I. C., Winderen, J., Zhang, X., and Juanes, F. (**2021**). "The soundscape of the Anthropocene ocean," Science **371**(6529), eaba4658.

Dylla, M., Hrnicek, A., Rice, C., and Ramachandran, R. (**2013**). "Detection of tones and their modification by noise in nonhuman primates," J. Assoc. Res. Otolaryngol. **14**, 547–560.

Erbe, C., Reichmuth, C., Cunningham, K., Lucke, K., and Dooling, R. (**2016**). "Communication masking in marine mammals: A review and research strategy," Mar. Pol. Bull. **103**, 15–38.

Fastl, H. (**1982**). "Fluctuation strength and temporal masking patterns of amplitude-modulated broadband noise," Hear. Res. **8**, 59–69.

Fay, R. R. (**2011**). "Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*," J. Acoust. Soc. Am. **129**, 3367–3372.

Finneran, J. J., and 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.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–66.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectro-temporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.

Hawkins, J. E., and Stevens, S. S. (**1950**). "The masking of pure tones and of speech by white noise," J. Acoust. Soc. Am. **22**, 6–13.

Johnson, C. S. (**1968**). "Masked tonal thresholds in the bottlenosed porpoise," J. Acoust. Soc. Am. **44**, 965–967.

Kastak, D., and Schusterman, R. J. (**1998**). "Low-frequency amphibious hearing in pinnipeds: Methods, measurements, noise, and ecology," J. Acoust. Soc. Am. **103**, 2216–2228.

Kastelein, R. A., Helder-Hoek, L., Cornelisse, S., Huijser, L. A. E., and Terhune, J. M. (**2019b**). "Temporary hearing threshold shift in harbor seals (*Phoca vitulina*) due to a one-sixth-octave noise band centered at 16 kHz," J. Acoust. Soc. Am. **146**, 3113–3122.

Kastelein, R. A., Helder-Hoek, L., and Gransier, R. (**2019c**). "Frequency of greatest temporary threshold shift in harbor seals (*Phoca vitulina*) depends on the fatiguing sound level," J. Acoust. Soc. Am. **145**, 1353–1362.

Kastelein, R. A., Helder-Hoek, L., Cornelisse, S. A., Defillet, L. N., Huijser, L. A. E., and Terhune, J. M. (**2020c**). "Temporary hearing threshold shift in harbor seals (*Phoca vitulina*) due to one-sixth-octave noise bands centered at 0.5, 1, and 2 kHz," J. Acoust. Soc. Am. **148**, 3873–3885.

Kastelein, R. A., Helder-Hoek, L., Cornelisse, S. A., Huijser, L. A. E., and Terhune, J. M. (**2020a**). "Temporary hearing threshold shift in harbor seals (*Phoca vitulina*) due to a one-sixth-octave noise band centered at 32 kHz," J. Acoust. Soc. Am. **147**, 1885–1896.

Kastelein, R. A., Helder-Hoek, L., Covi, J., Terhune, J. M., and Klump, G. (**2021**). "Masking release at 4 kHz in harbor porpoises (*Phocoena phocoena*) associated with sinusoidal amplitude-modulated masking noise," J. Acoust. Soc. Am. **150**, 1721–1732.

Kastelein, R. A., Gransier, R., Brouwers, M., and Helder-Hoek, L. (**2019a**). "Hearing thresholds of two harbor seals (*Phoca vitulina*) for helicopter dipping sonar signals (1.3-1.4 kHz)," Aquat. Mamm. **45**, 349–355.

Kastelein, R. A., Helder-Hoek, L., Gransier, R., Terhune, J. M., Jennings, N., and de Jong, C. A. F. (**2015**). "Hearing thresholds of harbor seals (*Phoca vitulina*) for playbacks of seal scarer signals, and effects of the signals on behavior," Hydrobiologia **756**, 75–88.

Kastelein, R. A., Helder-Hoek, L., and Terhune, J. M. (**2018a**). "Hearing thresholds, for underwater sounds of harbor seals (*Phoca vitulina*) at the water surface," J. Acoust. Soc. Am. **143**, 2554–2563.

Kastelein, R. A., Hoek, L., Gransier, R., de Jong, C. A. F., and Jennings, N. (**2013**). "Hearing thresholds of two harbor seals (*Phoca vitulina*) for playbacks of multiple pile driving strike sounds," J. Acoust. Soc. Am. **134**, 2307–2312.

Kastelein, R. A., Hoek, L., Kommeren, A., Covi, J., and Gransier, R. (**2018b**). "Effect of pile driving sounds on harbor seals (*Phoca vitulina*) hearing," J. Acoust. Soc. Am. **143**, 3583–3594.

Kastelein, R. A., Hoek, L., Wensveen, P. J., Terhune, J. M., and de Jong, C. A. F. (**2010**). "The effect of signal duration on the underwater hearing thresholds of two harbor seals (*Phoca vitulina*) for single tonal signals between 0.2 and 40 kHz," J. Acoust. Soc. Am. **127**, 1135–1145.

Kastelein, R. A., Parlog, C., Helder-Hoek, L., Cornelisse, S. A., Huijser, L. A. E., and Terhune, J. M. (**2020b**). "Temporary hearing threshold shift in harbor seals (*Phoca vitulina*) due to a one-sixth-octave noise band centered at 40 kHz," J. Acoust. Soc. Am. **147**, 1966–1976.

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

Kastelein, R. A., Wensveen, P., Hoek, L., and Terhune, J. M. (**2009a**). "Underwater hearing sensitivity of harbor seals (*Phoca vitulina*) for narrow noise bands between 0.2 and 80 kHz," J. Acoust. Soc. Am. **126**, 476–483.

Klump, G. M., and Langemann, U. (**1995**). "Comodulation masking release in a songbird," Hear. Res. **87**, 157–164.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Madsen, P. T. (**2005**). "Marine mammals and noise: Problems with root mean square sound pressure levels for transients," J. Acoust. Soc. Am. **117**, 3952–3957.

McDonald, M. A., Hildebrand, J. A., and Wiggins, S. M. (**2006**). "Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California," J. Acoust. Soc. Am. **120**, 711–718.

McDonald, M. A., Hildebrand, J. A., Wiggins, S. M., and Ross, D. (**2008**). "A 50 year comparison of ambient ocean noise near San Clemente Island: A bathymetrically complex coastal region off Southern California," J. Acoust. Soc. Am. **124**, 1985–1992.

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

Moore, B. C., and Schooneveldt, G. P. (**1990**). "Comodulation masking release as a function of bandwidth and time delay between on-frequency and flanking-band maskers," J. Acoust. Soc. Am. **88**, 725–731.

Murphy, C. T., Reichmuth, C., and Mann, D. (**2015**). "Vibrissal sensitivity in a harbor seal (*Phoca vitulina*)," J. Exp. Biol. **218**, 2463–2471.

Pranitha, B., and Anjaneyulu, L. (**2020**). "Analysis of underwater acoustic communication system using equalization technique for ISI reduction," Procedia Comput. Sci. **167**, 1128–1138.

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

Robinson, D. E., and Watson, C. S. (**1972**). "Psychophysical methods in modern psychoacoustics," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic Press, New York), Vol. 2, pp. 99–131.

Scharf, B. (**1970**). "Critical bands," in *Foundations of Modern Auditory Theory*, Vol. 1, edited by J. V. Tobias (Academic Press, New York), pp. 157–202.

Schooneveldt, G. P., and Moore, B. C. (**1989**). "Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration," J. Acoust. Soc. Am. **85**, 273–281.

Sills, J., Parnell, K., Ruscher, B., Lew, C., Kendall, T. L., and Reichmuth, C. (**2021**). "Underwater hearing and communication in the endangered Hawaiian monk seal *Neomonachus schauinslandi*," Endang. Species Res. **44**, 61–78.

Sills, J. M., Reichmuth, C., Southall, B. L., Whiting, A., and Goodwin, J. (**2020**). "Auditory biology of bearded seals (*Erignathus barbatus*)," Polar Biol. **43**, 1681–1691.

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

Sills, J. M., Southall, B. L., and Reichmuth, C. (**2015**). "Amphibious hearing in ringed seals (*Pusa hispida*): Underwater audiograms, aerial audiograms and critical ratio measurements," J. Exp. Biol. **218**, 2250–2259.

J. Acoust. Soc. Am. **154** (1), July 2023

Kastelein *et al.*   93

Sills, J. M., Southall, B. L., and Reichmuth, C. (**2017**). "The influence of temporally varying noise from seismic air guns on the detection of underwater sounds by seals," J. Acoust. Soc. Am. **141**, 996–1008.

Southall, B. L., Finneran, J. J., Reichmuth, C., Nachtigall, P. E., Ketten, D. R., Bowles, A. E., Ellison, W. T., Nowacek, D. P., and Tyack, P. L. (**2019**). "Marine mammal noise exposure criteria: Updated scientific recommendations for residual hearing effects," Aquat. Mamm. **45**, 125–232.

Southall, B. L., Nowacek, D. P., Senigaglia, A. E. B. V., Bejder, L., and Tyack, P. L. (**2021**). "Marine mammal noise exposure criteria: Assessing the severity of marine mammal behavioral response to human noise," Aquat. Mamm. **47**, 421–464.

Southall, B. L., Schusterman, R. J., and Kastak, D. (**2000**). "Masking in three pinnipeds: Underwater, low-frequency critical ratios," J. Acoust. Soc. Am. **108**, 1322–1326.

Terhune, J. M. (**1988**). "Detection thresholds of a harbor seal to repeated underwater high-frequency, short duration sinusoidal pulses," Can. J. Zool. **66**, 1578–1582.

Terhune, J. M. (**1991**). "Masked and unmasked pure tone detection thresholds of a harbour seal listening in air," Can. J. Zool. **69**, 2059–2066.

Terhune, J. M., and Ronald, K. (**1971**). "The harp seal, *Pagophilus groenlandicus (Erxleben*, 1777). X. The air audiogram," Can. J. Zool. **49**, 385–390.

Terhune, J. M., and Ronald, K. (**1975**). "Masked hearing thresholds of ringed seals," J. Acoust. Soc. Am. **58**, 515–516.

Trickey, J. S., Branstetter, B. K., and Finneran, J. J. (**2010**). "Auditory masking of a 10 kHz tone with environmental, comodulated, and Gaussian noise in a bottlenose dolphin (*Tursiops truncatus*)," J. Acoust. Soc. Am. **128**, 3799–3804.

Turnbull, S. D., and Terhune, J. M. (**1990**). "White noise and pure tone masking of pure tone thresholds of a harbour seal listening in air and underwater," Can. J. Zool. **68**, 2090–2097.

Turnbull, S. D., and Terhune, J. M. (**1993**). "Repetition enhances hearing detection thresholds in a harbour seal (*Phoca vitulina*)," Can. J. Zool. **71**, 926–932.

Urick, R. J. (**1983**). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).

Vélez, A., and Bee, M. A. (**2013**). "Signal recognition by green treefrogs (*Hyla cinerea*) and cope's gray treefrogs (*Hyla chrysoscelis*) in naturally fluctuating noise," J. Comp. Psychol. **127**, 166–178.

Verhey, J. L., Pressnitzer, D., and Winter, I. M. (**2003**). "The psychophysics and physiology of comodulation masking release," Exper. Brain Res. **153**, 405–417.

Wolski, L. F., Anderson, R. C., Bowles, A. E., and Yochem, P. K. (**2003**). "Measuring hearing in the harbor seal (*Phoca vitulina*): Comparison of behavioral and auditory brainstem response techniques," J. Acoust. Soc. Am. **113**, 629–637.

Zwicker, E., and Fastl, H. (**2007**). *Psychoacoustics: Facts and Models* (Springer, Heidelberg, Germany).

17 January 2024 10:58:28

94   J. Acoust. Soc. Am. **154** (1), July 2023

Kastelein *et al.*