# UC Santa Cruz
## UC Santa Cruz Previously Published Works

**Title**
Mechanisms of auditory masking in marine mammals.

**Permalink**
https://escholarship.org/uc/item/1dm5f2jf

**Journal**
Animal Cognition, 25(5)

**Authors**
Branstetter, Brian
Sills, Jillian

**Publication Date**
2022-10-01

**DOI**
10.1007/s10071-022-01671-z

Peer reviewed

**REVIEW**

# Mechanisms of auditory masking in marine mammals

Brian K. Branstetter[1] · Jillian M. Sills[2]

## Abstract

Anthropogenic noise is an increasing threat to marine mammals that rely on sound for communication, navigation, detecting prey and predators, and finding mates. Auditory masking is one consequence of anthropogenic noise, the study of which is approached from multiple disciplines including field investigations of animal behavior, noise characterization from in-situ recordings, computational modeling of communication space, and hearing experiments conducted in the laboratory. This paper focuses on laboratory hearing experiments applying psychophysical methods, with an emphasis on the mechanisms that govern auditory masking. Topics include tone detection in simple, complex, and natural noise; mechanisms for comodulation masking release and other forms of release from masking; the role of temporal resolution in auditory masking; and energetic vs informational masking.

**Keywords** Masking in marine mammals · Marine mammals and noise · Auditory masking · Informational masking

## Introduction

Auditory masking in marine mammals has garnered attention in recent years due to an increased awareness of the negative effects of anthropogenic noise on hearing. When one sound interferes with a listener's ability to detect, discriminate, or recognize another sound, auditory masking occurs. The marine environment has always been noisy, because sound travels exceptionally well in ocean environments. Anything that produces sound is a potential noise source (for a review, see Richardson et al. 1995). One type of noise source is non-biological, which includes wind, rain, ice movement, tides, and seismic events. Biological noise comprises any sounds that living organisms produce, such as snapping shrimp snaps, animal vocalizations, and sounds generated by physical interactions with the environment (e.g., fish chewing coral and whales slapping the water's surface). Non-biological and biological sounds have been part of the ocean's soundscape for millions of years

and marine mammals have evolved to be acoustic specialists in this environment. However, there has been a dramatic increase in anthropogenic noise since the industrial revolution (see Richardson et al. 1995; Southall et al. 2019). Noise sources include transportation (e.g., shipping, aircraft, icebreakers), construction (e.g., dredging, pile driving), geophysical surveys (e.g., airgun arrays), military (e.g., sonars, explosives), and ocean science surveys (e.g., seismology, acoustic tomography) (Richardson et al. 1995). Negative impacts from anthropogenic noise can include physical discomfort, pain, or death (e.g., Parsons 2017); permanent and temporary thresholds shifts in hearing sensitivity (e.g., Finneran 2015); increased stress (e.g., Wright et al. 2007; Houser et al. 2020); changes in behavior that may affect individual fitness (e.g., Southall et al. 2009); and auditory masking (for review, see Erbe et al. 2016).

Auditory masking can be defined as "the process by which the threshold of hearing for one sound is raised by the presence of another (masking) sound; and the amount by which the threshold of hearing for one sound is raised by the presence of another (masking) sound, expressed in dB" (American National Standard Institute (ANSI) 1995). The study of auditory masking in marine mammals can be divided into three primary categories, the first being behavioral response studies evaluating whether animals change their behavior to *presumably* mitigate auditory masking (Holt et al. 2009). Anti-masking strategies include, but are not limited to, the

✉ Brian K. Branstetter
brian.branstetter@nmmf.org

[1]  National Marine Mammal Foundation, 2240 Shelter Island Drive, #204, San Diego, CA 92106, USA

[2]  Institute of Marine Sciences, Long Marine Laboratory, University of California Santa Cruz, Santa Cruz, CA 95060, USA

Lombard effect (e.g., Holt et al. 2009; Scheifele et al. 2005) and animals re-locating to quieter environments (e.g., Frankel and Clark 2000; Southall 2005) or increasing the number of elements per call to improve detectability (e.g., Turnbull and Terhune 1993; Serrano and Terhune 2001). The second category is based on *in-situ* noise measurements, which typically inform quantitative modeling efforts. These models predict how communication space is reduced relative to a baseline level, often factoring in signal detection capabilities (e.g., Clark et al. 2009). Models may include hearing-related parameters, such as frequency sensitivity, directivity index, and critical ratios, all of which are derived from studying the hearing of animals. The last category, which is the topic of this paper, is the direct study of hearing in noise. This paper will describe the auditory mechanisms that govern masking patterns in marine mammals, reviewing relevant example studies with a focus on psychoacoustical experiments performed in the laboratory.

Psychoacoustics is the study of the relationship between sound and an animal's mental representation of that sound (Fechner 1860; Fastl and Zwicker 2007). Since mental representations cannot be directly observed, inferences must be made by observing animal behavior. The standard equation describing this relationship is

$$\Omega = f(S), \tag{1}$$

where $\Omega$ is the measured behavior, $S$ is the physical attribute of sound, and $f(S)$ represents the functional relationship between behavior and sound (Yost and Fay 2012). In auditory masking, at least two sounds are present: a signal, which is a sound of interest, and noise, which is a sound interfering with the detection, discrimination, or recognition of the signal. In studies of marine mammal auditory masking, the amplitude of sound is typically represented in decibel (dB) units related to the sound pressure level (SPL) of the stimulus waveform; reference pressures are 1 µPa in water and 20 µPa in air. Common metrics to describe the levels of sound include root-mean-square ($dB_{RMS}$), peak pressure ($dB_P$), peak-to-peak pressure ($dB_{p-p}$), pressure spectral density (dB re $\mu Pa^2$ / Hz), and 1/3 octave (OTO) band level (for a review, see Au and Hastings 2008). The majority of studies described below measure masked detection thresholds using a go/no-go signal detection task (Stebbins 1970). Animals are presented with a noise background and trained to produce a response (e.g., a paddle press or vocalization) when they hear a signal in the presence of that noise, or to forgo a response if they do not detect a signal. The level of the signal is manipulated using a variety of methods (e.g., adaptive staircase procedure, method of constant stimuli; Stebbins 1970; Levitt 1971) to measure the amplitude at which the animal can detect the signal at a prescribed percent correct (e.g., 50% correct).
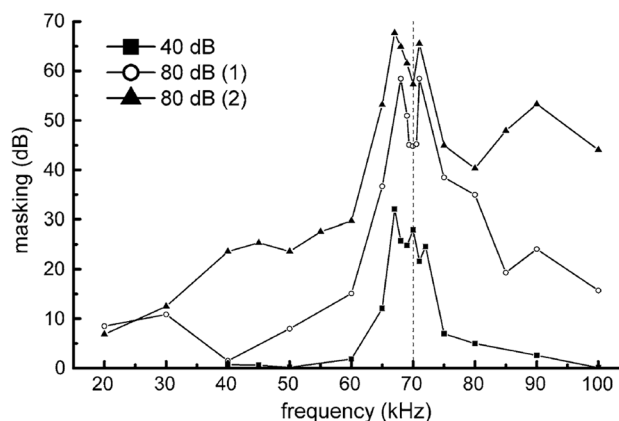


**Fig. 1** Two-tone masking experiment with a bottlenose dolphin, with masking (dB) shown as a function of frequency (kHz). The vertical dotted line represents the frequency of the masking tone at 70 kHz. Data points represent the dB level above absolute threshold for two levels of the masker: 40 dB and 80 dB re 1 µPa. The 1 and 2 represent two dolphin subjects (adapted from Johnson 1971).

## Detection of tones with simple masking noise

### Tone on tone masking

One of the earliest masking experiments to be conducted with a bottlenose dolphin (*Tursiops truncatus*) measured detection thresholds for a pure tone masked by another pure tone (Johnson 1971). In this study, a 70 kHz tone at two different SPLs (40 and 80 dB re 1 µPa) was used as a masking tone. Figure 1 plots how much masking occurred (i.e., the dB level above absolute threshold required to detect the signal at each frequency). Absolute threshold can be defined as the minimum level at which the signal can be detected half of the time without the influence of background noise. The observed masking patterns were consistent with human studies, where higher amplitude tones produced more masking and lower frequencies were better at masking higher frequencies.

Lower frequency tones are better at masking higher frequency tones due to basilar membrane mechanics. Hair cells sensitive to higher frequencies are positioned toward the basal end of the basilar membrane, while lower frequencies with longer wavelengths excite hair cells toward the apical end. Thus, a traveling wave from a low-frequency tone will traverse through and displace the high-frequency portion of the basilar membrane but not vice versa, resulting in the asymmetrical masking pattern.

The largest amount of masking in this experiment occurred when the signal and the masker were similar in frequency. Dips in the masking level near 70 kHz occurred when the signal and noise tones were almost identical in

frequency, as a result of the perception of "beats" or amplitude modulation (AM) due to the interaction of the two tones. The AM rate ($M_f$) in Hz is equal to

$$M_f = |f_1 - f_2|, \tag{2}$$

where $f_1$ and $f_2$ are the frequencies (in Hz) of the signal and noise tones, respectively. Mammalian auditory systems are more sensitive to low-frequency AM rates (Dolphin et al. 1995; Finneran et al. 2007; Viemeister, 1979), so as $M_f$ increases, the perception of beats diminishes and can no longer be used as a cue.

## Tone detection in white noise

Pure-tone signals and white-noise maskers are easy to generate and measure due to their relatively steady state, and are therefore commonly used in psychoacoustic experiments. White noise is a special type of noise where the pressure spectral density is equal or "flat" at each frequency. To digitally generate white noise, the instantaneous amplitude is randomly sampled from a normal or "Gaussian" distribution; hence, the term Gaussian noise is often used interchangeably with white noise.

In a seminal study with human listeners, Fletcher (1940) discovered that the level of a tonal signal at threshold ($S_{th}$) was proportional to the bandwidth of Gaussian noise ($\Delta f$) centered on the tone. As the bandwidth increased, detection thresholds increased proportionally, but only up to a *critical bandwidth* ($\Delta f_{CB}$). Noise frequencies beyond the critical bandwidth had no effect on detection thresholds. Fletcher envisioned a hypothetical band-pass filter or *critical band* centered on the signal, where only noise within the critical band contributed to the masking of the signal (Fig. 2A). The relationship between the signal at threshold ($S_{th}$) and the bandwidth of noise within the critical band ($\Delta f_{CB}$) can be formalized by

$$S_{th} = K \cdot N \cdot \Delta f_{CB}, \tag{3}$$

where $N$ is the spectral density (in $\mu Pa^2$ / Hz) of the noise and $K$ is a constant. If $K$ is assumed to be equal to one, Eq. 3 can be simplified to

$$\Delta f_{CR} = \frac{S_{th}}{N}, \tag{4}$$

where $\Delta f_{CR}$ is the critical ratio. Rather than performing the relatively time-consuming band-widening procedure necessary for critical band measurements, the auditory filter bandwidth can be estimated by simply measuring the detection threshold of a tonal signal in broadband noise to calculate the critical ratio (Fig. 2B). The term *broadband noise* is, within this paper, operationally defined as a bandwidth greater than the auditory filter bandwidth.
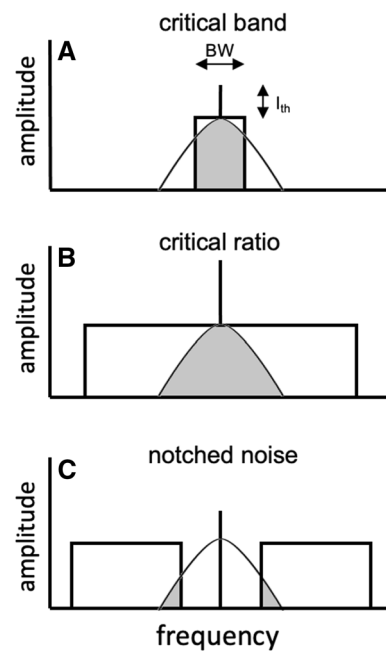


**Fig. 2** Relationship between (**A**) the critical band, (**B**) the critical ratio, and (**C**) notched-noise masking paradigms. In each panel, the height of the rectangle represents the spectral density level of the masking noise (dB re (1 $\mu Pa$)$^2$ / Hz in water) and the width represents the bandwidth (BW) in Hz. The vertical line ($I_{th}$) represents amplitude of the tonal signal at masked threshold (dB re 1 $\mu Pa$ in water). The curve represents a hypothetical auditory filter centered on the signal and the gray portion under the curve represents the amount of noise that contributes to the masking of the signal

For humans with a relatively narrow range of hearing (approximately 20 Hz to 20 kHz; Yost and Fay 2012), critical ratios can be estimated using a noise bandwidth that encompasses the full range of hearing. For marine mammals, some of which have an extremely broad hearing range (e.g., < 75 Hz to > 150 kHz for the bottlenose dolphin; Johnson 1967), critical ratios are often measured with noise bands that are one-third octave wide or greater (e.g., Au and Moore 1990; Branstetter et al. 2017, 2021). If the level of the signal ($S_{th}$) is expressed in terms of SPL and the level of the noise ($N$) is expressed as the pressure spectral density, the critical ratio (CR) can be calculated by

$$CR = S_{th} - N. \tag{5}$$

Critical ratios have been measured for five odontocete species: *Delphinapterus leucas* (Johnson et al. 1989), *Orcinus orca* (Branstetter et al. 2021), *Phocoena phocoena* (Kastelein and Wensveen, 2008; Kastelein et al. 2009), *Pseudorca crassidens* (Thomas et al. 1990), and *Tursiops truncatus* (Au and Moore 1990; Branstetter et al. 2017; Johnson 1968a; Lemonds et al. 2011). Auditory masking
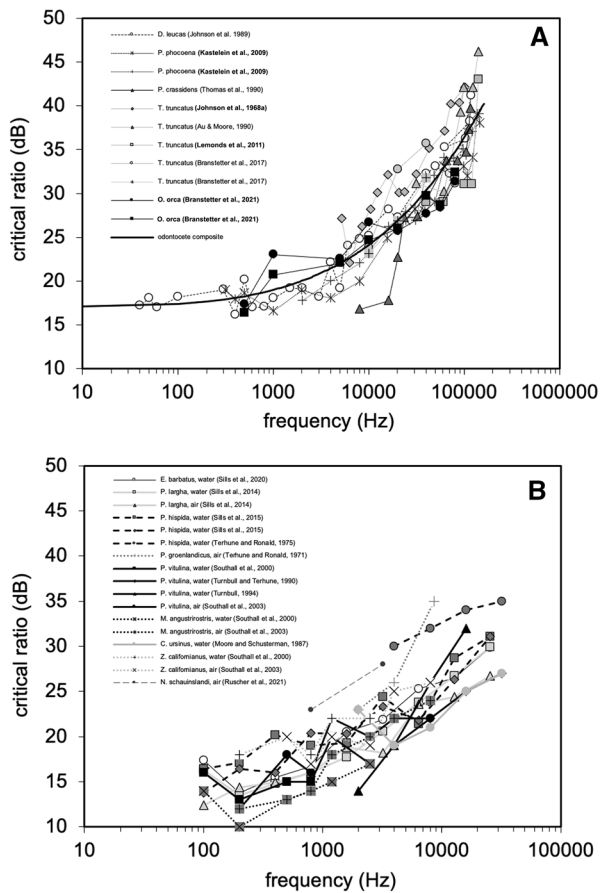
**Fig. 3** Critical ratios for odontocetes (**A**) and pinnipeds (**B**) as a function of frequency. The curve in panel **A** represents a model fit to the cumulative data. For pinniped critical ratios in panel **B**, the legend indicates the species and the sound medium (air or water) in which the critical ratios were measured. (Panel **A** was adapted from Branstetter et al. 2021; Finneran and Branstetter 2013)

(at least energetic masking) is thought to occur at the mechano-transduction level of the cochlea (Recio-Spinoso and Cooper 2013). Clear differences in basilar membrane size and morphology exist within odontocetes (Ketten 1992a); however, despite large differences in ear morphology (Ketten 1992b), functional head size (Heffner and Heffner 2008), and frequency sensitivity (see NOAA Fisheries 2018), critical ratios for toothed whales species are remarkably similar (Fig. 3A). Among pinnipeds, critical ratio measurements are available for nine species: *Callorhinus ursinus* (Moore and Schusterman, 1987), *Erignathus barbatus* (Sills et al. 2020), *Mirounga angustirostris* (Southall et al. 2000, 2003), *Neomonachus schauinslandi* (Ruscher et al. 2021), *Pagophilus groenlandicus* (Terhune and Ronald 1971), *Phoca largha* (Sills et al. 2014), *Phoca vitulina* (Renouf 1980; Southall et al. 2000, 2003; Terhune 1991; Turnbull 1994; Turnbull and Terhune 1990, 1993), *Pusa hispida* (Sills et al. 2015; Terhune and Ronald 1975), and *Zalophus*

*californianus* (Southall et al. 2000; 2003). Critical ratios measured for seals and sea lions (Fig. 3B) are consistently low across a wide range of frequencies relative to many terrestrial mammals (Fay 1988). This suggests that the auditory systems of these marine mammals possess a refined ability to detect signals within background noise.

A feature of mammalian auditory filters is that bandwidth typically increases as a function of the center frequency of the filter, represented by

$$Q = f_o / \triangle f, \tag{6}$$

where $f_o$ is the center frequency of the filter, $\Delta f$ is the filter bandwidth, and $Q$ is a quality factor. As a result of this relationship, the bandwidth of noise that contributes to the masking of a signal and the amount by which the signal must exceed that noise to be detected are both greater at higher frequencies. This is reflected by a general increase in both critical bands and critical ratios as a function of frequency (Fig. 3).

Fletcher's auditory filter model (Fletcher 1940) has developed into what is now referred to as the power spectrum model (PSM) of masking (Moore 1993). The PSM makes the following assumptions:

1. The auditory system can be modeled as a bank of continuously overlapping band-pass filters.
2. Only the spectral components within an auditory filter contribute to the masking of a signal centered within that filter.
3. Signal detection occurs by monitoring an energy detector at the output of the filter. A signal plus noise interval will result in an increase in filter output compared to a noise alone interval.
4. Signal threshold is proportional to the spectral density of noise within a filter centered on the signal. Noise is represented by its long-term spectra.

The model can be formalized by

$$P_s = K \int_0^\infty N(f)W(f)df, \tag{7}$$

where $P_S$ is the power of the signal, $N(f)$ is the noise power spectrum, and $W(f)$ is the auditory filter function, or the shape of the auditory filter. Auditory filter shapes have been measured behaviorally for two odontocete species (Finneran et al. 2002; Lemonds 1999; Lemonds et al. 2011) using a notched-noise masking paradigm (Fig. 2C). This involves measuring the threshold of a sinusoidal signal as a function of the width of a spectral notch in masking noise around the signal frequency. The masking data can then be fit to a two-parameter rounded exponential function (Patterson 1976; Patterson et al. 1982) or *roex* function

**Fig. 4** Auditory filter banks for (**A**) *Tursiops truncatus* (**B**), *Delphinapterus Leucas*, and (**C**) *Phocoena phocoena*. (Adapted from Finneran and Branstetter 2013; Au et al. 2009)
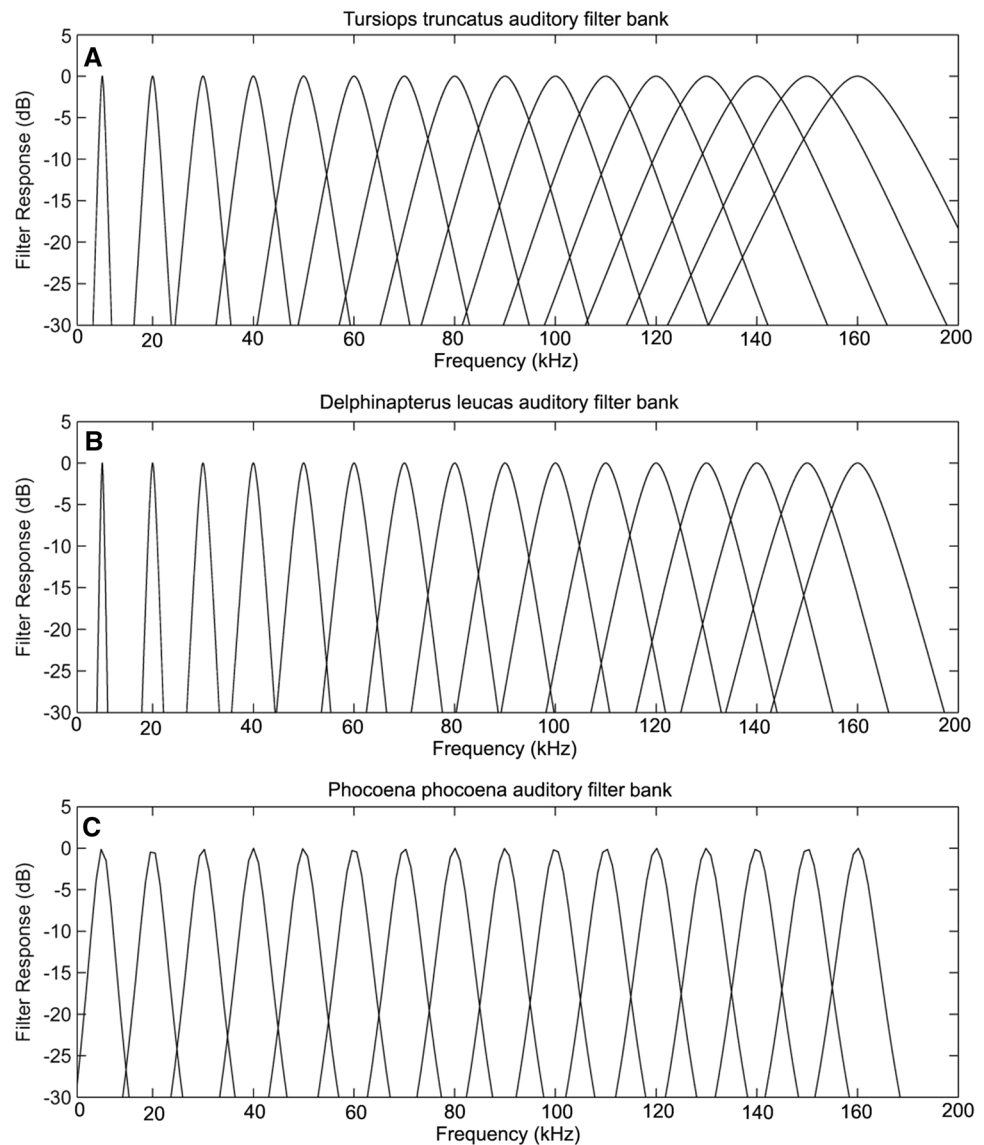


$$W(g) = (1 - r)(1 + pg)e^{(-pg)} + r, \qquad (8)$$

where $g$ is the normalized frequency deviation from the signal frequency

$$g = \frac{|f - f_o|}{f_o}. \qquad (9)$$

Here, $f$ is the cut-off frequency of the notched noise in Hz and $f_o$ is the center frequency of the notch, and $p$ and $r$ are fitting parameters. The roex filter is defined in the frequency domain, but lacks impulse response timing and phase characteristics found in physiological studies (Lyon et al. 2010). Conversely, the gammatone filter is virtually identical to the roex filter in the frequency domain, but is defined in the time-domain; it is often used for modeling purposes (Branstetter et al. 2020; Lyon et al. 2010; Roitblat et al. 1996).

Figure 4 displays gammatone auditory filter banks derived from roex filters (Branstetter et al. 2007; Roitblat et al. 1996) for a bottlenose dolphin (Lemonds 1999), a beluga whale (*Delphinapterus leucas*; Finneran et al. 2002), and a harbor porpoise (*Phocoena phocoena*; Popov et al. 2006). The gammatone filter bank models can be used in computational simulations of the hearing system to help describe the discrimination and classification capabilities of odontocetes (Branstetter et al. 2007, 2020; Roitblat 1998; Au et al. 2009), but have yet to be used in simulations of auditory masking.

When predicting auditory masking for mitigation purposes, critical ratios have become a standard metric used (e.g., Clark et al. 2009; Erbe et al. 2016)—along with the PSM of masking—due to their relative ease of data collection and model simplicity. For example, the odontocete critical ratio at 10 kHz is about 24 dB (Fig. 3A). If broadband ocean noise was recorded with a spectral density level of
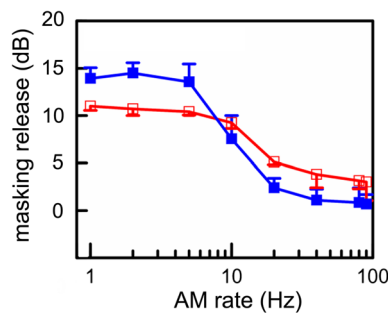
**Fig. 5** Masking release (dB) as a function of AM rate with SAM maskers and a 4 kHz tonal signal. The two colors represent two different harbor porpoise participants (Adapted from Kastelein et al. 2021)

100 dB re $(1 \ \mu Pa)^2$ /Hz @ 10 kHz, a received 10 kHz signal would need to be at least 124 dB re 1 µPa to be detected by most odontocetes. Although the critical ratio and PSM are simple and useful, the spectral–temporal complexity of many biological signals and ocean noise sources can lead to both masking release (MR) and elevated levels of masking, which will be discussed in subsequent sections.

## Signal detection with complex sounds

### Amplitude modulation, temporal resolution, and multiple looks

Many sounds in nature will have a spectral–temporal pattern more complex than Gaussian noise (Nelken et al. 1999). Thus, masking patterns derived from Gaussian noise may be considered a special case. One of the most common features of natural ocean noise is amplitude modulation caused by a reverberant environment, or the animals themselves. Kastelein et al. (2021) measured the harbor porpoise's ability to detect a 4 kHz tone embedded in sinusoidal amplitude modulated (SAM) noise. A release from masking up to 14 dB relative to Gaussian noise resulted for AM rates between 1 and 5 Hz. From 5 to 20 Hz, MR gradually decreased to levels similar to masking with Gaussian noise (Fig. 5). The results are consistent with a within-valley or a "dip" listening model (Buus 1985) where listening occurs in the troughs of the amplitude modulated noise where the signal-to-noise ratio is highest. As the modulation rate increases, the temporal resolution of the animal's hearing system will begin to "smear" the temporal envelope of the noise, perceptually reducing the troughs in the noise, thus increasing tone detection thresholds.

In humans, temporal resolution of SAM noise can be described by the temporal modulation transfer function (TMTF), which has a low-pass shape. The ability to discriminate between Gaussian and SAM noise decreases

rapidly above 100 Hz (Viemeister 1979). With higher AM rates and lower AM depths, SAM noise and Gaussian noise become indistinguishable. In odontocetes, temporal resolution is complex, and there appears to be at least two distinct mechanisms: one for transient sonar signals (Johnson et al. 1988; Moore et al. 1984) and another for longer duration signals (Johnson 1968b). For broadband transient signals similar to biosonar clicks, temporal resolution has been measured with an integration time of approximately 264 µs (Johnson et al. 1988; Moore et al. 1984). However, for tonal signals, integration times are frequency dependent and on the order of tens to hundreds of milliseconds. For example, the harbor porpoise has an integration time of approximately 277 ms at 4 kHz (Kastelein et al. 2021). If a 277 ms integration window were applied to amplitude modulated noise, AM rates above 4 Hz (4 Hz period = 250 ms) would likely be smeared, suggesting that different mechanisms govern the integration of tonal signals and the resolution of AM noise. Kastelein et al. also demonstrated that signals with longer durations (i.e., 1000 and 2000 ms) were easier to detect than a 500 ms signal, despite the fact that all three exceed the 277 ms integration time. The longer duration signals provide more opportunities for detecting the signal, which is consistent with a statistical "multiple looks" model (Viemeister and Wakefield 1991).

Currently, there is no consistent or single model of temporal integration for odontocetes. The idea that the odontocete auditory system analyses sound on different timescales (i.e., one for sonar-like signals and one for long-duration signals) is not without precedence. Evidence suggest that songbirds may have two distinct temporal integration times, one for short-duration signals between 10 and 30 ms, and another for longer duration signals between 500 and 700 ms (Narayan et al. 2006). Human speech may even be processed on a dual scale (Teng et al. 2016). How temporal integration affects auditory masking in marine mammals is a topic in need of investigation and will be revisited below in the section on comodulation masking release.

Sills et al. (2017) measured detection thresholds for a spotted seal (*Phoca largha*) and a ringed seal (*Pusa hispida*) listening for tonal signals in seismic airgun noise. Seismic airgun noise is an impulsive sound produced by a rapid release of compressed air (Vaage et al. 1983). Airgun arrays typically have a 10 s duty cycle, with an interval between pulses during which noise may or may not return to background levels (e.g., Guan et al. 2015). Sills et al. applied a time-window analysis, where windows of varying duration were used to measure the signal-to-noise ratio. The results indicate that the animals were likely using a "dip-listening" strategy. Detection thresholds were well predicted by critical ratios applied to short-duration temporal windows, particularly when noise amplitude fluctuated substantially over the duration of the signal. When noise levels were less variable,

**Fig. 6** Example of time-window analysis. Noise levels were calculated over different window durations between 50 and 500 ms. The levels were used to predict auditory masking when a 500 ms signal occurred either in the onset, intermediate, or terminal interval of the seismic airgun noise. (Adapted from Sills et al. 2017)
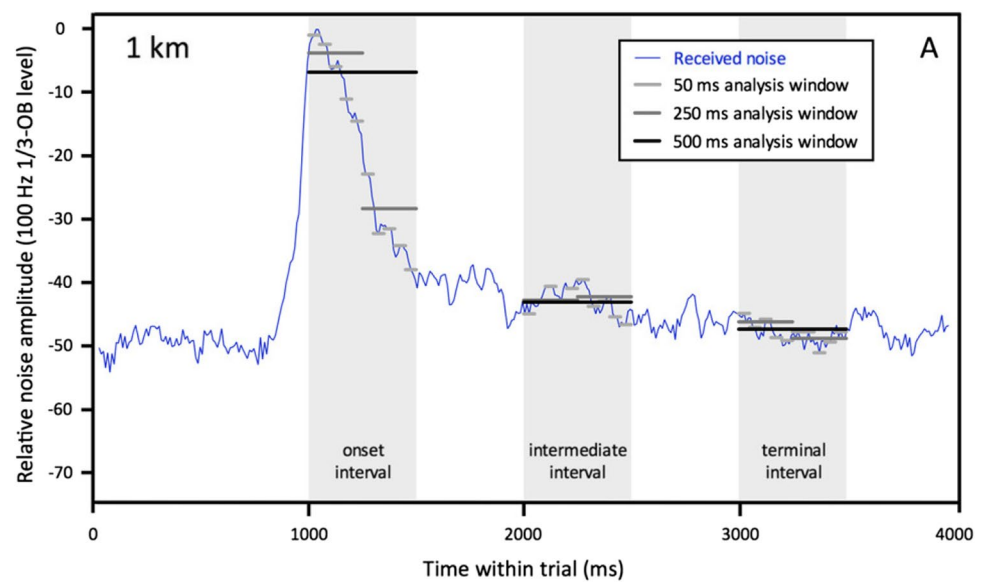


as in the 'intermediate' and 'terminal' intervals farther from pulse onset, analyses across longer temporal windows were also good predictors of masking (Fig. 6).

## Comodulation masking release

Many natural sounds are both broadband and amplitude modulated (Nelken et al. 1999). When the amplitude modulation is coherent across frequency regions (i.e., multiple auditory filters), the noise is called comodulated and can result in comodulation masking release (CMR) (Branstetter and Finneran 2008; Buschermohle et al. 2007; Hall and Grose 1988). Animal auditory systems appear to use comodulation to group and separate sounds in complex auditory scenes (Klump and Nieder 2001; Nelken et al. 1999). In a band-widening experiment, Branstetter and Finneran (2008) compared detection thresholds of a 10 kHz tone masked by Gaussian noise (G) and by comodulated noise (CM), which was produced by multiplying Gaussian noise by low-pass noise. When the bandwidths of G and CM noise were narrower than the width of an auditory filter, thresholds increased as a function of bandwidth in a manner consistent with the PSM. However, when bandwidth exceeded the width of an auditory filter (i.e., it was processed by more than one auditory filter), thresholds for tones within CM noise decreased as a function of noise bandwidth. The effect was pronounced, with a noise bandwidth of 8 kHz resulting in substantial release from masking relative to Gaussian noise of the same bandwidth and pressure spectral density (Fig. 7). In other words, more total noise energy resulted in less masking, a result that is not intuitive and is at odds with the PSM. The breakpoints for both G and CM noise at a bandwidth of 1 kHz (Fig. 7), which is the bandwidth of the auditory filter centered on the 10 kHz tone, suggest that
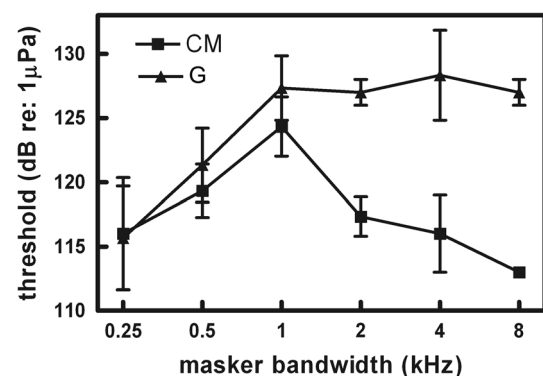


**Fig. 7** Comparison between masking patterns from comodulated (CM) noise and Gaussian (G) noise for one bottlenose dolphin. Masked thresholds at 10 kHz are shown (with standard deviation) as a function of masker bandwidth. (Adapted from Branstetter and Finneran 2008)

a processing transition occurs. For Gaussian noise, spectral frequencies beyond the auditory filter no longer contribute to masking of the signal (i.e., the PSM holds). However, CM noise beyond the auditory filter appears to enhance the detectability of the signal.

To test the effects of both amplitude modulation and the degree of coherence across frequency regions, an additional experiment was conducted (Branstetter et al. 2013a). Masking noise was divided into a signal band, which was the width of an auditory filter and centered on a 10 kHz signal, and two flanking bands (see Fig. 8). The signal band was then progressively delayed relative to the flanking bands, thus decreasing the level of across-channel coherence. This procedure was repeated for both G and CM noise. For the G noise, the delay had no effect on detection thresholds, since Gaussian noise is not coherent. However, for CM noise,
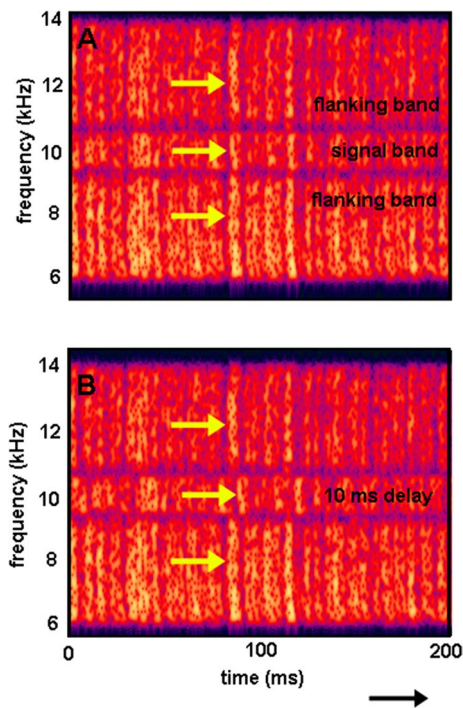
**Fig. 8** Spectrograms of masking noise with a signal band and two flanking bands. The signal band in Panel B is delayed by 10 ms, which decorrelates the envelopes. (Adapted from Branstetter et al. 2013a)
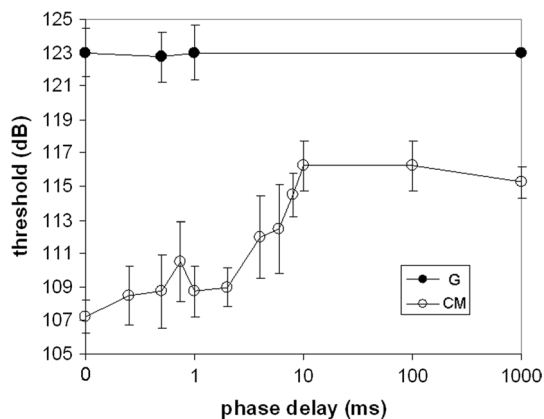


**Fig. 9** Detection thresholds in Gaussian (G) and comodulated (CM) noise as a function of the phase delay of the signal band, as shown in Fig. 8. The phase delay had no effect on thresholds measured in the presence of G noise. However, for CM noise, the phase delay accounted for approximately 7 dB of masking release and AM depth accounted for approximately 6 dB of masking release. Error bars represent standard deviations. (Adapted from Branstetter et al. 2013a)

the delay disrupted the CMR effect, resulting in increasing thresholds for delays between 2 and 10 ms (Fig. 9). The increase in thresholds as a function of delay was approximately 0.85 dB/ms delay (Branstetter et al. 2013a). Although CMR was disrupted by decorrelating the envelopes, the

effect was only about 7 dB of masking release. This suggested an additional mechanism accounted for the remaining CMR effect, namely "dip" listening. This hypothesis was tested by adjusting the depth of modulation of CM noise from 100 to 0% modulated. When AM depth decreased from 100 to 90%, thresholds increased by 6 dB and then remained stable from 90 to 0% (Branstetter et al. 2013a). These results provided evidence that the total CMR effect was the combination of two factors, across channel coherence (about 7 dB of MR) and dip-listening (about 6 dB of MR).

Temporal resolution plays a central role in CMR and has been shown to be related to the TMTF (Berg 1996). As mentioned above, CM noise can be synthesized by multiplying low-pass noise by Gaussian noise. The cut-off frequency of the low-pass filter affects the AM rates of CM noise (Fig. 10), which provides a mechanism for evaluating the effect of AM rate on detection thresholds. In an experiment to this effect, Branstetter and Finneran (2008) found that lower AM rates—associated with lower-frequency cutoffs—resulted in a greater amount of CMR (Fig. 11). The same pattern was displayed in an AM masking study by Kastelein et al. (2021); however, at very different AM rates. In Kastelein et al. (2021), MR did not occur for AM rates above 20 Hz. In Branstetter and Finneran (2008), the CMR effect was still noticeable with AM rates up to at least 500 Hz (Fig. 11). The source of the variability between the two studies is unknown; however, a bottlenose dolphin was the subject in Branstetter and Finneran (2008), while a harbor porpoise was the subject in Kastelein et al. (2021), suggesting species-specific differences. More studies are required to determine the relationship between temporal resolution and masking release.

## Masking with environmental noise

The spectral-temporal properties of environmental noise can be varied and complex (Erbe et al. 2016), and as a result, lead to different levels of auditory masking. Comodulation masking release has been observed with different environmental noise types. For example, a beluga whale was trained to detect conspecific vocalizations within four different noise types: G noise, ice-creaking noise, propeller cavitation, and a bubbler noise system (Erbe and Farmer 1998, 2000). There was an 11-dB spread in masked detection thresholds, despite equal spectral densities between the noise types. Both dip-listening and CMR were likely candidates for the observed masking release. In a related experiment, two bottlenose dolphins demonstrated a release from masking with "environmental noise," which was a recording of ambient noise in San Diego Bay (Trickey et al. 2011). The recording was dominated by snapping shrimp and resulted in a 6-dB release from masking compared to G noise of the same spectral density. In a band-widening experiment

**Fig. 10** Power spectra from envelopes of different comodulated (CM) noise. Panels **A**, **B**, and **C** show comodulated noise produced by multiplying Gaussian noise with low-pass noise where the low-pass cutoff was 100, 500, and 4000 Hz, respectively. Panel **D** is from Gaussian noise. (Adapted from Branstetter and Finneran, 2008)
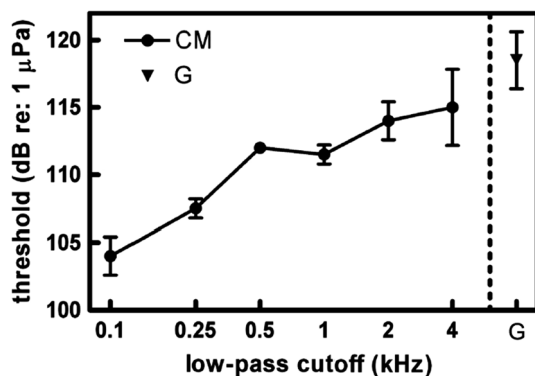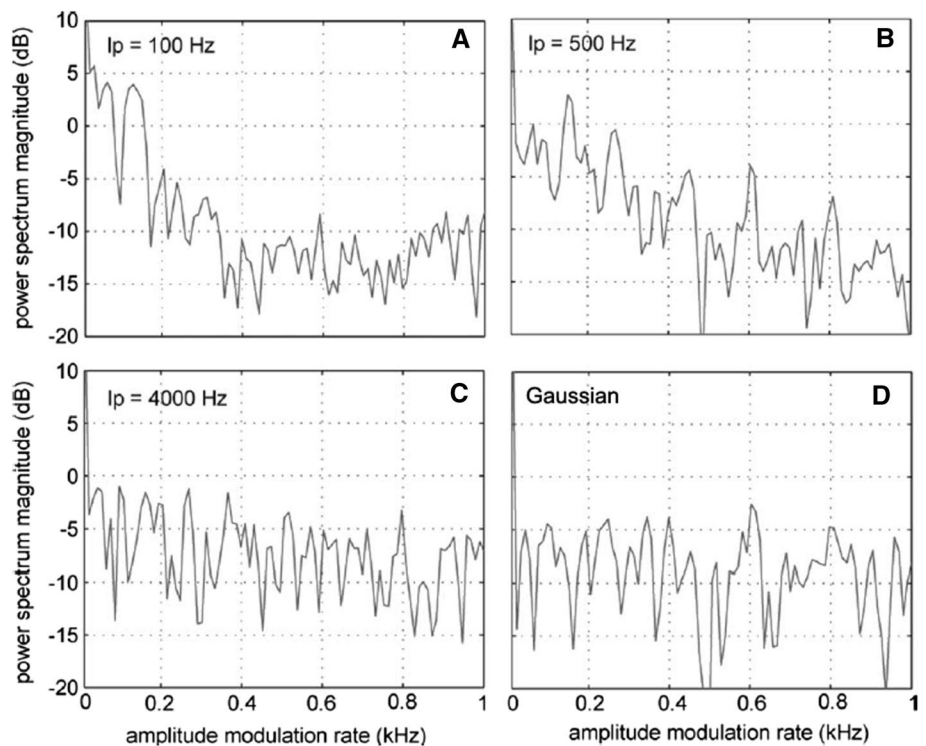


**Fig. 11** Detection thresholds in comodulated (CM) noise compared to Gaussian (G) noise. CM noise is generated by multiplying G noise by low-pass noise. Detection thresholds increase as a function of the low-pass filter cutoff. (Adapted from Branstetter and Finneran 2008)

with a bottlenose dolphin (Branstetter et al. 2013a), detection thresholds were measured for a 10 kHz tone masked by seven noise types: G noise, CM noise, snapping shrimp (SS), rain (RN), boat noise (BT), a pile saw (PS), and ice squeaks (IS). Spectrograms of each noise type are displayed in Fig. 12. For narrow-band maskers centered on 10 kHz (i.e., 1 kHz bandwidth and below), all noise types had a masking pattern consistent with the PSM, in which detection thresholds increased as a function of bandwidth (Fig. 13). However, when bandwidth exceeded the width of the auditory filter, masking patterns diverged. Two noise types, SS and CM, resulted in a release from masking, while RN and

G produced a masking pattern consistent with the PSM. Interestingly, both PS and IS noise, which have strong frequency-modulated components, resulted in elevated masking thresholds. To a human listener, both PS and IS have tonal qualities, making them categorically similar to the tonal signal. Informational masking may have been responsible for these elevated thresholds, which will be discussed below.

Although a mechanistic model has not been applied to describe and predict auditory masking in marine mammals with different noise types, Branstetter et al (2013b) produced a linear model to describe masking patterns with 12 different noise types and detection thresholds from three different bottlenose dolphins. Predictor variables were noise statistics, and they were divided into three categories: (1) waveform metrics (e.g., RMS and kurtosis), (2) frequency spectrum metrics (e.g., spectral density), and (3) temporal envelope metrics (e.g., envelope kurtosis, comodulation index). The comodulation index (CI) is a measure of how correlated temporal envelopes are across auditory filters. CI is calculated by band-pass filtering the noise sample into three bands that simulate auditory filters. The three bands are a signal band that is centered on the signal, and two flanking bands. The Hilbert envelope $env(t)$ is then calculated from the output of each filter

$$\text{env}(t) = \sqrt{f^2(t) + h^2(t)}, \tag{10}$$

where $f(t)$ is the time-domain waveform and $h(t)$ is the Hilbert transform. The magnitude squared coherence is then
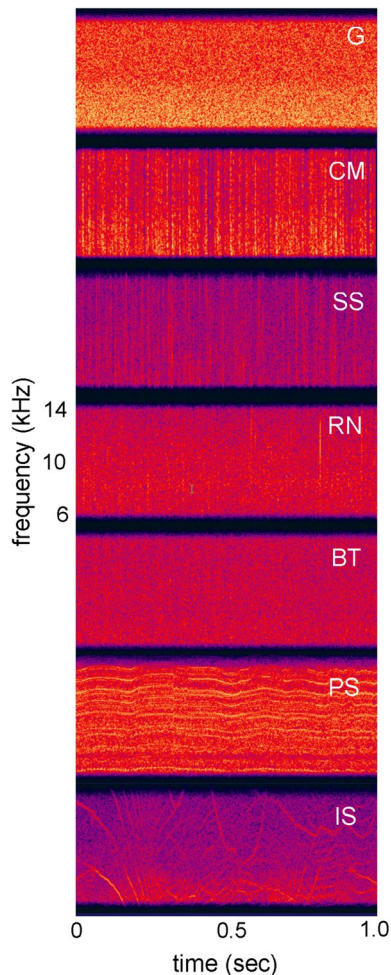
**Fig. 12** Spectrograms of different noise types: Gaussian (G), comodulated (CM) noise, snapping shrimp (SS), rain (RN), boat noise (BT), a pile saw (PS), and ice squeaks (IS). All noise types had flat spectral densities of 95 dB re (1 μPa)²/Hz. (Adapted from Branstetter et al. 2013a)
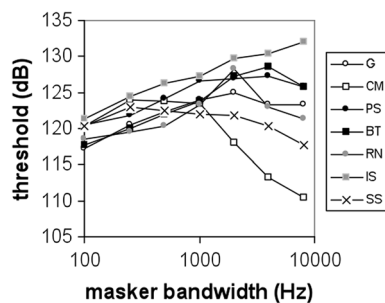


**Fig. 13** Masking patterns for a 10 kHz tone in the presence of different noise types. All noise types display similar masking patterns up to a bandwidth of 1 kHz. For noise bandwidths greater than 1 kHz, masking patterns diverge, representing different processing mechanisms. Noise types are Gaussian (G), comodulated (CM) noise, snapping shrimp (SS), rain (RN), boat noise (BT), a pile saw (PS), and ice squeaks (IS). (Adapted from Branstetter et al. 2013a)

calculated between the Hilbert envelopes of the signal band and each flanking band. The process is displayed in Fig. 14. An exponential decay model produced the best fit to the cumulative masking data

$$y = b_1 PSD + b_2 e^{(-CI/b_3)}, \tag{11}$$

where $y$ is the predicted detection threshold, PSD is the pressure spectral density, and $b_1$, $b_2$, and $b_3$ are fitting parameters. Values for $b_1$, $b_2$, and $b_3$ were 1.11, 31.54, and 0.37, respectively. A surface plot of the model can be found in Fig. 15. The PSD factor is linear, which is consistent with the PSM of masking: for every dB increase in the spectral density, there is a 1.11 dB increase in the detection threshold. The detection threshold is then mediated by the exponential decay function associated with the value of the CI (Fig. 15).

## Detection of complex signals in noise

The ability of the critical ratio to predict auditory masking levels is limited with non-Gaussian noise (e.g., Branstetter et al. 2013a, b). Similarly, complex signals can result in a departure from critical ratio predictions. Although the majority of auditory masking experiments have employed pure-tone signals, pure-tone sounds in nature are rare. Sounds that are perceived to have a tonal quality, such as a dolphin whistle, invariably contain multiple harmonics (Lammers et al. 2003), amplitude modulation (Jones et al. 2022), and frequency modulation (Janik and Sayigh 2013).

Cunningham et al. (2014) tested whether critical ratio predictions would generalize to complex signals. Masked detection thresholds were measured for a California sea lion (*Zalophus californianus*) and a harbor seal (*Phoca vitulina*), where the signals were either AM, frequency-modulated (FM), or contained multiple harmonics. For Gaussian-noise maskers, all signal types resulted in enhanced detectability compared to pure-tone signals (Fig. 16). For a shipping noise masker, which was AM and CM, the FM signal again resulted in enhanced detectability for both subjects. Conversely, for the harbor seal listener, detection of the AM signal in shipping noise resulted in more masking than expected (Fig. 16). The similarity between the AM signal and AM shipping noise may be an example of informational masking, which is covered below.

## Spatial release from masking

Current models of auditory masking (e.g., critical ratios and the power spectrum model) assume that both signal and noise are emitted from the same location, resulting in worst-case-scenario masking (Clark et al. 2009). In real ocean environments, noise sources are typically spatially separated
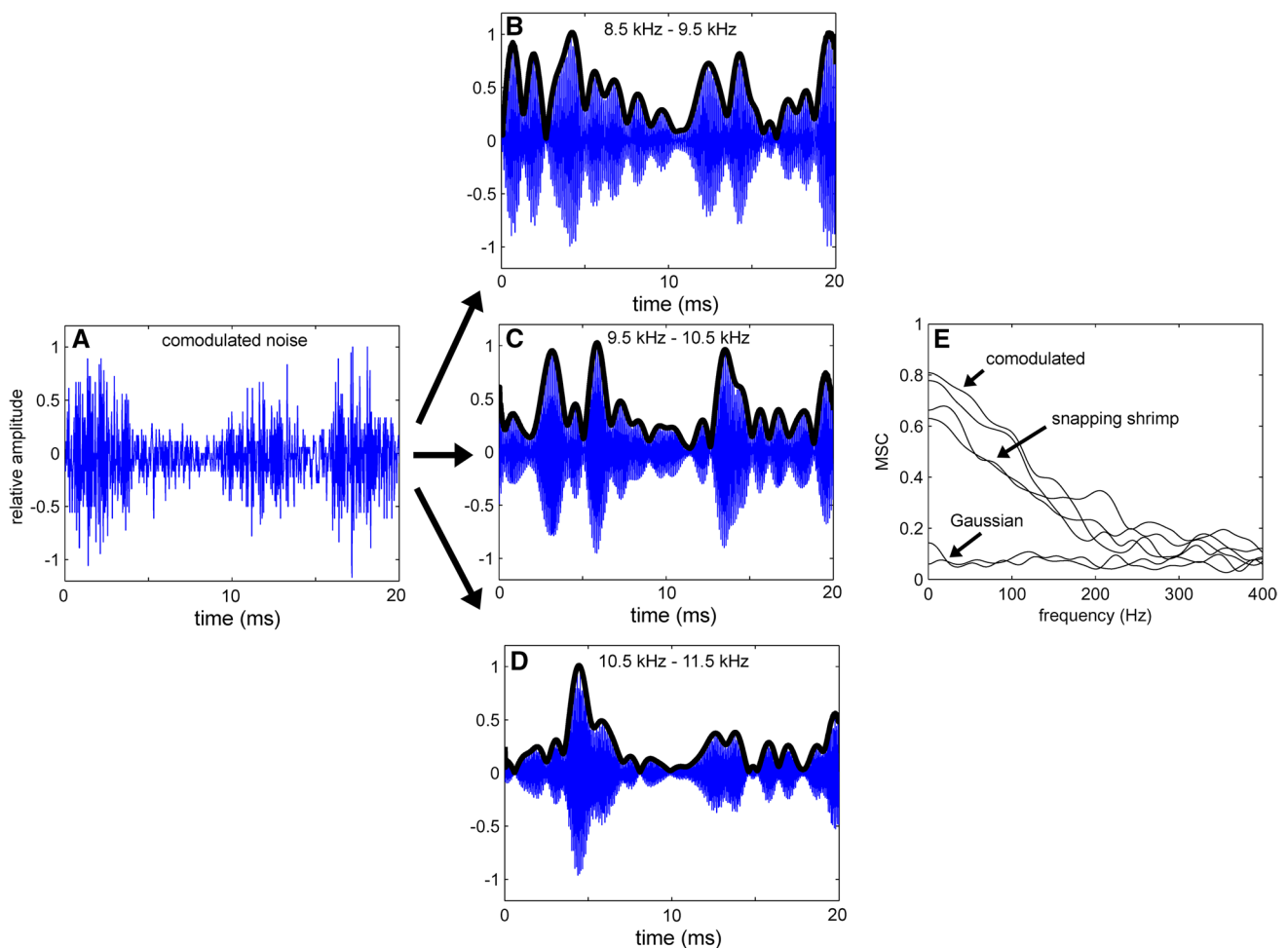
**Fig. 14** Processing stages used to calculate the comodulation index. Panel A is the time-domain waveform. Panels B, C, and D represent the outputs of the band-pass-filtered waveform. Thick black lines represent the Hilbert envelope. Panel E represents the magnitude squared coherence (MSC) between the temporal envelope of the signal band and the flanking bands for three different noise types. (Adapted from Branstetter et al. 2013b)

from biologically relevant signals. Under such conditions, spatial release from masking (SRM) can occur (e.g., Au and Moore, 1984; Holt and Schusterman, 2007; Popov et al. 2020). In humans, where research on this topic is extensive, the position of sound sources relative to a listener can act as one of the most salient cues to segregate multiple sounds in a complex auditory scene (Bregman 1990). Marine mammals have excellent sound localization abilities and directional receiving beam patterns (Accomando et al. 2020; Bodson et al. 2007; Au and Moore 1984; Branstetter and Mercado III 2006; Holt et al. 2004) which likely combine to aid the animal in separating auditory events, thus improving detection performance.

Au and Moore (1984) measured detection thresholds for a bottlenose dolphin listening for on-axis tones in the presence of Gaussian noise, where the angular location of the noise source varied. The most masking occurred when the signal and noise were co-located. As the angle between the signal and

noise increased, a release from masking occurred that became more salient with increasing frequency (Fig. 17). While testing was not conducted below 30 kHz, the trend in the data suggests that SRM would be less pronounced for low-frequency communication signals.

Holt and Schusterman (2007) measured SRM for airborne sounds in a harbor seal and a California sea lion. In this experiment, the location of the noise source was fixed at the on-axis position, while the location of tonal signals varied. To account for the directional sensitivity of each animal (i.e., detection thresholds vary as a function of angular location even without masking noise), the masking level difference (MLD) in dB was calculated

$$\text{MLD} = \left(M_q - M_o\right) - \left(U_q - U_o\right), \tag{12}$$

where $U_o$ and $U_q$ were the unmasked detection thresholds at 0° and $q°$, respectively, and $M_o$ and $M_q$ were the masked

$$y = 1.11PSD + 31.54e^{(-CI/0.37)}$$
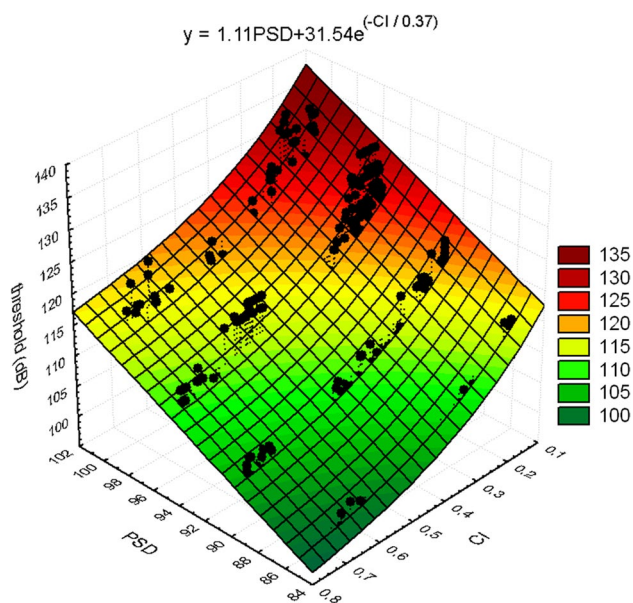
**Fig. 15** Surface plot of the exponential model used to describe dolphin detection thresholds in 12 different noise types with different levels of masking. PSD is the pressure spectral density and CI is the comodulation index. (Adapted from Branstetter et al. 2013b)

detection thresholds at 0° and $q$°. The pattern for MLD was complex, but greater levels were generally measured when signal and noise were separated (Fig. 18). Overall, there was an improvement in sensitivity due to spatial separation of signal and noise sources of up to 19 dB for the harbor seal and 12 dB for the sea lion.

The effect of SRM is profound (greater than 20 dB in dolphins), especially for higher frequencies. This effect is likely due to inter-aural loudness differences resulting from sound shadowing. However, current data are limited. For the dolphin in Au and Moore (1984), the noise source was always fixed at 0° azimuth, and for the pinnipeds in Holt and Schusterman (2007), the tonal source was always fixed at 0° azimuth. Although the focus in the current paper is on behavioral studies, it is important to note that an evoked potential study did measure SRM where the locations of both the signal and noise varied on the horizontal axis (Popov et al. 2020). In this study, the signal was restricted to one frequency, a 64-kHz tone pip, and the masker was band-pass noise between 40 and 90 kHz. The most SRM occurred when the signal and noise were ipsilateral to each other; however, the effect was only 8 dB. Because sound sources are much more likely to be separated in space than co-located, more research on SRM is warranted, especially for lower frequencies similar to both communication signals and anthropogenic noise.

# Energetic vs informational masking

The definitions of energetic and informational masking vary (Branstetter et al. 2016; Kidd et al. 2008; Pollack 1975; Wilson et al. 2012). However, many definitions suggest that energetic masking is a sensory phenomenon that occurs at the auditory periphery (e.g., Recio-Spinoso and Cooper 2013), while informational masking encompasses a wide variety of perceptual faculties including attention, memory, and recognition, and occurs at later stages of auditory processing (e.g., Branstetter et al. 2016; Kidd et al. 2008). For example, the ability to recognize human speech may be degraded in a noisy background. A human listener may clearly detect that a sound is present and even identify the sound as human speech without comprehending the semantic component of the sentence. Similarly, a bottlenose dolphin may be able to detect a whistle masked by noise, but unable to recognize the frequency contour and, thus, unable to recognize the identity of the caller. Dolphins use "signature whistles" or stereotyped whistles that uniquely identify the caller to conspecifics (Caldwell et al. 1990; Janik and Sayigh 2013). The stereotyped frequency-modulated pattern is used for recognition, while amplitude may encode other information such as emotional state (Jones et al. 2022). In an experiment in Branstetter et al. (2016), whistle-like FM signals were created to be equal in duration (500 ms) and bandwidth (8 kHz to 12 kHz) but vary in frequency contour (Fig. 19). Masked detection thresholds (energetic masking) were first measured for each signal using a standard go/no-go detection task with four noise types: ice squeaks, Gaussian, snapping shrimp and comodulated. The dolphin was then trained to associate each FM signal with a specific object using a matching-to-sample (MTS) paradigm. For example, when a linear FM sweep was presented (Fig. 19C), the dolphin was trained to swim and touch a rope. When a one-cycle sinusoidal FM signal was presented (Fig. 19A), the dolphin was taught to swim and touch a steel ball. Masked recognition thresholds were then measured using a three-alternative, forced-choice MTS test, using the same noise types from the signal detection task. Although the dolphin performed a forced-choice task where guessing would result in a correct answer 33% of the time, the dolphin did not respond at all for lower signal-to-noise ratios. As a result, no-response thresholds—defined as the average SPL where the dolphin only responded 50% of the time—were also reported. On average, recognition thresholds were 4 dB greater than no-response thresholds (Fig. 20). This result suggests that the cognitive processing needed to recognize the FM pattern and match it to a physical object required an additional 4 dB. This result is consistent with a 6-dB difference found between recognition and detection thresholds in birds (Dooling et al. 2009). Curiously, signal detection

**Fig. 16** Sensitivity difference from predicted (dB) for a harbor seal and a California sea lion listening for complex signals in the presence of flat-spectrum (Gaussian) noise or shipping noise. Positive values indicate that measured sensitivity was greater than predicted by critical ratios and a power spectrum model of masking. (Adapted from Cunningham et al. 2014)
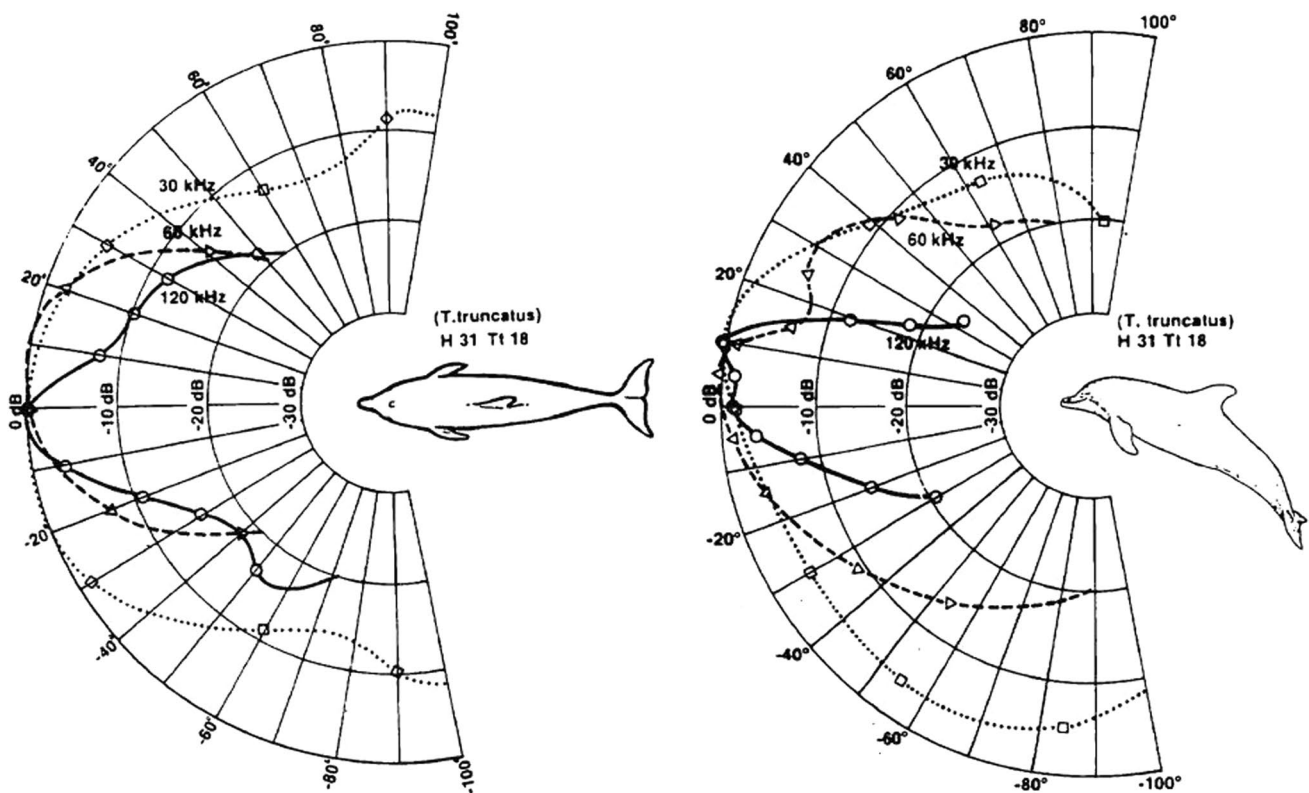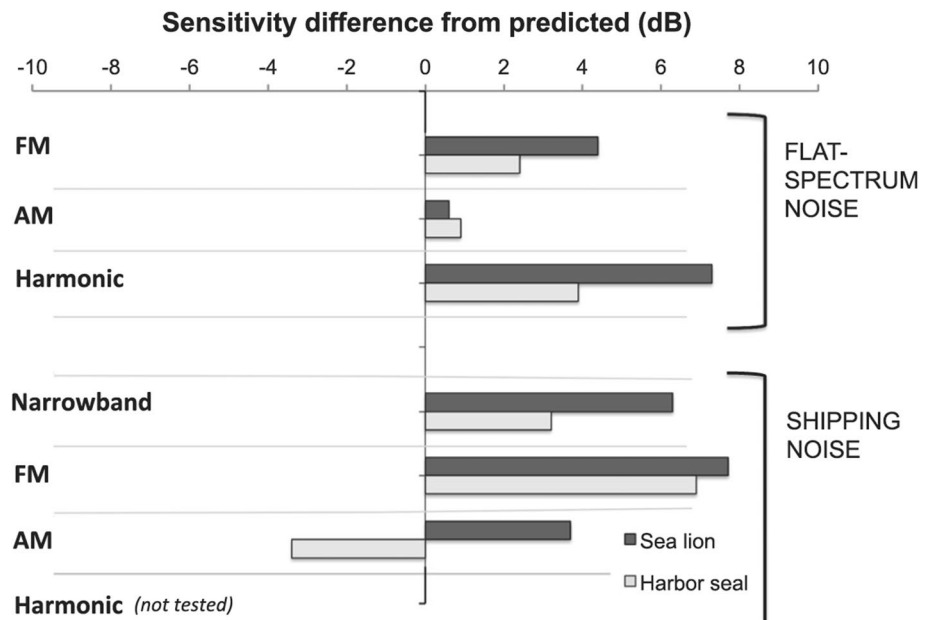
**Fig. 17** Receiver beam patterns for the bottlenose dolphin in the horizontal plane (left) and vertical plane (right) (Adapted from Au and Moore 1984)

thresholds were slightly elevated relative to the no-response thresholds, even though, hypothetically, they are measuring the same thing. However, methodological differences related to the go/no-go vs the MTS tasks may have had different

sustained attention demands that accounted for this difference (Branstetter et al. 2016). For example, in the MTS task, the signal was guaranteed to be presented within 5 s after the dolphin stationed on the bite plate, which is a relatively
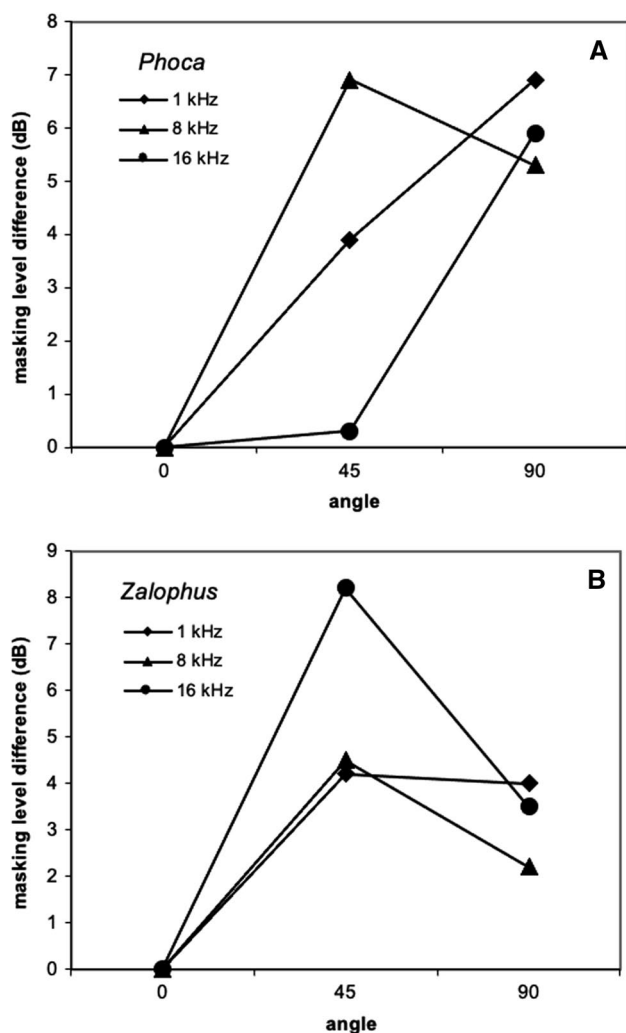
**Fig. 18** Masking level differences (in dB) relative to those measured at 0° for a harbor seal (Panel A) and California sea lion (Panel B) listening for tones projected from varying angles in the presence of on-axis masking noise. (Adapted from Holt and Schusterman 2007)

short amount of time to remain vigilant. However, for the go /no-go task, signals could occur anytime during a dive, and each dive lasted between 5 and 60 s, a much longer duration to remain vigilant. Another hypothesis is that the forced-choice task increases motivation to respond (compared to a go/no-go task) because guessing results in fish 33% of the time.

Ice squeak noise consistently produced the largest masked detection and recognition thresholds in this experiment (Figs. 13 and 20). Curiously, detection thresholds in IS noise were higher than recognition thresholds, which is a non-intuitive result. However, the similarity between the FM signals and the FM ice squeaks suggests that during the detection task, the signal likely registered in the dolphin's auditory system, but was misclassified as noise. To test this hypothesis, two FM noise types were created where one had

a predictable FM pattern, a linear down sweep (Fig. 21A), and the second noise type resembled IS noise—the frequency contour was random (Fig. 21B). Both noise types were presented at the same spectral density level. Detection thresholds were measured for two 500 ms signals with bandwidths between 8 and 12 kHz: a linear upsweep and a linear down sweep. Detection thresholds for both signal types in predictable FM noise were similar to detection thresholds in Gaussian noise (Fig. 22). However, detection thresholds in random FM noise produced strikingly elevated levels of masking, even higher than IS noise. When the noise had a predictable pattern, the presentation of the signal disrupted the pattern and was easy to identify. Masking levels were consistent with the PSM and represented energetic masking. However, when the background noise was random, the presentation of the signal was wrongfully classified as part of the background noise. In this case, thresholds were elevated even though the dolphin could likely hear the signal, potentially due to informational masking. In this study with simple linear upsweeps and down sweeps, there was a 12 dB difference between apparent signal detection and signal recognition; however, the frequency contour of dolphin whistles tends to be much more complex, which could possibly allow enhanced recognition in noisy backgrounds.
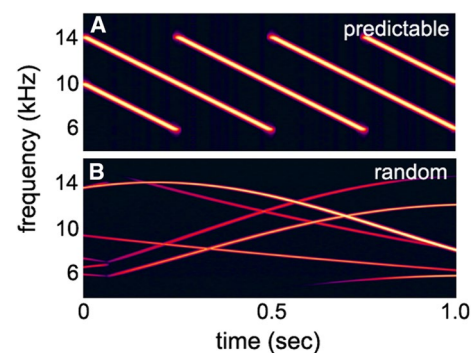
## Conclusions

Energetic masking likely occurs at the level of the cochlea (Recio-Spinoso and Cooper 2013) and can be divided into within-channel processing such as that described by the power spectrum model, and across-channel processing such as that which enables comodulation masking release. Noise types that lack significant amplitude modulation are often well described by the power spectrum model. Masking by impulsive sounds, such as seismic air guns, appears to be well predicted by the PSM with the addition of a time-window analysis that considers varying noise levels within shorter temporal intervals (Sills et al. 2017). For continuous amplitude-modulated noise, a dip-listening or valley-listening strategy is most effective for relatively lower AM rates, depending on the species. As AM rates increase, temporal resolution "smears" the valleys and detection thresholds approach levels similar to Gaussian noise (Branstetter and Finneran 2008; Kastelein et al. 2021). When broadband noise is coherently amplitude modulated across auditory filters, comodulation masking release can occur. The CMR effect can be disrupted by decorrelating the noise across channels, providing strong evidence that the auditory system compares temporal envelopes across auditory filters—or possibly groups sounds with similar temporal envelope patterns across frequency regions—to enhance signal detection (Branstetter et al. 2013a; Dau and Kollmeier 1997).

Fig. 19 Spectrograms of FM signals. The title of each panel represents the physical object that was associated with each signal in a matching-to-sample task. (Adapted from Branstetter et al. 2016)
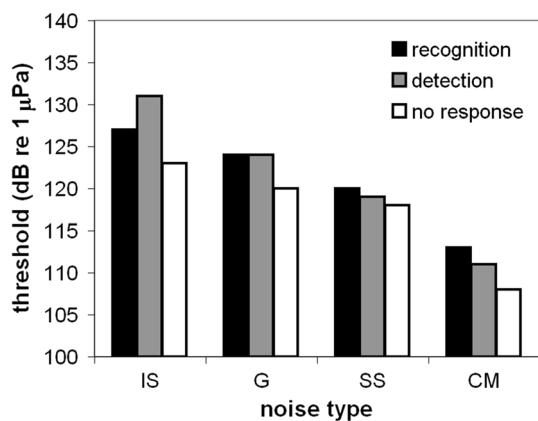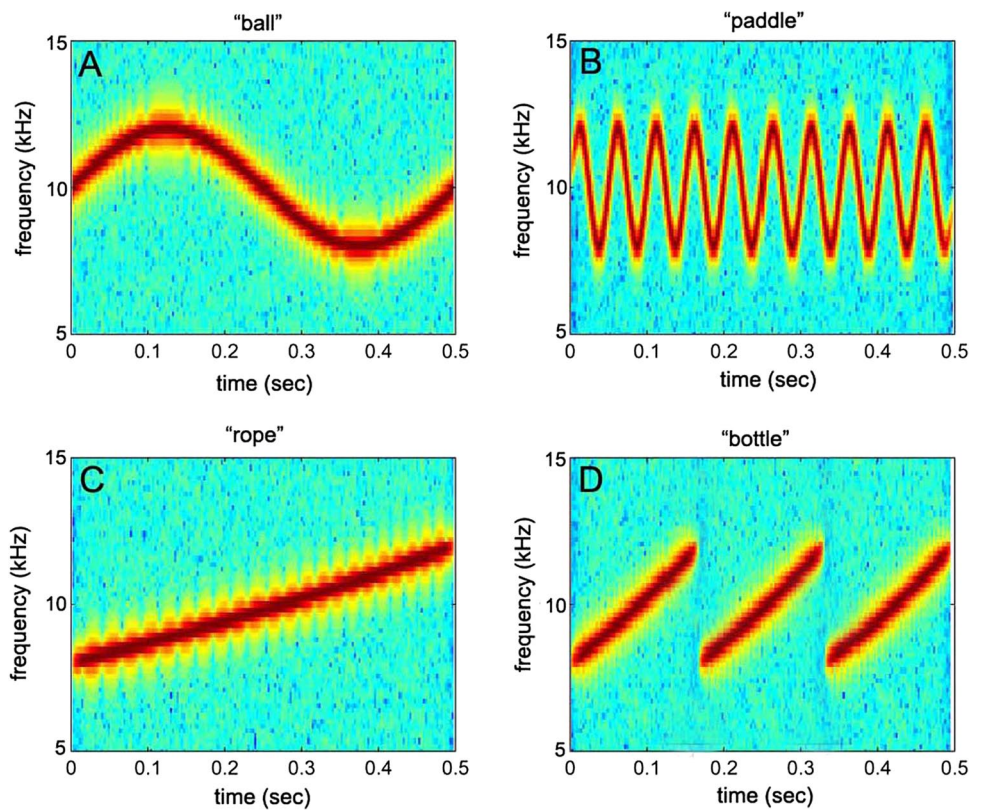


Fig. 20 Recognition, detection, and no-response thresholds in four different noise types: ice squeaks (IS), Gaussian (G), snapping shrimp (SS), and comodulated (CM) noise. (Adapted from Branstetter et al. 2016)
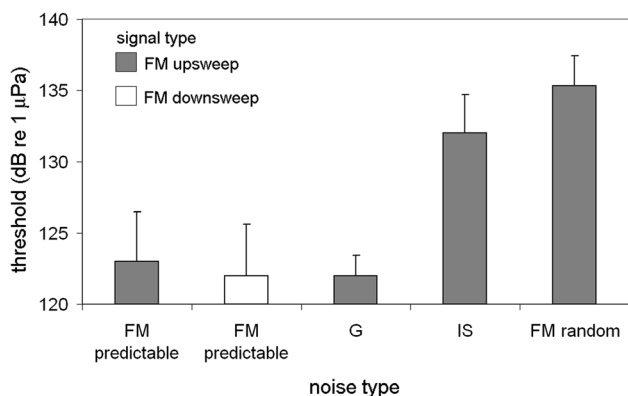


Fig. 21 Spectrograms of two different FM noise types, one with a predictable FM pattern and one with a random frequency contour. (Adapted from Branstetter et al. 2016)

Although studies of masked signal detection or energetic masking are useful for informing communication space models (Clark et al. 2009; Erbe et al. 2016; Jensen et al. 2012), animals not only need to detect a signal, they need to recognize the signal, as well. Being able to recognize a signal 80% of the time would have a much higher fitness utility than detection of a signal 50% of the time, yet 50% detection thresholds are the standard metric that

inform most models of masking (Branstetter et al. 2016; Clark et al. 2009). Currently, only one marine mammal study has investigated informational masking, resulting in recognition thresholds that are approximately 4–12 dB greater than detection thresholds (Branstetter et al. 2016). The exact mechanisms responsible for informational masking are unknown. However, failure to attend to the signal, or to segregate the signal from the background due to similarity, are likely candidates. A realistic masking scenario could be investigated by measuring recognition thresholds for a complex signal (e.g., a conspecific

**Fig. 22** Detection thresholds for two different signals presented within four different noise types. Spectrograms of FM predictable and FM random noise are displayed in Fig. 21. G and IS are Gaussian and Ice Squeak noise, respectively. (Adapted from Branstetter et al. 2016)

vocalization) in naturally occurring comodulated noise (e.g., snapping shrimp), where the noise is spatially separated from the signal. This is the type of acoustic environment that many species of coastal odontocetes and pinnipeds likely experience; however, this masking scenario has not been evaluated experimentally . A significant release from masking would be predicted from the combined effects of (1) enhanced  detection of the complex signal and (2) masking release from the comodulated noise and the spatially separated sound sources.

Accurate predictions of auditory masking are necessary to inform best management practices for marine mammals. Relatively simple masking models such as the power spectrum model—based on auditory filters and combined with standard critical ratios—can be applied effectively in certain situations. However, masking release (or conversely, elevated levels of masking) due to the spectral-temporal features of signals and noise and the spatial relationship of sound sources can complicate efforts to predict masking in realistic scenarios. More research is needed to better understand the mechanisms of auditory masking in marine mammals, and to improve the accuracy of masking predictions in the marine environment.

## References

Accomando AW, Mulsow J, Branstetter BK, Schlundt CE, Finneran JJ (2020) Directional hearing sensitivity for 2–30 kHz sounds in the bottlenose dolphin (Tursiops truncatus). J Acoust Soc Am 147:388–398. https://doi.org/10.1121/10.0000557

American National Standard Institute (ANSI) (1995) "Bioacoustical Terminology"

Au WWL, Moore PWB (1984) Receiving beam patterns and directivity indices of the Atlantic bottlenosed dolphin (*Tursiops truncatus*). J Acoust Soc Am 75:255–262

Au WWL, Moore PWB (1990) Critical ratio and critical bandwidth for the Atlantic bottlenose dolphin. J Acoust Soc Am 88:1635–1638

Au WWL, Branstetter BK, Benoit-Bird KJ, Kastelein RA (2009) Acoustic basis for fish prey discrimination by echolocating dolphins and porpoises. J Acoust Soc Am 126:460–467

Au WWL, Hastings MC (2008) Principles of Marine Bioacoustics, Springer-Verlag, New York. Retrieved from https://link.springer.com/book/https://doi.org/10.1007/978-0-387-78365-9

Berg BG (1996) On the relationship between comodulation masking release and temporal modulation transfer functions. J Acoust Soc Am 100:1013–1023

Bodson A, Miersch L, Dehnhardt G (2007) Underwater localization of pure tones by harbor seals (*Phoca vitulina*). J Acoust Soc Am 122:2263–2269

Branstetter BK, Finneran JJ (2008) Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). J Acoust Soc Am 124:625–633

Branstetter B, Mercado E III (2006) Sound localization by cetaceans. Int J Comp Psychol 19:26–61

Branstetter BK, Mercado E III, Au WWL (2007) Representing multiple discrimination cues in a computational model of the bottlenose dolphin auditory system. J Acoust Soc Am 122:2459–2468

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

Branstetter BK, Trickey JS, Aihara H, Finneran JJ, Liberman TR (2013b) Time and frequency metrics related to auditory masking of a 10 kHz tone in bottlenose dolphins (*Tursiops truncatus*). J Acoust Soc Am 134:4556–4565

Branstetter BK, Bakhtiari K, Black A, Trickey JS, Finneran JJ, Aihara H (2016) Energetic and informational masking of complex

sounds by a bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 140:1904. https://doi.org/10.1121/1.4962530

Branstetter BK, Van Alstyne KR, Wu TA, Simmons RA, Curtis LD, Xitco MJ Jr (2017) Critical ratio functions for odontocete cetaceans. J Acoust Soc Am 142:1897–1900

Branstetter BK, Van Alstyne KR, Strahan MG, Tormey MN, Wu T, Breitenstein RA, Houser DS et al (2020) Spectral cues and temporal integration during cylinder echo discrimination by bottlenose dolphins (*Tursiops truncatus*). J Acoust Soc Am 148:614–626. https://doi.org/10.1121/10.0001626

Branstetter BK, Felice M, Robeck T (2021) Auditory masking in killer whales (Orcinus orca): Critical ratios for tonal signals in Gaussian noise. J Acoust Soc Am 149:2109. https://doi.org/10.1121/10.0003923

Bregman AS (1990) Auditory Scene Analysis: The Perceptual Organization of Sound. The MIT Press, Massachusetts

Buschermohle M, Verhey JL, Feudel U, Freund JA (2007) The role of the auditory periphery in comodulation detection difference and comodulation masking release. Biol Cybern 97:397–411. https://doi.org/10.1007/s00422-007-0179-8

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

Caldwell MC, Caldwell DK, Tyack TL (1990) Review of the signature-whistle hypothesis for the Atlantic Bottlenose Dolphin. In: Leatherwood S, Reeves RR (eds) The Bottlenose Dolphin. Academic Press Inc, San Diego, pp 199–234

Clark CW, Ellison WT, Southall BL, Hatch L, Van Parijs SM, Frankel A, Ponirakis D (2009) Acoustic masking in marine ecosystems: intuitions, analysis, and implication. Mar Ecol Prog Ser 395:201–222

Cunningham KA, Southall BL, Reichmuth C (2014) Auditory sensitivity of seals and sea lions in complex listening scenarios. J Acous Soc Am 136(6):3410–3421. https://doi.org/10.1121/1.4900568

Dau T, Kollmeier B (1997) Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. J Acoust Soc Am 102:2892–2905

Dolphin WF, Au WW, Nachtigall PE, Pawloski J (1995) Modulation rate transfer functions to low-frequency carriers in three species of cetaceans. J Comp Physiol A 177:235–245

Dooling RJ, West EW, Leek MR (2009) "Conceptual and computational models of the effects of anthropogenic noise on birds." Proc Inst Acoust 31(1)

Erbe C, Farmer DM (1998) Masked hearing thresholds of a beluga whale (*Delphinapterus leucas*) in icebreaker noise. Deep-Sea Res 45:1373–1378

Erbe C, Farmer DM (2000) Zones of impact around icebreakers affecting beluga whales in the Beaufort Sea. J Acoust Soc Am 109:1332–1340

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

Fastl H, Zwicker E (2007) Psychoacoustics - Facts and Models, Springer Series in Information Sciences, Springer-Verlag, Berlin, 3rd ed., Vol. 22, 462 pages

Fay RR (1988) Hearing in vertebrates: A psychophysics handbook, Hill-Fay Associates, Winnetka, Illinois, 621 pages

Fechner GT (1860) Elemente der Psychophysik, Breitkopf u. Härtel, 590 pages

Finneran JJ (2015) Noise-induced hearing loss in marine mammals: A review of temporary threshold shift studies from 1996 to 2015. J Acoust Soc Am 138:1702–1726. https://doi.org/10.1121/1.4927418

Finneran JJ, Schlundt CE, Carder DA, Ridgway SH (2002) Auditory filter shapes for the bottlenose dolphin (*Tursiops truncatus*) and the white whale (*Delphinapterus leucas*) derived with notched noise. J Acoust Soc Am 112:322–328

Finneran JJ, London HR, Houser DS (2007) Modulation rate transfer functions in bottlenose dolphins (*Tursiops truncatus*) with normal hearing and high-frequency hearing loss. J Comp Physiol A 193:835–843

Finneran JJ, Branstetter B (2013) Effects of noise on sound perception in marine mammals. Animal Commun Noise. https://doi.org/10.1007/978-3-642-41494-7_10

Fletcher H (1940) Auditory patterns. Rev Mod Phys 12:47–65

Frankel AS, Clark CW (2000) Behavioral responses of humpback whales (*Megaptera novaeangliae*) to full-scale ATOC signals. J Acoust Soc Am 108:1930–1937

Guan S, Vignola J, Judge J, Turo D (2015) Airgun inter-pulse noise field during a seismic survey in an Arctic ultra shallow marine environment. J Acoust Soc Am 138:3447. https://doi.org/10.1121/1.4936904

Hall JW, Grose JH (1988) Comodulation masking release: Evidence for multiple cues. J Acoust Soc Am 84:1669–1675

Heffner HE, Heffner RS (2008) High-frequency hearing. In: Dallos P, Oertel D, Hoy R (eds) Handbook of the Senses: Audition. Elsevier, NY., pp 55–60

Holt MM, Schusterman RJ (2007) Spatial release from masking of aerial tones in pinnipeds. J Acoust Soc Am 121:1219–1225

Holt MM, Schusterman RJ, Southall BL, Kastak D (2004) Localization of aerial broadband noise by pinnipeds. J Acoust Soc Am 115:2339–2345

Holt M, Noren DP, Veirs V, Emmons CK, Veirs S (2009) Speaking up: Killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. J Acoust Soc Am 125:EL27–EL32

Houser DS, Martin S, Crocker DE, Finneran JJ (2020) Endocrine response to simulated U.S. Navy mid-frequency sonar exposures in the bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am 147:1681–1687. https://doi.org/10.1121/10.0000924

Janik VM, Sayigh LS (2013) Communication in bottlenose dolphins: 50 years of signature whistle research. J Comp Physiol A 199:479–489

Jensen FH, Beedholm K, Wahlberg M, Bejder L, Madsen PT (2012) Estimating communication range and energetic cost of bottlenose dolphin whistles in a tropical habitat. J Acoust Soc Am 131:582–592

Johnson CS (1967) Sound detection thresholds in marine mammals. In: Tavolga WN (ed) Marine Bioacoustics. Pergamon Press, Oxford, pp 247–260

Johnson CS (1968a) Masked tonal thresholds in the bottlenosed porpoise. J Acoust Soc Am 44:965–967

Johnson CS (1968b) Relation between absolute threshold and duration of tone pulses in the bottlenosed dolphin. J Acoust Soc Am 43:757–763

Johnson CS (1971) Auditory masking of one pure tone by another in the bottlenosed porpoise. J Acoust Soc Am 49:1317–1318

Johnson RA, Moore PWB, Stoermer MW, Pawloski JL, Anderson LC (1988) Temporal order discrimination within the dolphin critical interval. In: Nachtigall PE, Moore PWB (eds) Animal Sonar Processes and Performance. Plenum Press, New York, pp 317–321

Johnson CS, McManus MW, Skaar D (1989) Masked tonal hearing thresholds in the beluga whale. J Acoust Soc Am 85:2651–2654

Jones B, Tufano S, Daniels R, Mulsow J, Ridgway S (2022) Nonstereotyped amplitude modulation across signature whistle contours. Behav Proc 194:104561. https://doi.org/10.1016/j.beproc.2021.104561

Kastelein RA, Wensveen PJ (2008) Effect of two levels of masking noise on the hearing threshold of a harbor porpoise (*Phocoena phocoena*) for a 4.0 kHz signal. Aquat Mamm 34:420–425

Kastelein RA, Helder-Hoek L, Covi J, Terhune JM, 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. https://doi.org/10.1121/10.0006103

Kastelein RA, Wensveen PJ, Hoek L, Au WWL, Terhune JM, De Jong CAF (2009) J Acoust Soc Am 126(3):1588–1597. https://doi.org/10.1121/1.3177274

Ketten DR (1992a) The marine mammal ear: Specializations for aquatic audition and echolocation. In: Webster D, Fay R, Popper A (eds) The Biology of Hearing. Springer-Verlag, New York, pp 717–754

Ketten DR (1992b) The cetacean ear: form, frequency, and evolution. In: Thomas JA, Kastelein RA, Supin AY (eds) Marine Mammal Sensory Systems. Plenum Press, New York, pp 53–75

Kidd G, Mason CR, Richards VM, Gallun FJ, Durlach NI (2008) "Informational Masking," In W. A. Yost, A. N. Popper, and R. R. Fay (Eds.), Auditory Perception of Sound Sources, Springer Handbook of Auditory Research, Springer US, Boston, MA, pp. 143–189. https://doi.org/10.1007/978-0-387-71305-2_6

Klump GM, Nieder A (2001) Release from masking in flucuating background noise in a songbird's auditory forbrain. Neuroethology 12:1825–1829

Lammers MO, Au WWL, Herzing DL (2003) The broadband social acoustic signaling behavior of spinner and spotted dolphins. J Acoust Soc Am 114:1629–1639

Lemonds DW, Kloepper LN, Nachtigall PE, Au WWL, Vlachos SA, Branstetter BK (2011) A re-evaluation of auditory filter shape in delphinid odontocetes: Evidence of constant-bandwidth filters. J Acoust Soc Am 130:3107–3114. https://doi.org/10.1121/1.3644912

Lemonds DW (1999) Auditory filter shapes in an Atlantic bottlenose dolphin (Tursiops truncatus) University of Hawaii, 74 pages

Levitt H (1971) Transformed up-down methods in psyhcoacoustics. J Acoust Soc Am 49:467–477

Lyon RF, Katsiamis AG, Drakakis EM (2010) "History and future of auditory filter models," Proceedings of 2010 IEEE International Symposium on Circuits and Systems, 3809–3812. Presented at the Proceedings of 2010 IEEE International Symposium on Circuits and Systems. https://doi.org/10.1109/ISCAS.2010.5537724

Moore BCJ (1993) Frequency analysis and pitch perception. In: Yost WA, Popper AN, Fay RR (eds) Human Psychophysics. Springer-Verlag, New York, pp 56–115

Moore PWB, Schusterman RJ (1987) Audiometric assessment of northern fur seals, Callorhinus ursinus. Mar Mamm Sci 3:31–53

Moore PWB, Hall RW, Friedl WA, Nachtigall PE (1984) The critical interval in dolphin echolocation: What is it? J Acoust Soc Am 76:314–317

Narayan R, Graña G, Sen K (2006) Distinct time scales in cortical discrimination of natural sounds in songbirds. J Neurophysiol 96:252–258. https://doi.org/10.1152/jn.01257.2005

Nelken I, Rotman Y, Bar Yosef O (1999) Response of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157

NOAA Fisheries (2018). Technical Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammal Hearing | NOAA Fisheries, NOAA, Available: https://www.fisheries.noaa.gov/resource/document/technical-guidance-assessing-effects-anthropogenic-sound-marine-mammal-hearing, (date last viewed: 20-Oct-20). Retrieved October 20, 2020, from https://www.fisheries.noaa.gov/resource/document/technical-guidance-assessing-effects-anthropogenic-sound-marine-mammal-hearing

Parsons ECM (2017) Impacts of navy sonar on whales and dolphins: now beyond a smoking gun? Front Mar Sci. https://doi.org/10.3389/fmars.2017.00295

Patterson RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640–654

Patterson RD, Nimmo-Smith I, Weber DL, Milroy R (1982) The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. J Acoust Soc Am 72:1788–1803

Pollack I (1975) Auditory informational masking. J Acoust Soc Am 57:S5

Popov VV, Supin AY, Wang D, Wang K (2006) Nonconstant quality of auditory filters in the porpoises, Phocoena phocoena and Neophocaena phocaenoides (Cetacea, Phocoenidae). J Acoust Soc Am 119:3173–3180. https://doi.org/10.1121/1.2184290

Popov VV, Supin AY, Gvozdeva AP, Nechaev DI, Tarakanov MB, Sysueva EV (2020) Spatial release from masking in a bottlenose dolphin Tursiops truncatus. J Acoust Soc Am 147:1719. https://doi.org/10.1121/10.0000909

Recio-Spinoso A, Cooper NP (2013) Masking of sounds by a background noise–cochlear mechanical correlates. J Physiol 591:2705–2721. https://doi.org/10.1113/jphysiol.2012.248260

Renouf D (1980) Masked hearing thresholds of harbour seals (Phoca vitulina) in air. J Audit Res 20:263–269

Richardson WJ, Greene CR, Malme CI, Thomson DH (1995) Marine Mammals and Noise. Academic Press

Roitblat HL, Ketten D, Au WWL, Nachtigall PE (1996) A computational model of early stages of dolphin hearing (A). J Acoust Soc Am 100:2643

Roitblat H (1998) "Recognition processes in echolocation." Presented at the Proceedings of the Biological Sonar Conference

Ruscher B, Sills JM, Richter BP, Reichmuth C (2021) In-air hearing in Hawaiian monk seals: implications for understanding the auditory biology of Monachinae seals. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 207:561–573. https://doi.org/10.1007/s00359-021-01498-y

Scheifele PM, Andrew S, Cooper RA, Darre M, Musiek FE, Max L (2005) Indication of a Lombard vocal response in the St. Lawrence River beluga. J Acoust Soc Am 117:1486–1492

Serrano A, Terhune JM (2001) Within-call repetition may be an anti-masking strategy in underwater calls of harp seals (Pagophilus groenlandicus). Can J Zool 79(8):1410–1413. https://doi.org/10.1139/cjz-79-8-1410

Sills JM, Southall BL, Reichmuth C (2014) Amphibious hearing in spotted seals (Phoca largha): underwater audiograms, aerial audiograms and critical ratio measurements. J Exp Biol 217:726–734. https://doi.org/10.1242/jeb.097469

Sills JM, Southall BL, Reichmuth C (2015) J Exp Biol. https://doi.org/10.1242/jeb.120972

Sills JM, Southall B, 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

Sills JM, Reichmuth C, Southall BL, Whiting A, Goodwin J (2020) Auditory biology of bearded seals (Erignathus barbatus). Polar Biol 43:1681–1691. https://doi.org/10.1007/s00300-020-02736-w

Southall BL, Schusterman RJ, Kastak D (2000) Masking in three pinnipeds: underwater, low-frequency critical ratios. J Acoust Soc Am 108:1322–1326

Southall BL, Schusterman RJ, Kastak D (2003) Auditory masking in three pinnipeds: Aerial critical ratios and direct critical bandwidth measurements. J Acoust Soc Am 114:1660–1666

Southall BL, Finneran JJ, Reichmuth C, Nachtigall PE, Ketten DR, Bowles AE, Ellison WT et al (2019) Marine mammal noise exposure criteria: updated scientific recommendations for residual hearing effects. Aquat Mamm 45:125–232. https://doi.org/10.1578/AM.45.2.2019.125

Southall BL, Tyack PL, Moretti D, Clark C, Claridge D, Boyd I (2009) Behavioral responses of beaked whales and other

cetaceans to controlled exposures of simulated sonar and other sounds. Presented at the 18th Biennial Conference on the Biology of Marine Mammals

Southall BL (2005) Final Report of the 2004 National Oceanic and Atmospheric Administration (NOAA) International Symposium: Shipping Noise and Marine Mammals: A Forum for Science, Management, and Technology Arlington, Virginia: NOAA Fisheries Acoustics Program, Office of Protected Resources, NMFS, NOAA

Stebbins WC (1970) "Principles of Animal Psychophysics," In: Stebbins WC (Ed.), Animal Psychophysics: the design and conduct of sensory experiments, Springer US, Boston, MA, pp. 1–19. https://doi.org/10.1007/978-1-4757-4514-6_1

Teng X, Tian X, Poeppel D (2016) Testing multi-scale processing in the auditory system. Sci Rep 6:34390. https://doi.org/10.1038/srep34390

Terhune JM (1991) Masked and unmasked pure tone thresholds of a harbour seal listening in air. Can J Zool 69:2059–2066

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

Terhune JM, Ronald K (1975) Masked hearing thresholds of ringed seals. J Acoust Soc Am 58:515–516

Thomas JA, Pawloski JL, Au WWL (1990) Masked hearing abilities in a False Killer Whale (Pseudorca crassidens). In: Thomas JA, Kastelein RA (eds) Sensory Abilities of Cetaceans: Laboratory and Field Evidence. Plenum Press, New York, pp 395–404

Trickey JS, Branstetter BK, Finneran JJ (2011) Auditory masking with environmental, comodulated, and Gaussian noise in bottlenose dolphins (*Tursiops truncatus*). J Acoust Soc Am 128:3799–3804

Turnbull SD (1994) Changes in masked thresholds of a harbour seal (Phoca vitulina) associated with angular separation of signal and noise sources. Can J Zool 72:1863–1866. https://doi.org/10.1139/z94-253

Turnbull SD, Terhune JM (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 SD, Terhune JM (1993) Repetition enhances hearing detection thresholds in a harbour seal (*Phoca vitulina*). Can J Zool 71:926–932

Vaage S, Haugland K, Utheim T (1983) Signatures from single airguns. Geophys Prospect 31:87–97

Viemeister NF (1979) Temporal modulation transfer functions based on modulation thresholds. J Acoust Soc Am 66:1364–1380

Viemeister NF, Wakefield GH (1991) Temporal integration and multiple looks. J Acoust Soc Am 90:858–865

Wilson RH, Trivette CP, Williams DA, Watts KL (2012) The effects of energetic and informational masking on The Words-in-Noise Test (WIN). J Am Acad Audiol 23:522–533. https://doi.org/10.3766/jaaa.23.7.4

Wright AJ, Aguilar Soto N, Linda Baldwin A, Bateson M, Beale CM, Clark C, Deak T et al (2007) Anthropogenic noise as a stressor in animals: A multidisciplinary perspective. Int J Comp Psychol 20:250–273

Yost WA, Fay RR (2012) Human Psychophysics, Springer Science & Business Media, 253 pages