



**HAL**  
open science

## **Bone-conducted sound in a dolphin's mandible: Experimental investigation of elastic waves mediating information on sound source position**

Michael Reinwald, Quentin Grimal, Jacques Marchal, Stefan Catheline, Lapo  
Boschi

► **To cite this version:**

Michael Reinwald, Quentin Grimal, Jacques Marchal, Stefan Catheline, Lapo Boschi. Bone-conducted sound in a dolphin's mandible: Experimental investigation of elastic waves mediating information on sound source position. *Journal of the Acoustical Society of America*, 2018, 144 (4), pp.2213-2224. 10.1121/1.5063356 . hal-01981636

**HAL Id: hal-01981636**

**<https://hal.sorbonne-universite.fr/hal-01981636>**

Submitted on 15 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Bone-conducted sound in a dolphin's mandible: Experimental investigation of elastic waves mediating information on sound source position**

Michael Reinwald,<sup>1, a)</sup> Quentin Grimal,<sup>1, b)</sup> Jacques Marchal,<sup>2, c)</sup> Stefan Catheline,<sup>3, d)</sup> and Lapo Boschi<sup>4, e)</sup>

<sup>1</sup>*Sorbonne Université, CNRS, INSERM, Laboratoire d'Imagerie Biomédicale, LIB, F-75006, Paris, France.*

<sup>2</sup>*Sorbonne Université, CNRS, Institut Jean le Rond d'Alembert, F-78210, Saint-Cyr-l'École, France.*

<sup>3</sup>*LabTAU, INSERM, Centre Léon Bèrard, Université Lyon 1, Univ. Lyon, F-69003, Lyon, France.*

<sup>4</sup>*Sorbonne Université, CNRS-INSU, Institut des Sciences de la Terre Paris, ITeP UMR 7193, F-75005 Paris, France.*

(Dated: 29 September 2018)

1 Mammals use binaural or monaural (spectral) cues to localize acoustic sources. While  
2 the sensitivity of terrestrial mammals to changes in source elevation is relatively poor,  
3 the accuracy achieved by the odontocete cetaceans' biosonar is high, independently of  
4 where the source is. Binaural/spectral cues are unlikely to account for this remarkable  
5 skill. We study bone-conducted sound in a dolphin's mandible, investigating its  
6 possible contribution to sound localization. Experiments are conducted in a water  
7 tank by deploying, on the horizontal and median planes of the skull, ultrasound  
8 sources that emit synthetic clicks between 45-55 kHz. Elastic waves propagating  
9 through the mandible are measured at the pan bones and used to localize source  
10 positions via either binaural cues or a correlation-based full-waveform algorithm.  
11 Exploiting the full waveforms and, most importantly, their reverberated coda, we  
12 can enhance the accuracy of source localization in the vertical plane, and achieve  
13 similar resolution of horizontal- vs. vertical-plane sources. Our results need to be  
14 substantiated by further experimental work, accounting for soft tissues and making  
15 sure that the data are correctly mediated to the internal ear. If confirmed, they would  
16 favor the idea that dolphin's echolocation skills rely on the capability to analyze the  
17 coda of biosonar echoes.

---

a) [michael.reinwald@upmc.fr](mailto:michael.reinwald@upmc.fr)

b) [quentin.grimal@upmc.fr](mailto:quentin.grimal@upmc.fr)

c) [jacques.marchal@upmc.fr](mailto:jacques.marchal@upmc.fr)

d) [stefan.catheline@inserm.fr](mailto:stefan.catheline@inserm.fr)

e) [lpo.boschi@upmc.fr](mailto:lpo.boschi@upmc.fr)

## 18 I. INTRODUCTION

19 The acoustic environment of marine mammals is very different from that of humans and  
20 other terrestrial mammals. Water is much denser than air, and sound travels five times  
21 faster through water than through air and is less strongly attenuated; the energy carried  
22 by acoustic waves is more efficiently transferred to bone tissue from water than from air;  
23 presumably because they would be disadvantageous from the point of view of locomotion in  
24 water, marine mammals have lost pinnae through evolution; their ear canals are typically  
25 filled with cellular debris and appear to play no functional role in hearing ([Ketten, 1997](#)).

26 Marine mammals use echolocation to navigate and hunt. For about two centuries ([Hunter  
27 and Banks, 1787](#)), they have been known to complete such tasks with remarkable accuracy  
28 and efficiency. The contribution of dolphin’s anatomy to audition-related tasks was first  
29 evaluated by Kenneth Norris in a suite of groundbreaking studies ([Norris, 1964; 1968a,b;  
30 Norris and Harvey, 1974](#)). A dolphin’s mandible is very thin, almost “translucent,” at its  
31 posterior end (0.5 mm to 3.0 mm thickness, depending on the species), and is overlain by  
32 an oval fatty volume, which connects the posterior jaw bone, also named pan bone, with the  
33 tympano-periotic complex (TPC). Norris suggested that sound propagates through the thin  
34 pan bone, entering the fats which possibly act as a low-impedance wave guide that directs  
35 sound towards the inner ear. This is still the most widely accepted theory of the sound  
36 propagation pathway for hearing in cetaceans ([Au, 2012; Au and Hastings, 2008; Brill \*et al.\*,  
37 2001; Mooney \*et al.\*, 2012](#)) and is supported by experimental ([Brill \*et al.\*, 1988; Norris and  
38 Harvey, 1974](#)) and numerical ([Aroyan, 2001](#)) results. Norris’ “jaw-bone theory” has been

39 further developed in more recent studies: While high-frequency sounds could propagate  
40 through the jaw bone, low-frequency sounds (below 30 kHz) could propagate through bone-  
41 free fat channels just below the eyes and posterior to the lower jawbone (Brill *et al.*, 2001;  
42 Ketten, 1994; Popov and Supin, 1990). Cranford et al. simulated sound propagation in the  
43 head of a Cuvier’s beaked whale and introduced the idea of a “gular pathway” for sound  
44 reception (Cranford *et al.*, 2008), with sound entering the head from below and between the  
45 mandibular walls rather than through the posterior mandible. Some studies propose that the  
46 teeth play an important role in sound reception; due to their periodic placement, they could  
47 act as an acoustic metamaterial, resonating and amplifying sound for specific frequencies  
48 (Dible *et al.*, 2009; Graf *et al.*, 2009). Others suggest that the morphology of the mental  
49 foramens found in the mandible helps sound enter the fatty tissues (Ryabov, 2010). Møhl  
50 et al. (Møhl *et al.*, 1999) measured the auditory brainstem evoked potential for an Atlantic  
51 bottlenose dolphin by placing suction cups on various positions on the animal’s head and  
52 emitting acoustic click stimuli via embedded piezoelectric transducer elements. Their results  
53 suggest that the area of maximum sensitivity is slightly forward of the pan bone area. This  
54 does not contradict Norris’ jaw bone theory since the outer mandibular fat channel extends  
55 forward to the skin and towards the area where Møhl et al. found sensitivity to be highest.

56 It has been suggested by experiments (Blauert, 1997; Moore *et al.*, 1995; Renaud and  
57 Popper, 1975) that dolphins locate sound sources via binaural cues known to be employed  
58 by terrestrial animals, i.e. interaural time differences (ITD), which describe the delay of a  
59 signal arriving at the two ears at two different times, and interaural level differences (ILD),  
60 which describe the difference in intensity between the signals perceived at the two ears.

61 While ITD are relatively easy to reproduce theoretically, ILD effects are more complex, as  
62 they cannot be modeled by simply accounting for differences in source-receiver distance:  
63 they are importantly affected by sound shadowing due to the impedance mismatch between  
64 the subject’s head and the surrounding propagation medium (Mooney *et al.*, 2012), while  
65 a significant fraction of acoustic energy traveling from one ear to the other is diffracted by  
66 the head’s surface, thus following a complex propagation path. In any case, binaural cues  
67 are only relevant as long as a sound originates from somewhere else than the median plane.  
68 Because median-plane sources are equidistant from both ears, no phase or amplitude (if  
69 the subject’s anatomical features are symmetric with respect to the median plane, as they  
70 most often are) differences exist between the signals perceived at the two ears, i.e. zero ITD  
71 and ILD (e.g. Butler *et al.*, 1990; Hartmann, 1999). ITD and ILD are naturally nonzero  
72 whenever the source is not on the median plane, so that the ears lay at different distances  
73 from the source.

74 The only (rare) cases of median-plane sources generating nonzero ITD/ILD are species  
75 characterized by asymmetrically positioned ears, e.g., the barn owl (Keller *et al.*, 1998).  
76 Other species, including humans, are not very effective at differentiating sound source po-  
77 sitions within the median plane (Butler and Belendiuk, 1969; Heffner and Heffner, 1992;  
78 Van Opstal, 2016), or, more generally, within a “cone of confusion” (e.g. Van Opstal, 2016).  
79 Their (limited) ability at this task must be explained in terms of non-binaural cues (e.g.,  
80 acoustical clues not simply related to a difference between right and left signals). It has been  
81 suggested that sound-localizing animals learn to interpret certain acoustical cues associated  
82 with their anatomy in order to solve this ambiguity (Batteau, 1967; Blauert, 1969; Hart-

83 [mann, 1999](#); [Macpherson and Sabin, 2013](#); [Van Opstal, 2016](#)). Anatomy can be thought of  
84 as a spectral filter (the head-related transfer function, or HRTF), which will change depend-  
85 ing on source position: because the back of our head is different from our face, it interacts  
86 differently with an incoming wave field, which consequently sounds different to our ears.  
87 The HRTF associated with a human skull has been found to provide, in principle, sufficient  
88 information for a source to be localized with fairly high accuracy, independent of the lo-  
89 cation of the source, even when data from only one ear are used ([Catheline et al., 2007](#));  
90 yet, psychoacoustics studies ([Van Opstal, 2016](#)) have shown that the performance of the,  
91 e.g., human ear-brain system at localizing median-plane sources is relatively poor: we are  
92 much more effective at discriminating sources within the horizontal plane. Other terrestrial  
93 species show the same limitations. It has also been found experimentally that humans are  
94 relatively poor at source localization tasks if only one ear is used; subjects with unilateral  
95 hearing loss apparently learn to function with one ear only, but their performance at sound  
96 localization has been found to remain significantly poorer than that of subjects with no  
97 hearing loss ([Agterberg et al., 2011](#); [Van Opstal, 2016](#)). It is inferred that, while humans  
98 and other terrestrial species certainly use HRTF information in sound-localization tasks,  
99 they exploit only a subset of the information provided by the HRTF itself. The consensus  
100 is that the only monaural cues that they are actually capable of using are certain “notches”  
101 of the frequency spectrum perceived by the ears, or “spectral cues,” whose amplitude, and  
102 location along the frequency axis, are controlled by the complex shape of the pinnae and  
103 depend on the position of the source ([Van Opstal, 2016](#), Chapter 7).

104 Cetaceans are also characterized by a salient HRTF (Aroyan, 2001; Au and Fay, 2012;  
105 Supin and Popov, 1993); how and to what extent they make use of it, is still unclear. Simple  
106 physical considerations suggest that anatomical features characterized by relatively strong  
107 density contrasts with respect to the surrounding medium (water) most significantly con-  
108 tribute to characterizing the HRTF, and thus to sound localization. Since the density of  
109 soft tissues found in marine mammal bodies is close to that of water (Norris and Harvey,  
110 1974; Reysenbach de Haan, 1957), it is inferred that features such as the mandible, the  
111 cranium or small air sacs play the most important roles, similar to the external ears of ter-  
112 restrial mammals (Aroyan *et al.*, 1992; Song *et al.*, 2017; Wei *et al.*, 2016). One important  
113 difference in the sound localization performance of terrestrial mammals vs cetaceans is the  
114 latter's ability to localize sound sources within the median plane with a very high accuracy  
115 (Renaud and Popper, 1975). This can be quantified by the minimum audible angle (MAA),  
116 i.e., the minimum angular distance between two sources of sound, still allowing to discrim-  
117 inate them as two different sources. Signals emitted by two sources separated by an angle  
118 smaller than the MAA are perceived as coming from only one source. The MAA changes  
119 depending on the azimuth and elevation of the sources, and on the nature of the emitted sig-  
120 nal. By studying the behavior of live dolphins when exposed to sound coming from different  
121 locations, their MAA in the median plane has been estimated around  $0.7^\circ$  for broadband  
122 clicks. Similar values are observed for sources positioned on the horizontal plane (Au and  
123 Hastings, 2008; Nachtigall, 2016). In comparison, psychoacousticians estimate the MAA of  
124 human subjects at around  $7^\circ$  in the vertical plane, as opposed to only  $\sim 1^\circ$  in the horizontal  
125 one (Nachtigall, 2016), while other terrestrial mammals perform more poorly than humans



126 (Heffner and Heffner, 2016, Figure 3). It can be inferred from these observations that, when  
127 echolocating, dolphins are capable of extracting from their HRTF more information than  
128 terrestrial mammals in sound localization tasks (Branstetter and Mercado III, 2006). The  
129 acoustic environment of cetaceans would indeed favor animals capable of localizing sound,  
130 whether it be emitted or reflected from prey or predators, regardless of their position in  
131 space. Dolphins' MAA grows to  $2.3^{\circ}$ - $3.5^{\circ}$  for narrow-band signals (Au and Hastings, 2008;  
132 Nachtigall, 2016), which do not mimic typical echolocation clicks.

133 This study addresses the question of how a dolphin's head inner anatomy may contribute  
134 to sound localization, and in particular to echo-localization, by means of a suite of physical  
135 acoustics experiments conducted on one skull specimen of short-beaked common dolphin  
136 (*Delphinus Delphis*). The HRTF of the short-beaked common dolphin has so far only been  
137 addressed in a limited number of studies. Most of our previous knowledge results from  
138 numerical models; Krysl and Cranford carried out vibroacoustic simulations on a CT scan of  
139 a full head of a common dolphin showing single-frequency HRTFs for 5.6 kHz, 22.5 kHz and  
140 38 kHz (Krysl and Cranford, 2016). In their work, single-frequency HRTFs were equivalent  
141 to amplitude or intensity variations with respect to the source position. They simulated  
142 sound pressure levels at two virtual positions on the surface of the TPCs (one dorsal, one  
143 ventral), caused by monochromatic plane waves traveling along a suite of different azimuths  
144 and elevation angles. Strong variations in modeled data were found between the dorsal and  
145 the ventral receiver positions. The spatial pattern of the HRTFs was also found to depend  
146 strongly on the source frequency used. The results were not conclusive as to what extent the  
147 calculated asymmetries in the receiving pressure pattern were due to the inexact placement

148 of the receivers, versus asymmetries in the specimen’s anatomy. Another, similarly minded  
149 study (Aroyan, 2001) showed that the mandible and its surrounding fats focus acoustic  
150 waves toward the TPC, therefore playing an important role in sound conduction through  
151 the head. Receptivity patterns at two virtual ear positions showed high asymmetry and  
152 complexity, and varied significantly depending on which parts of the head were simulated.

153 In our experimental study, we attempt to evaluate the specific contribution of bone con-  
154 duction to sound localization. To this goal, we conduct experiments on a dolphin’s skull  
155 specimen immersed in water, in the absence of soft tissues. It has been suggested (e.g.,  
156 Aroyan, 2001; McCormick *et al.*, 1970; Song *et al.*, 2016; Wei *et al.*, 2018) that bone conduc-  
157 tion affects significantly the dolphin’s HRTF. Bone conduction seems to be a driving force for  
158 hearing mechanisms in baleen whales (Cranford and Krysl, 2015). Despite the isolation of  
159 the skull from the TPC, Ketten also theorized some influence of bone conduction in dolphin  
160 hearing, i.e., on sound reception pathways (Ketten, 2000). Cranford *et al.* suggested that  
161 a complex wave propagation pattern including flexural waves along the mandible bone of a  
162 Cuvier’s beaked whale might contribute to the received pressure at its two ears (Cranford  
163 *et al.*, 2008). In summary, while it is likely that the mandible plays an important role in this  
164 context, the exact nature of its contribution to sound localization remains to be determined.

165 Using accelerometers glued to the pan bone, we measure elastic waves traveling through  
166 a mandible specimen immersed in water; we record the signal generated by different sound  
167 sources, positioned at many different locations within a large water tank; we measure the  
168 ITD and ILD (*binaural cues*) resulting from such recordings and estimate their potential  
169 performance as source-localization cues. Finally, we study in much detail how the wave-

170 form of the recorded signal depends on source position, and use a correlation-based method  
 171 (known in physical acoustics as “acoustic time reversal” (Fink *et al.*, 2000)) to numerically  
 172 reconstruct the location of sources via full-waveform data. In this endeavour, we take the  
 173 standpoint of physicists, attempting to quantify the information carried by our data, inde-  
 174 pendent of how these would be received and processed by the auditory system of a living  
 175 organism. Our main goal is to contribute some new information on the potential contribution  
 176 of certain features of dolphins’ skulls (in particular, their mandible) to sound localization  
 177 performance.

## 178 II. EXPERIMENTAL SETUP & DATA ACQUISITION

179 All our experiments are conducted on the skull (cranium and mandible) of a male adult  
 180 short-beaked common dolphin, shown in Figure 1a. The skull is  $\sim 50$  cm long and  $\sim 20$  cm  
 181 wide. The specimen was acquired on loan from the French National Museum of Natural His-

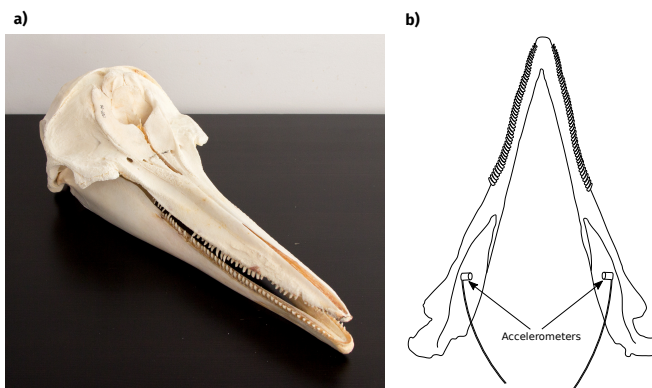


FIG. 1. (Color online) (a) Photograph of the sample (cranium and mandible) used in this study. (b) Sketch of the mandible and the accelerometers glued to it. The accelerometers are approximately 11 cm apart.

183 tory (Muséum National d'Histoire Naturelle, Paris, France), inventory number 1989-06 from  
184 the Collection of Comparative Anatomy (Collection d'Anatomie Comparée - Mammifères  
185 et Oiseaux). Two miniature piezoelectric charge accelerometers (*Brüel & Kjaer* Type 4374)  
186 are glued to the inside of the pan bone by a common cyano-acrylate adhesive as shown in  
187 Figure 1b. Aroyan ([Aroyan, 1996](#)) showed via numerical simulations that acoustic waves  
188 entered the head forward of the pan bone, propagated through the outer mandibular fat,  
189 to and through the pan bone, continued along intramandibular fats, and converged at the  
190 TPC. Measuring the vibration of the pan bone should, therefore, be representative of the  
191 sound received at the ears. These sensors weigh 0.75 g and are characterized by a flat  
192 frequency response curve in the frequency range of interest. They are both waterproofed  
193 by applying a layer of flexible adhesive. Measurements are conducted in a water tank (6  
194 meters in width, 12 meters in length and 3 meters in depth) filled with chlorinated water  
195 kept at the temperature of  $\sim 12^\circ\text{C}$  throughout the duration of the experiment; the specimen  
196 is immersed in the water, centered in depth and in width. Both cranium and mandible  
197 are independently suspended and aligned with each other according to the real anatomy.  
198 The geometry of the experimental setup is shown in Figure 2. Let us take the midpoint of  
199 the segment defined by the accelerometer positions as the origin of a Cartesian reference  
200 frame; let the  $y$ -axis be defined by the accelerometer positions, while the  $x$ -axis is identified  
201 by the tip of the mandible and the origin. The horizontal plane consequently lies roughly  
202 on the tooth lines. A broadband marine transducer (*Airmar* B75L) with an active area of  
203  $9.6\text{ cm}^2$  and a transmitting voltage response of around 155 dB (re  $1\mu\text{Pa}$  per volt at 1 m)  
204 throughout the used frequency range is placed at a distance of 2 meters away from the origin  
205

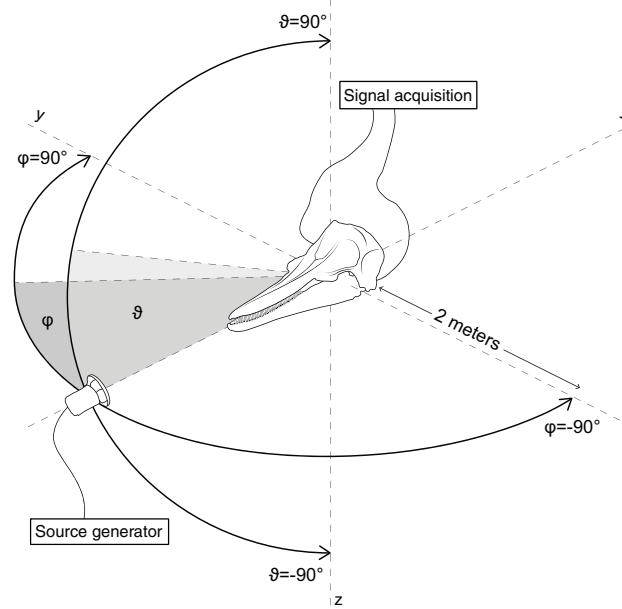


FIG. 2. Sketch of the experimental setup. The sound source moves along two half circles, either in the median or horizontal plane at a distance of 2 m from the origin.

206 in front of the skull along the  $x$ -axis. The skull is then rotated around either the  $z$ -axis,  
 207 which corresponds to an angular movement of the transducer in the horizontal plane (i.e.  
 208 constant source elevation  $\vartheta = 0^\circ$ ), while azimuth  $\varphi$  changes from  $-90^\circ$  nearest the left “ear”  
 209 to  $+90^\circ$  closest to the right “ear”), or around the  $y$ -axis, which corresponds to an angular  
 210 movement of the transducer in the median plane (i.e. constant  $\varphi = 0^\circ$ , while  $\vartheta$  changes  
 211 from  $-90^\circ$  directly below to  $+90^\circ$  directly above the origin). Data are recorded first for a  
 212 discrete set of source azimuths on the skull’s horizontal plane, spaced  $1^\circ$  from one another,  
 213 from  $\varphi=-90^\circ$  to  $\varphi=+90^\circ$ , and then for a discrete set of source elevations on the vertical plane,  
 214 again  $1^\circ$  from one another, from  $\vartheta=-90^\circ$  to  $\vartheta=+90^\circ$ . For each source location, the transducer  
 215 emits two different source signals which are digitally generated through a desktop computer  
 216 and recorded and processed separately. Each source signal is amplified by 30 dB through a

217 home-made power supply resulting in an emitted sound level of about 185 dB (re  $1\mu\text{Pa}$  per  
 218 volt at 1 m).

219 The source signals are

220 1. a sinusoidal burst, i.e.

$$c(t) = \sin[\phi_0 + 2\pi ft] w(t), \quad (1)$$

221 where  $\phi_0$  denotes the initial phase of the signal,  $f=45$  kHz and  $w(t)$  is a Tukey window  
 222 ([Harris, 1978](#)), which has a total duration of  $100\ \mu\text{s}$  and tapers the first and the last  
 223  $15\ \mu\text{s}$  of the signals, to ensure their smooth on- and offset;

224 2. a linear chirp

$$c(t) = \sin\left[\varphi_0 + 2\pi\left(f_0 t + \frac{k}{2} t^2\right)\right] w(t), \quad (2)$$

225 with minimum frequency  $f_0=45$  kHz, maximum frequency  $f_1=55$  kHz, chirpiness (i.e.,  
 226 rate of frequency change across the chirp)  $k=\frac{f_1-f_0}{t}$ , and  $w(t)$  the same Tukey window  
 227 as above.

228 The sampling frequency for both signals is 2 MHz. The signals and their normalized fre-  
 229 quency spectra are shown in [Figure 3](#). Peak frequencies are 45 kHz (sinusoid) an 49 kHz  
 230 (chirp), and 3dB bandwidths are 6 kHz (sinusoid) and 10 kHz (chirp). Since our skull  
 232 specimen belongs to a short-beaked common dolphin, the duration and peak frequency of  
 233 source signals are chosen to be in the range of that of echolocation clicks of common dol-  
 234 phins ([Richardson et al., 2013](#); [Soldevilla et al., 2008](#)). However, our synthetic signals have  
 235 a smaller bandwidth since we cannot emit broadband clicks due to the specifications of the  
 236 equipment. While the sinusoidal burst is used for the investigation of binaural and monau-

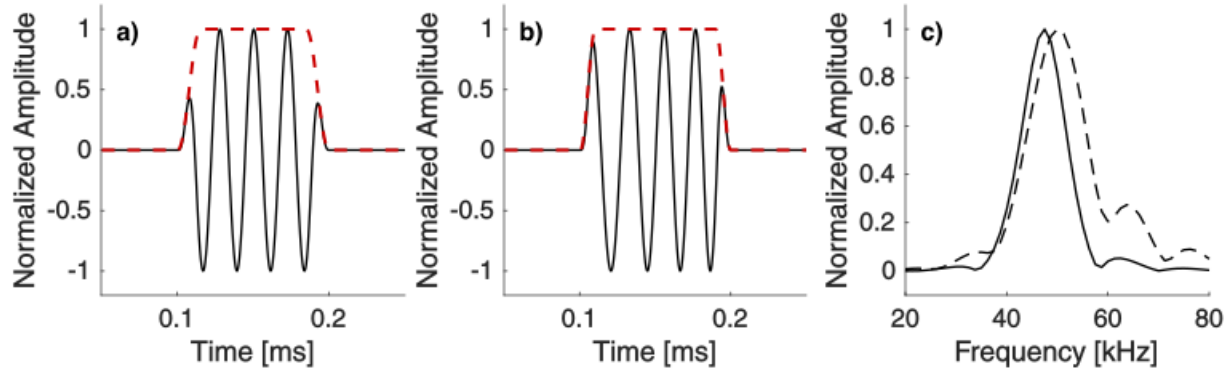


FIG. 3. (Color online) Source signals and their frequency spectra. The Tukey window function used to taper the signals in time is also shown, and is denoted by a dashed red line. a) Sinusoidal burst with a duration of  $100 \mu\text{s}$  and frequency of 45 kHz. b) Narrowband chirp with a duration of  $100 \mu\text{s}$  and a frequency range of 45 to 55 kHz. c) Normalized spectrum of the sinusoidal burst (solid line) and the chirp (dashed line).

237 ral cues in both planes, the chirp is solely used for monaural cues. The accelerometers are  
 238 calibrated to synchronously measure the acceleration of the pan bone on each side of the  
 239 mandible. At each realization of the experiment, they record for  $800 \mu\text{s}$  at a sampling rate  
 240 of 2 MHz using a 16 bit analog-to-digital converter (*National Instruments PXIe-6366*). The  
 241 duration of our recordings coincides with the time needed for an acoustic wave to travel  
 242 1.2 m in water, which means that signals reflected from the sides, bottom or surface of the  
 243 tank are well separated and can be easily identified; we systematically cut our data so that  
 244 such signals are not taken into account. All recordings are Butterworth bandpass filtered,  
 245 with cutoff frequencies of 40 kHz and 60 kHz to further reduce unwanted noise.

246 The entire experiment was repeated three times, including setup and wiring, in order to  
 247 check consistency and minimize the effect of random errors. All measurements presented

248 in the following are obtained by averaging the outcomes of the three experiments, for each  
 249 combination of source and receiver positions. The associated standard deviation is used  
 250 as an estimate of measure uncertainty. Throughout this study, we dub “direct” signal  
 251 the waveform defined by Equations 1 or 2, as it is recorded at the accelerometers after  
 252 having propagated through water and bone, and being accordingly attenuated. We dub  
 253 “reverberated” the signal recorded after the direct signal, refracted, reflected, diffracted by  
 254 and through bone tissue. The so defined reverberated signal is, alone, always longer than  
 255 Au et al.’s estimate of  $\sim 250 \mu\text{s}$  for a dolphin’s integration time (Au *et al.*, 1988). For the  
 256 sake of simplicity, we neglect reverberations occurring before the end of the direct signal;  
 257 visual inspection (e.g., Figure 4) shows that their effect is indeed minor, compared to the  
 258 complex, relatively long coda.

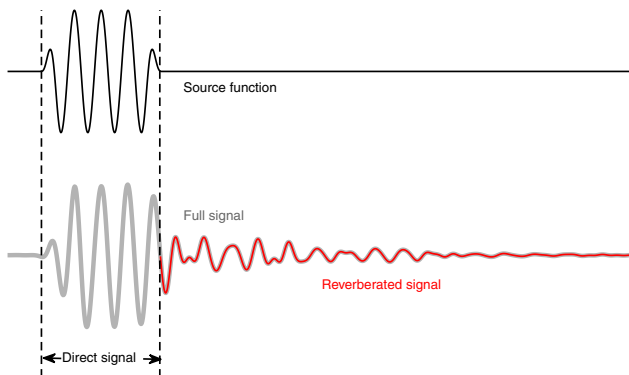


FIG. 4. (Color online) Examples of (top) signal as emitted at the source; (bottom) the same signal, as recorded by one receiver. The recorded trace consists of “direct” (grey) and “reverberated” (red) signals, defined in Section II.

259

260



### 261 III. ITD- AND ILD-BASED SOURCE LOCALIZATION

262 We define ITD as the onset time of the direct signal measured at the left accelerometer  
 263 minus the onset time of the same signal, measured at the right accelerometer. We measure  
 264 the ITD associated to all our recordings of horizontal- and median-plane sinusoidal sources  
 265 (Equation 1). This is done by means of a matlab routine that identifies the shape of the  
 266 source signal in the recorded signal through cross correlation. We show in Fig. 5 the results  
 267 of this exercise, as functions of source azimuth (if the source is on the horizontal plane) or  
 268 elevation (if on the vertical plane). For median-plane sources, the ITD should be approx-  
 269 imately zero; measured values of ITD accordingly never exceed  $6 \mu\text{s}$ , corresponding to an  
 270 error of 0.9 cm in space. For horizontal-plane sources, by simple geometrical considerations  
 271 and neglecting HRTF-related diffraction effects (which is reasonable given the absence of  
 272 soft tissues in our experiment), ITD is expected to approximately coincide with

$$\text{ITD}(\varphi) = (a/c) \sin(\varphi), \quad (3)$$

273 where  $a$  is inter-receiver distance and  $c$  the speed of sound in water. Again, Figure 5 shows a  
 274 good agreement between our data and theoretical predictions. Importantly, our measure of  
 275 ITD should not be taken as an estimate of ITD as perceived by live dolphins, which might  
 276 be significantly affected by the presence of soft tissues and other anatomical features.

278 We define ILD as the ratio of the maximum amplitudes (Figure 6) of the direct signal as  
 279 recorded by left vs. right receivers, in dB, i.e.,

$$\text{ILD}(\varphi, \vartheta) = 20 \log_{10} \left\{ \frac{\max [s(\vartheta, \varphi, \mathbf{r}_L, t)]}{\max [s(\vartheta, \varphi, \mathbf{r}_R, t)]} \right\} \quad [\text{dB}], \quad (4)$$

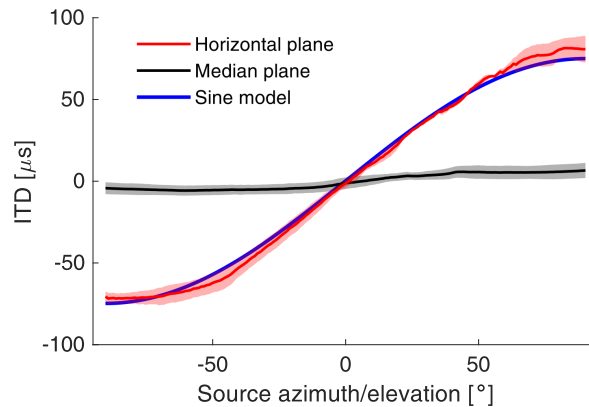


FIG. 5. (Color online) Measured ITD from our binaural recordings of sources deployed on the horizontal plane (red solid line), as a function of source azimuth, and on the vertical plane (black solid line), as a function of source elevation. Color-shaded areas around each solid line denote standard deviation. Expected horizontal-plane ITD based on the theoretical model of Equation 3 is shown as a blue solid line.

280 where, for the sake of clarity, the signal  $s$  is explicitly written as a function of source azimuth  
 281 and elevation, and receiver position (its only possible values being  $\mathbf{r}_L$ ,  $\mathbf{r}_R$  for left and right  
 282 receiver, respectively). Although other definitions of ILD have been proposed, e.g. in the  
 283 field of robotics (Youssef *et al.*, 2012), Equation 4 has been used in similar bioacoustic  
 284 research (Moore and Au, 1975) and can be interpreted similarly to peak values of electro-  
 285 physiological audiograms (Mulsow *et al.*, 2014; Supin and Popov, 1993). We cannot relate  
 287 our ILD observations to a simple theoretical model as for the ITD, because of (i) the inherent  
 288 complexity of waveforms resulting from multiple reverberations within the pan bone, and  
 289 (ii) our neglect of anatomical features, other than the mandible and skull bones, including  
 290 cranial air sacks, the albuminous foam (which separates the middle and inner ear from

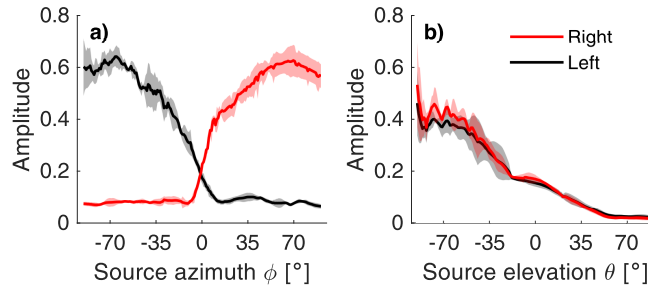


FIG. 6. (Color online) Maximum amplitudes recorded at the left (black) and right (red) receivers of sources deployed on (a) the horizontal plane, as a function of source azimuth, and (b) the vertical plane, as a function of source elevation. Color-shaded areas around each solid line denote standard deviation.

291 the skull) and acoustically functional fats, that are likely to contribute to ILD (Ketten,  
 292 1992; Supin and Popov, 1993) and, interestingly, introduce significant dispersion (Aroyan,  
 293 2001). Also, because our setup does not account for such complexity, our data cannot be  
 294 directly compared to experimental data or realistic numerical ILD models. Figure 7 shows  
 295 our measures of ILD, derived from waveform data via Equation 4, as a function of source  
 296 azimuth and elevation.

298 As expected, ILD values associated with median-plane sources are close to 0, with fluc-  
 299 tuations of less than 2 dB. For horizontal-plane sources, the ILD ranges between 18 dB and  
 300 -18 dB, changing most rapidly directly in front of the dolphin's beak, at  $\varphi$  between  $-10^\circ$  and  
 301  $10^\circ$ . In this range of  $\varphi$ , ILD decreases from 13 dB down to -12 dB, losing more than 1 dB  
 302 per degree. This is an effect of sound shadowing by bone tissues, as the receiver at  $x_L$  loses  
 303 direct acoustic sight of the sound source when this is rotated to the opposite side of the

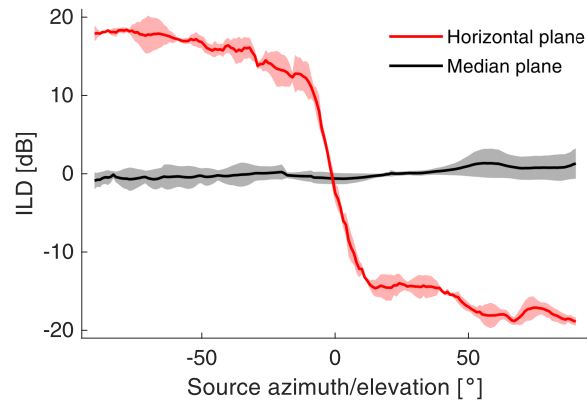


FIG. 7. (Color online) ILD in both planes. Mean ILD (solid lines) and their standard deviation (color shaded areas) of three independent measurements are shown in red (horizontal plane) and black (median plane).

304 mandible. At larger, positive or negative, azimuths, the ILD grows less rapidly, at a rate of  
 305 less than 1 dB per degree, and fluctuations (standard deviation) up to  $\pm 2$  dB.

306 The results in Figures 5 through 7 are not new or surprising per se, but confirm some  
 307 simple, well known properties of all binaural auditory systems. Importantly, the left-right  
 308 symmetries of our data and the fit between data and a simple ITD model confirm that our  
 309 setup is correct, and adequate to the applications that follow.

#### 310 IV. CORRELATION-BASED SOURCE LOCALIZATION

311 Waves that interact with a complex HRTF carry a great wealth of information, that could  
 312 in principle be exploited to localize their sources. Both binaural and monaural cues discussed  
 313 so far only exploit a small portion of such information. While it has been established that hu-  
 314 mans and other terrestrial species localize via those cues alone, the echolocation performance

315 observed in dolphins suggests that their auditory system might include a more sophisticated  
 316 localization mechanism. We implement a simple algorithm to localize sources, based on  
 317 the time-reversal concept developed by Mathias Fink and co-workers (e.g., [Catheline et al.,](#)  
 318 [2007](#); [Fink et al., 2000](#)).

### 319 **A. Accuracy of source localization by correlation**

320 We conduct a “time-reversal” exercise based on the theoretical formulation developed in  
 321 the appendix. Specifically, we implement the right-hand side of Equation [A.5](#) and study its  
 322 effectiveness as a source-localization algorithm. As explained in detail in the appendix, in  
 323 the context of echolocation “time reversal” as defined e.g. by Catheline et al. ([Catheline](#)  
 324 [et al., 2007](#)) is equivalent to a simple correlation of each newly perceived signal with a library  
 325 of echoes previously heard and “stored.” Accordingly, pairs of traces  $s(\mathbf{r}_R, \mathbf{r}_A, t)$ ,  $s(\mathbf{r}_R, \mathbf{r}_B, t)$   
 326 recorded at  $\mathbf{r}_R$  as described in Section [II](#), are cross-correlated to one another, for all possible  
 327 pairs of source locations  $\mathbf{r}_A$ ,  $\mathbf{r}_B$ . The same is done for traces recorded at  $\mathbf{r}_L$ . As a result,  
 328 for each source location  $\mathbf{r}_B$ , we obtain the correlation between the corresponding recorded  
 329 signal and the signal associated to all other possible sources ( $\mathbf{r}_A$ ). Because it is closely  
 330 related to how sharply a time-reversed wave field would focus at  $\mathbf{r}_B$  (see appendix), we dub  
 331 it “focusing function”. Since, in this study, we are looking at sources on the horizontal and  
 332 median planes only, the focusing function depends on either  $\vartheta$  or  $\varphi$  only; by definition, it is  
 333 exactly 1 when both  $\vartheta$  and  $\varphi$  are the same as those of the actual source.

334 For the sake of simplicity (and speed), cross correlation is implemented by first shifting  
 335 each pair of signals to have zero lag, and then calculating the correlation between the shifted

336 traces. Intensity differences between the two correlated signals are also irrelevant, as the  
 337 convolution product is normalized so that the auto-correlation at zero lag equals 1.

338 We next visualize how well a source is localized by our algorithm as a function of its true  
 339 location. This is shown in Figure 8 through 11 where the horizontal and vertical axes of  
 340 each plot correspond to the azimuth  $\varphi_0$  or elevation  $\vartheta_0$  of the true source and of all recorded  
 341 sources  $(\varphi_i, \vartheta_i)$ . Specifically, focusing functions obtained based on the chirp-like source in the  
 342 median plane are plotted in Figure 8, while Figure 9 shows the corresponding results for the  
 343 sinusoidal source. By definition, values on the diagonal of all panels in both figures are 1; near  
 344 the diagonal, correlations decrease monotonously in all panels; some relevant fluctuations  
 345 are then observed in both figures for  $\vartheta_i$  far from  $\vartheta_0$  when both direct and reverberated signals  
 346 are correlated, but not when the reverberated signal alone is considered. In the latter case,  
 347 the focusing function is much sharper, particularly in the  $-50^\circ$  to  $20^\circ$  elevation range, and  
 348 its sharpness does not seem to depend on source elevation  $\vartheta_0$ .

351 To study how the resolution of our algorithm depends on the true source position in the  
 352 median plane, we visualize (Figure 10a for the chirp-like source function, Figure 10b for the  
 353 sinusoidal source function) the increment in  $\vartheta$  needed for the focusing function to decrease to  
 354 70% of its maximum, i.e. the -3 dB width of the focusing function, which is a rule-of-thumb  
 355 criterion frequently used in time-reversal acoustics (Catheline *et al.*, 2007; Ing *et al.*, 2005;  
 356 Kim *et al.*, 2003). The smaller the value of the -3 dB width, the higher the resolution, and  
 357 the performance that can be expected in identifying the true source location. The value  
 358 of 3 dB is of no particular physical or biological significance: it is only chosen in analogy  
 359 with the mentioned studies. This is adequate to our goals, as we are not attempting to

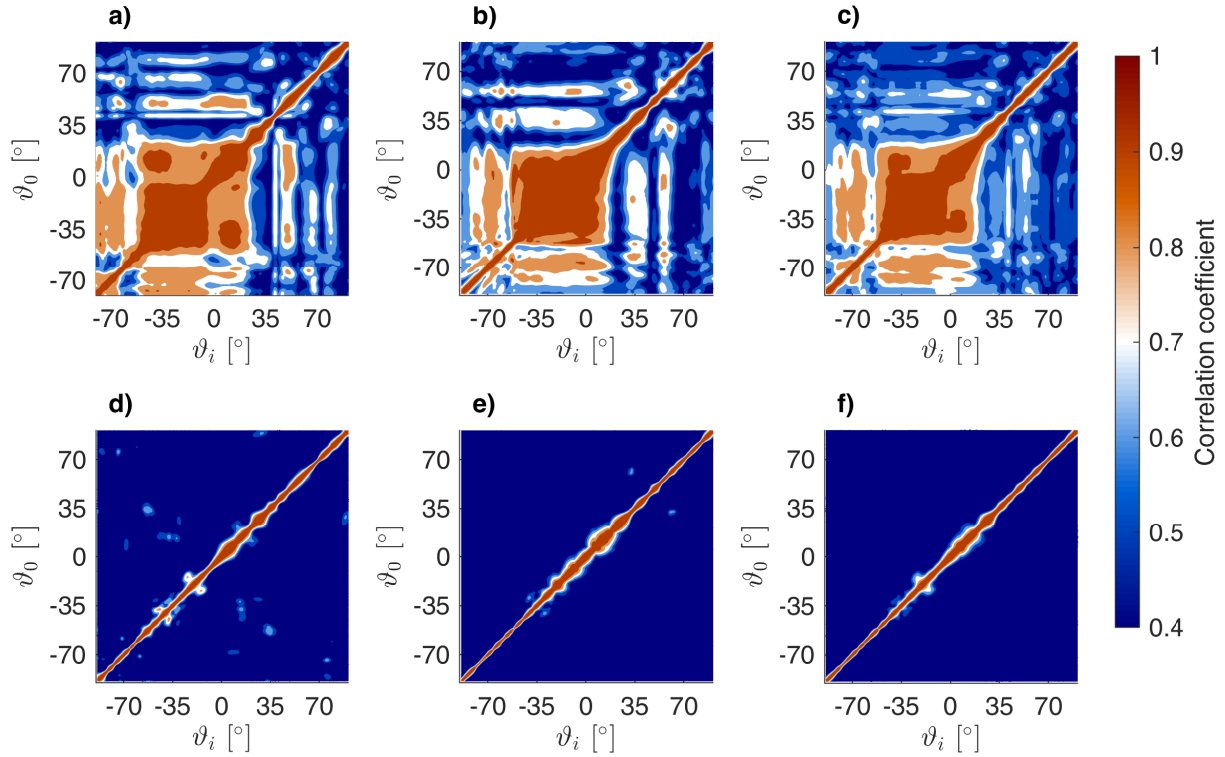


FIG. 8. (Color online) Focusing functions in the median plane using the chirp-like source function as determined from the entire waveform, recorded by the (a) left, (b) right, and (c) both (sum of (a) and (b)) accelerometers, and from the reverberated waveform alone, again at (d) left, (e) right, and (f) both accelerometers. Each row of a given panel shows, accordingly, the maximum cross correlation value between the signal associated with one particular source (defined by its elevation  $\vartheta_0$ ), and those of all other sources (elevations  $\vartheta_i$  on the horizontal axis).

360 reproduce absolute, observed MAA values, but rather to estimate the relative changes in  
 362 the resolution in source localization. Figures 8 through 10 show that direct signal alone does  
 363 not provide sufficient information to discriminate sources in the  $-50^\circ$  to  $20^\circ$  elevation range;  
 364 on the contrary, it obscures the information contained in the reverberated signal, which,

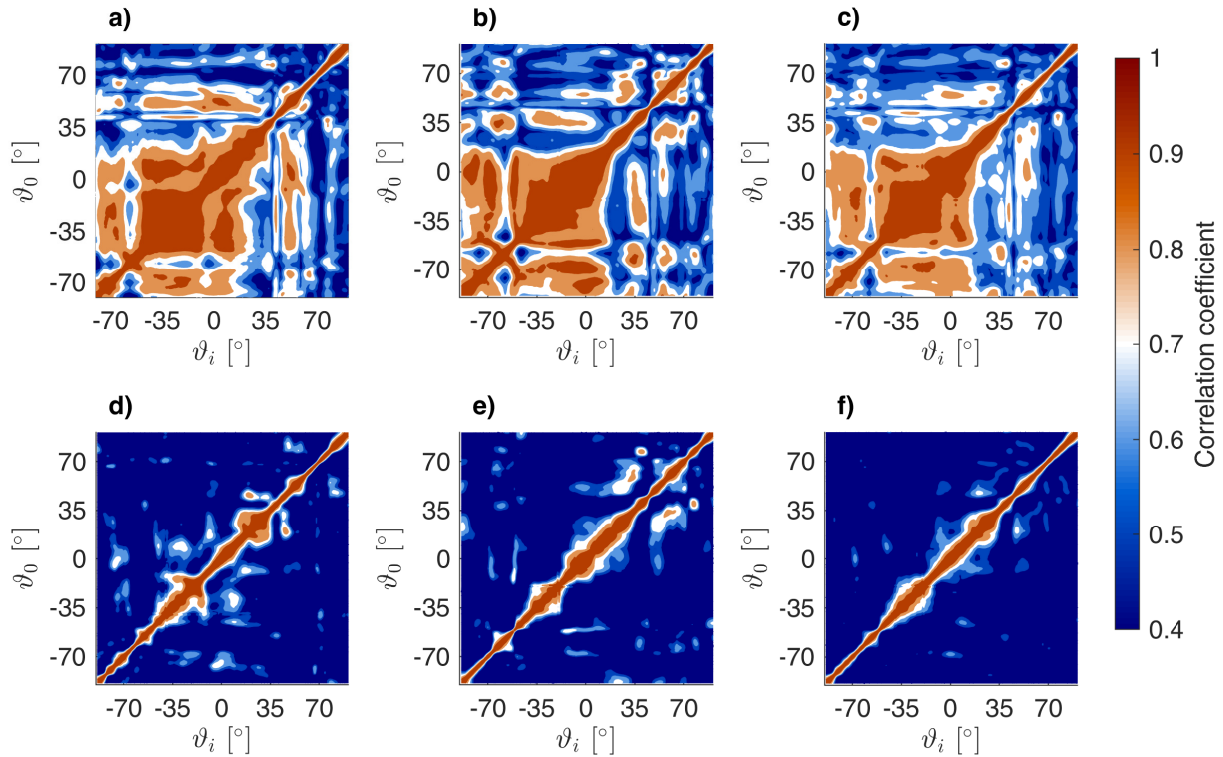


FIG. 9. (Color online) Focusing functions in the median plane using the sinusoidal source function.

Panels are structured the same way as in Figure 8.

365 if used by itself, actually results in much sharper focusing functions. It is apparent from  
 366 our results that our algorithm achieves approximately equal accuracy for monochromatic  
 367 vs multi-frequency signals (Figure 8&9). Figure 10 shows that localization of a sinusoidal  
 368 source affords slightly lower resolution (larger -3 dB widths) throughout all elevations. As  
 369 to be expected, widening the frequency band of the source increases the resolution of this  
 370 algorithm. Similar inferences can be made based on the focusing functions obtained from  
 372 horizontal-plane sources, which are shown in Figure 11. In this case, the resolution highly  
 373 benefits from analyzing the reverberated signal alone, if the source is on the same side of the



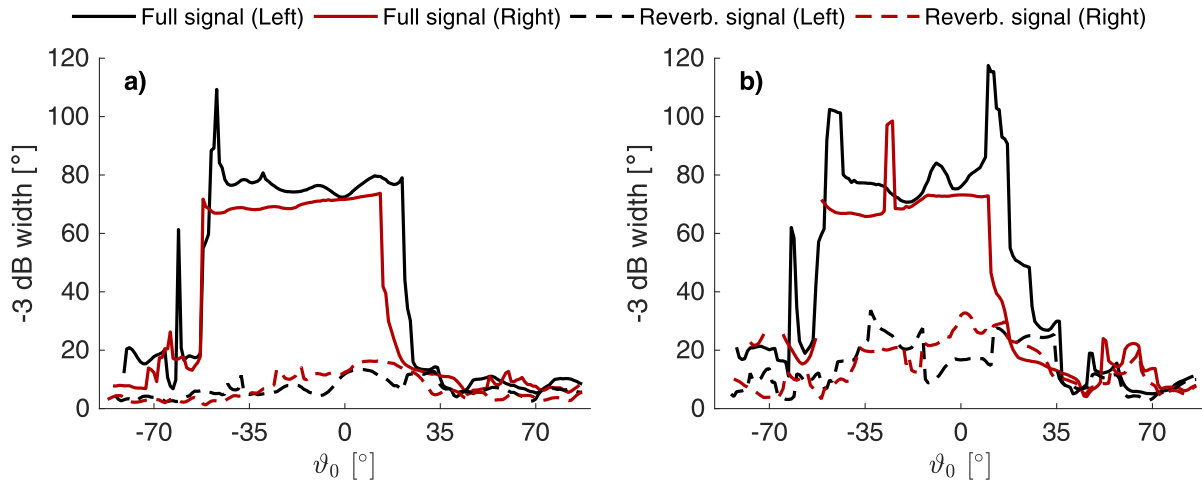


FIG. 10. (Color online) -3 dB widths of the focusing functions in the median plane using a) the chirp-like source function and b) the sinusoidal source function.

374 skull as the respective receiver. Interestingly, the -3dB width is similar to that extrapolated  
 375 from Figures 8&9, i.e. our algorithm is about equally sensitive to changes in azimuth vs  
 376 elevation of the source.

## 377 V. SUMMARY AND CONCLUSIONS

378 We have developed a source localization algorithm (Section IV) based on the cross cor-  
 379 relation of an observed signal with a library of known signals, each corresponding to a  
 380 different source location. We have implemented the algorithm in the context of a biosonar  
 381 application (Equation A.5 and related discussion), and “source” should be interpreted here  
 382 as synonymous with biosonar “target” (or “secondary” source). We have substantiated  
 383 our source-localization metric from a theoretical standpoint, by drawing an analogy be-  
 384 tween cross correlation and the theory of acoustic time reversal. We have evaluated the

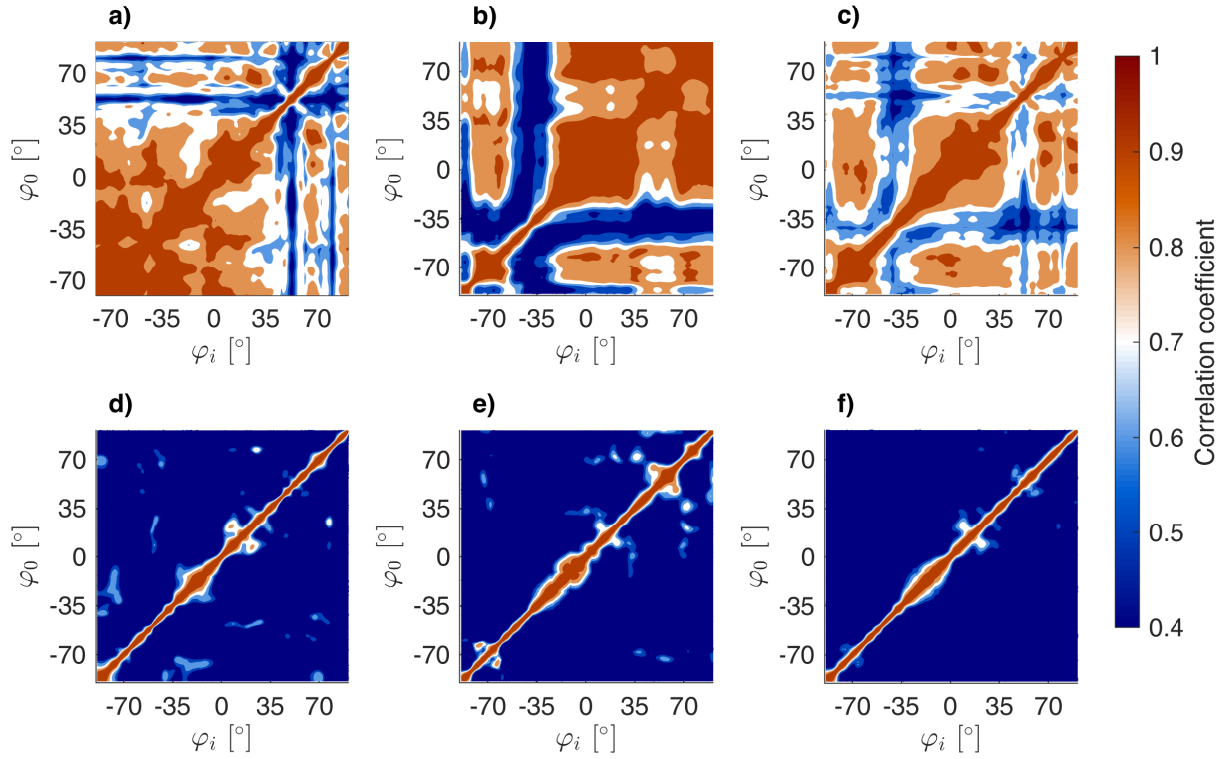


FIG. 11. (Color online) Focusing functions in the horizontal plane, i.e.  $\varphi$  defines the azimuth, using the chirp-like source function. Panels are structured the same way as in Figure 8.

385 performance of our algorithm, as applied to a particular setup, via a suite of experiments.  
 386 The setup consists of two accelerometers installed on the mandible of a dolphin skull, fully  
 387 immersed in a large water tank, and recording signals similar to a dolphin’s echolocation  
 388 “clicks.”

389 We quantify the performance of our algorithm via the width of the the focusing function,  
 390 or, in other words, the rate at which correlation decreases, as an observed signal is compared  
 391 with library signals associated with sources increasingly far from the true one. We find  
 392 that this width is significantly reduced (the rate of correlation loss is accelerated) when

393 the direct signal, which is simply an attenuated version of the original chirp/sinusoidal  
394 burst, is subtracted from the recorded waveform before cross correlation. This way, only  
395 the reverberated coda, most strongly affected by the shape and properties of the skull, is  
396 actually employed in localization: localizing by reverberated signal alone (rather than the  
397 entire wavetrain) sharpens source resolution.

398 The spatial accuracy of source localization by dolphins has been observed, through be-  
399 havioral experiments, to be equally accurate independent of source azimuth and elevation,  
400 i.e., it has approximately constant resolution over the entire solid angle (Nachtigall, 2016).  
401 This property of dolphins is counter-intuitive, if one considers that humans and other species  
402 have presumably evolved pinnae to help determine the elevation of sound sources (Section I),  
403 while cetaceans have actually lost them. We infer that, to achieve such performance, the  
404 dolphin’s auditory system might make use of a unique, and currently unknown localization  
405 tool, particularly effective for sources in the median plane or along the “cone of confusion.”  
406 Our results do not directly constrain the nature of the sound localization system imple-  
407 mented in a dolphin’s brain; yet, they do show that signal reverberated within the dolphin’s  
408 skull (with the mandible playing the most prominent role) varies significantly as a function  
409 of (that is to say, is very sensitive to) source location. The “direct” signal, i.e. signal without  
410 reverberation, appears to be much less sensitive to source location. We have shown that re-  
411 verberated signal contains sufficient information to discriminate median-plane sources, and  
412 that this could be achieved by simply cross-correlating any newly perceived sound with a  
413 library of previously recorded data.

414 While our model shares with dolphins some relevant features, we are hardly reproducing  
415 the signals that would be perceived by actual, live specimens. Importantly, we conduct our  
416 experiments on the skull alone, neglecting the effects of muscles, fats, and other soft tissues.  
417 While we plan to surpass this limitation in our future work, at this point we consider it to  
418 be partly justified, first, by the fact that acoustic waves propagate through soft tissues at  
419 about the same speed as through water (e.g., [Gray and Rogers, 2017](#); [Soldevilla \*et al.\*, 2005](#)),  
420 which limits wave-propagation effects. Mandible and skull, on the contrary, provide a strong  
421 wave speed contrast (a factor of about two) resulting in significant diffraction, reverberation,  
422 etc. Cranford et al. show that the amplitude of perceived signals is significantly affected by  
423 the anatomical features of soft tissues in the head ([Cranford \*et al.\*, 2008](#)), but we speculate  
424 that phase, more than amplitude, is relevant to source localization (correlation being mostly  
425 sensitive to phase). Secondly, the neglect of soft tissues allows us to isolate the specific effects  
426 of bone-conducted waves, before additional experiments are conducted on whole heads. In  
427 future studies, the issue of a dolphin’s resolution of the fine features of acoustic signals  
428 with respect to time will also need to be addressed, taking into account e.g. the concept of  
429 cochlear integration time as defined by Au et al. ([Au \*et al.\*, 1988](#)). Finally, we envisage to  
430 extend our analysis to a broader frequency range, whether by numerical modeling, or new  
431 experiments relying on better, currently unavailable hardware.

432 In summary, our results indicate that, within a good approximation, a one-to-one cor-  
433 respondence exists between the waveform of the bone-conducted, reverberated coda as  
434 recorded at a dolphin’s ear locations, and the locations of the source (or, in principle,  
435 the reflecting target) that originally generated (or reflected) the signal. While we have no

436 knowledge of how such information might be processed and exploited by the brain, we spec-  
437 ulate that bone-conducted, reverberated sound could contribute to explaining the peculiar,  
438 poorly understood accuracy of sound localization in odontocete cetaceans.

#### 439 **ACKNOWLEDGMENTS**

440 This project has received funding from the European Union’s Horizon 2020 research  
441 and innovation programme under the Marie Skłodowska-Curie grant agreement No 641943  
442 (ITN WAVES). We are grateful to Christine Lefèvre, and to the Collection of Comparative  
443 Anatomy at the French National Museum of Natural History, for providing the skull speci-  
444 men on which all our experiments were conducted. We are likewise grateful to the editor W.  
445 W. L. Au and two anonymous reviewers, for their insightful comments to our manuscript.

446 **APPENDIX: TIME-REVERSAL THEORY AND CROSS CORRELATION**

447 It is well known that, if attenuation is neglected, the imaginary part ( $\Im$ ) of the acoustic  
 448 Green’s function (i.e., impulse response)  $G$  associated with a source at  $\mathbf{r}_A$  and a receiver at  
 449  $\mathbf{r}_B$  (or vice-versa) can be obtained by the frequency-domain relationship

$$\frac{\rho c}{\omega} \Im[G(\mathbf{r}_A, \mathbf{r}_B)] = - \int_{\partial V} d^2\mathbf{r} [G^*(\mathbf{r}, \mathbf{r}_B)G(\mathbf{r}, \mathbf{r}_A)], \quad (\text{A.1})$$

450 (e.g., [Boschi and Weemstra, 2015](#), Equation (103)), where  $G$  is the 3-D Green’s function,  $\partial V$   
 451 is an arbitrary closed surface surrounding  $\mathbf{r}_A$  and  $\mathbf{r}_B$ , and  $\rho$ ,  $c$ ,  $\omega$  denote density, speed of  
 452 sound and frequency, respectively.  $*$  stands for complex conjugation, so that the integrand  
 453 at the right-hand side of Equation [A.1](#) is the Fourier transform of the time-domain cross  
 454 correlation of  $G(\mathbf{r}, \mathbf{r}_B, t)$  and  $G(\mathbf{r}, \mathbf{r}_A, t)$ .

455 Think now of  $\mathbf{r}_B$  as the location of an acoustic source (e.g., [Boschi and Weemstra, 2015](#));  
 456  $G(\mathbf{r}, \mathbf{r}_B, \omega)$  is the Fourier-transform of an impulse generated at  $\mathbf{r}_B$  and recorded by a receiver  
 457 at  $\mathbf{r}$ ;  $G^*(\mathbf{r}, \mathbf{r}_B, \omega)$  is the Fourier transform of the same signal, reversed in time. Imagine that  
 458 the time-reversed signal be then emitted from  $\mathbf{r}$  and recorded at another point  $\mathbf{r}_A$ : this  
 459 amounts to convolving (in the frequency domain, multiplying) the time-reversed signal with  
 460 the Green’s function  $G(\mathbf{r}_A, \mathbf{r}, \omega)$ . Eq. [\(A.1\)](#) then shows that by repeating time reversal and  
 461 propagation (“backward in time”) for all points  $\mathbf{r}$  on  $\partial V$ , and summing all the resulting  
 462 traces at  $\mathbf{r}_B$ , the imaginary part of the Green’s function between  $\mathbf{r}_B$  and  $\mathbf{r}_A$  is obtained.  
 463 Note that the imaginary part of the frequency-domain  $G$  coincides, in the time domain,  
 464 with the inverse Fourier transform

$$F^{-1} \{ \Im [G(\mathbf{r}_A, \mathbf{r}_B, \omega)] \} = G(\mathbf{r}_A, \mathbf{r}_B, -t) - G(\mathbf{r}_A, \mathbf{r}_B, t), \quad (\text{A.2})$$

465 i.e., as  $t$  grows from  $-\infty$  to 0, a time-reversed Green’s function, followed by a regular  $G$  with  
 466 its sign reversed (e.g. [Fink, 2006](#)).

467 It is inferred that the Green’s function between  $\mathbf{r}_A$  and  $\mathbf{r}_B$  can be reconstructed from an  
 468 impulse emitted at  $\mathbf{r}_B$  and recorded at a set of points  $\mathbf{r}$  that sample  $\partial V$ , by (i) time-reversing  
 469 the signal  $G(\mathbf{r}, \mathbf{r}_B, \omega)$  emitted by  $\mathbf{r}_B$  and recorded at  $\mathbf{r}$ ; (ii) convolving the time-reversed  
 470 signal  $G^*(\mathbf{r}', \mathbf{r}_B, \omega)$  with the impulse response  $G(\mathbf{r}_A, \mathbf{r}, \omega)$  between  $\mathbf{r}$  and  $\mathbf{r}_A$ ; (iii) iterating  
 471 over all receivers  $\mathbf{r}$ ; (iv) summing the resulting signals. This procedure is usually referred to  
 472 as “acoustic time reversal,” because the wave field so obtained is essentially a time-reversed,  
 473 backward propagated version of the original impulse response  $G$  ([Fink, 2006](#)); as such, it  
 474 will naturally focus at the original source location, where it will show a very prominent  
 475 maximum. An important consequence of this is that time reversal can be used as a source  
 476 localization tool: if a signal generated by a source at an *unknown* location  $\mathbf{r}_B$  is recorded  
 477 by an array of receivers forming a closed surface  $\partial V$ , by implementing numerically steps (i)  
 478 through (iv) above and looking for the maximum of the resulting numerical wave field, the  
 479 source location can be determined.

480 While in principle  $G$  is accurately reconstructed (the time-reversed wave field focuses at  
 481 the original source location) only if recordings made at a dense, uniform array of receivers  
 482 are time-reversed and backward propagated, many studies have shown that focusing can also  
 483 be achieved using a much smaller receiver array, provided that the medium of propagation  
 484 has some relevant 3-D structure, and that this structure is known and properly accounted  
 485 for when modeling wave propagation. In particular, it was shown ([Catheline et al., 2007](#))  
 486 that a pair of receivers, deployed at ear locations on a human skull, are enough for the

487 time-reversed, backward-propagated signal to sharply focus at the source; since our setup is  
 488 essentially the same, we can reasonably expect two-receiver time reversal to perform about  
 489 equally well in our case. Equation A.1 can thus be simplified to

$$\Im[G(\mathbf{r}_A, \mathbf{r}_B)] \propto G^*(\mathbf{r}_L, \mathbf{r}_B)G(\mathbf{r}_L, \mathbf{r}_A) + G^*(\mathbf{r}_R, \mathbf{r}_B)G(\mathbf{r}_R, \mathbf{r}_A), \quad (\text{A.3})$$

490 where, for the sake of simplicity, and since we are only interested in finding the maxima of  
 491 the expressions in question, we have dropped absolute amplitude information. Equation A.3  
 492 is only valid for impulsive signals, but it is straightforward to generalize it to an arbitrary  
 493 signal  $s(\omega)$ . Write  $s$  as the convolution  $s(\mathbf{x}_1, \mathbf{x}_2, \omega) = h(\omega)G(\omega, \mathbf{x}_1, \mathbf{x}_2)$ , with  $h$  an arbitrary  
 494 “source time function” independent of the source and receiver positions  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . If one  
 495 multiplies both sides of Equation A.3 by  $h^*(\omega)$ ,

$$h^*(\omega)\Im[G(\mathbf{r}_A, \mathbf{r}_B)] \propto s^*(\mathbf{r}_L, \mathbf{r}_B)G(\mathbf{r}_L, \mathbf{r}_A) + s^*(\mathbf{r}_R, \mathbf{r}_B)G(\mathbf{r}_R, \mathbf{r}_A). \quad (\text{A.4})$$

496 The convolution of  $s^*$  with  $G$  at the right-hand side of Equation A.4 should be interpreted,  
 497 again, as backward propagation of the time-reversed recorded signal  $s$ ; Equation A.4 stipu-  
 498 lates that, by this procedure (in the assumption that sufficient information about the wave  
 499 field be recorded by a pair of receivers alone), a source of arbitrary complexity (with respect  
 500 to time) can be reconstructed: the time-reversed signal will focus at the source, where a  
 501 receiver would approximately record the original source time function  $h(t)$ , reversed in time.

502 We take here a slightly different approach (Catheline *et al.*, 2007). Let us multiply both  
 503 sides of Equation A.3 by  $|h(\omega)|^2$ ,

$$|h(\omega)|^2\Im[G(\mathbf{r}_A, \mathbf{r}_B)] \propto s^*(\mathbf{r}_L, \mathbf{r}_B)s(\mathbf{r}_L, \mathbf{r}_A) + s^*(\mathbf{r}_R, \mathbf{r}_B)s(\mathbf{r}_R, \mathbf{r}_A). \quad (\text{A.5})$$



504 Note that the products at the right-hand side of Equation A.5 can be interpreted, in the  
505 time domain, as both the convolution of  $s(\mathbf{r}_{L,R}, \mathbf{r}_A, t)$  with the time-reversed counterpart of  
506  $s(\mathbf{r}_{L,R}, \mathbf{r}_B, t)$ , and the *cross correlation* of  $s(\mathbf{r}_{L,R}, \mathbf{r}_A, t)$  and  $s(\mathbf{r}_{L,R}, \mathbf{r}_B, t)$  (Derode *et al.*, 2003;  
507 Draeger and Fink, 1999). As opposed to Equation A.4, the right-hand side of Equation A.5  
508 does not allow one to reconstruct, from the data, the signal as originally emitted at  $\mathbf{r}_B$   
509 (because  $(s_{\mathbf{r}_{R,L}, \mathbf{r}_A})$  are unknown and cannot be computed). Equation A.5 can be relevant,  
510 however, if the time function  $h(t)$  is known, while the location of the source is to be deter-  
511 mined. This applies, for instance, to echolocating species, that identify and analyze echoes  
512 of signals that they have themselves emitted. Echolocation can presumably be learned by  
513 training, which is equivalent to forming a “library” of observed echoes  $s(\mathbf{r}_{R,L}, \mathbf{r}_A)$  associated  
514 with a given emitted signal and known target locations  $\mathbf{r}_A$ : each time a relevant signal is  
515 perceived, the echolocating agent would then systematically compare it to all recorded traces  
516  $s(\mathbf{r}_{R,L}, \mathbf{r}_A)$ , each corresponding to a different value of  $\mathbf{r}_A$  eventually covering the entire solid  
517 angle. Imagine that this comparison be implemented via cross correlation: this is equiva-  
518 lent to implementing the right-hand side of Equation A.5, and the same equation implies  
519 that cross correlation should be maximum when  $\mathbf{r}_A = \mathbf{r}_B$ ; the sharpness of focusing at the  
520 source, and thus the accuracy of source localization, is strictly related to how well a time-  
521 reversed, backward propagated wave field would focus at the original source. Importantly,  
522 however, the proposed algorithm does not involve any wave propagation modeling, but is  
523 based entirely on signal processing of measurements at two receivers.

524

- 525 Agterberg, M. J., Snik, A. F., Hol, M. K., van Esch, T. E., Cremers, C. W., Van Wanrooij,  
526 M. M., and Van Opstal, A. J. (2011). “Improved horizontal directional hearing in bone  
527 conduction device users with acquired unilateral conductive hearing loss,” *Journal of the*  
528 *Association for Research in Otolaryngology* **12**(1), 1–11.
- 529 Aroyan, J. L. (1996). *Three-dimensional numerical simulation of biosonar signal emission*  
530 *and reception in the common dolphin* (University of California, Santa Cruz).
- 531 Aroyan, J. L. (2001). “Three-dimensional modeling of hearing in delphinus delphis,” *The*  
532 *Journal of the Acoustical Society of America* **110**(6), 3305–3318.
- 533 Aroyan, J. L., Cranford, T. W., Kent, J., and Norris, K. S. (1992). “Computer modeling of  
534 acoustic beam formation in d elphinusdelphis,” *The Journal of the Acoustical Society of*  
535 *America* **92**(5), 2539–2545.
- 536 Au, W. W. (2012). *The sonar of dolphins* (Springer Science & Business Media).
- 537 Au, W. W., and Fay, R. R. (2012). *Hearing by whales and dolphins*, **12** (Springer Science  
538 & Business Media).
- 539 Au, W. W., and Hastings, M. C. (2008). *Principles of marine bioacoustics* (Springer).
- 540 Au, W. W., Moore, P. W., and Pawloski, D. A. (1988). “Detection of complex echoes in  
541 noise by an echolocating dolphin,” *The Journal of the Acoustical Society of America* **83**(2),  
542 662–668.
- 543 Batteau, D. W. (1967). “The role of the pinna in human localization,” *Proc. R. Soc. Lond.*  
544 *B* **168**(1011), 158–180.

- 545 Blauert, J. (1969). "Sound localization in the median plane," *Acta Acustica united with*  
546 *Acustica* **22**(4), 205–213.
- 547 Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization* (MIT  
548 press).
- 549 Boschi, L., and Weemstra, C. (2015). "Stationary-phase integrals in the cross correlation of  
550 ambient noise," *Reviews of Geophysics* **53**(2), 411–451.
- 551 Branstetter, B. K., and Mercado III, E. (2006). "Sound localization by cetaceans," *Inter-*  
552 *national Journal of Comparative Psychology* **19**(1).
- 553 Brill, R. L., Moore, P. W., and Dankiewicz, L. A. (2001). "Assessment of dolphin (*tursiops*  
554 *truncatus*) auditory sensitivity and hearing loss using jawphones," *The Journal of the*  
555 *Acoustical Society of America* **109**(4), 1717–1722.
- 556 Brill, R. L., Sevenich, M. L., Sullivan, T. J., Sustman, J. D., and Witt, R. E. (1988). "Be-  
557 havioral evidence for hearing through the lower jaw by an echolocating dolphin (*tursiops*  
558 *truncatus*)," *Marine Mammal Science* **4**(3), 223–230.
- 559 Butler, R. A., and Belendiuk, K. (1969). "Monaural and binaural localization of noise bursts  
560 vertically in the median sagittal plane," *Journal of the Auditory Research* **3**, 230–235.
- 561 Butler, R. A., Humanski, R. A., and Musicant, A. D. (1990). "Binaural and monaural  
562 localization of sound in two-dimensional space," *Perception* **19**(2), 241–256.
- 563 Catheline, S., Fink, M., Quieffin, N., and Ing, R. K. (2007). "Acoustic source localiza-  
564 tion model using in-skull reverberation and time reversal," *Applied physics letters* **90**(6),  
565 063902.

- 566 Cranford, T. W., and Krysl, P. (2015). “Fin whale sound reception mechanisms: skull  
567 vibration enables low-frequency hearing,” *PloS one* **10**(1), e0116222.
- 568 Cranford, T. W., Krysl, P., and Hildebrand, J. A. (2008). “Acoustic pathways revealed:  
569 simulated sound transmission and reception in cuvier’s beaked whale (*ziphius cavirostris*),”  
570 *Bioinspiration & Biomimetics* **3**(1), 016001.
- 571 Derode, A., Larose, E., Tanter, M., De Rosny, J., Tourin, A., Campillo, M., and Fink, M.  
572 (2003). “Recovering the greens function from field-field correlations in an open scattering  
573 medium (I),” *The Journal of the Acoustical Society of America* **113**(6), 2973–2976.
- 574 Dible, S. A., Flint, J. A., and Lepper, P. A. (2009). “On the role of periodic structures in  
575 the lower jaw of the atlantic bottlenose dolphin (*tursiops truncatus*),” *Bioinspiration &*  
576 *biomimetics* **4**(1), 015005.
- 577 Draeger, C., and Fink, M. (1999). “One-channel time-reversal in chaotic cavities: Theoret-  
578 ical limits,” *The Journal of the Acoustical Society of America* **105**(2), 611–617.
- 579 Fink, M. (2006). “Time-reversal acoustics in complex environments,” *Geophysics* **71**(4),  
580 SI151–SI164.
- 581 Fink, M., Cassereau, D., Derode, A., Prada, C., Roux, P., Tanter, M., Thomas, J.-L., and  
582 Wu, F. (2000). “Time-reversed acoustics,” *Reports on Progress in Physics* **63**(12), 1933.
- 583 Graf, S., Blondel, P., Megill, W. M., and Clift, S. E. (2009). “Acoustic modelling of dolphin  
584 sound reception and implications for biosonar design,” in *OCEANS 2009-EUROPE*, IEEE,  
585 pp. 1–6.
- 586 Gray, M. D., and Rogers, P. H. (2017). “In vivo ultrasonic attenuation in cetacean soft  
587 tissues,” *The Journal of the Acoustical Society of America* **141**(2), EL83–EL88.

- 588 Harris, F. J. (1978). “On the use of windows for harmonic analysis with the discrete fourier  
589 transform,” *Proceedings of the IEEE* **66**(1), 51–83.
- 590 Hartmann, W. M. (1999). “How we localize sound,” *Physics today* **52**, 24–29.
- 591 Heffner, H. E., and Heffner, R. S. (2016). “The evolution of mammalian sound localization,”  
592 *Acoustics Today* **12**, 20–27.
- 593 Heffner, R. S., and Heffner, H. E. (1992). “Evolution of sound localization in mammals,”  
594 in *The evolutionary biology of hearing* (Springer), pp. 691–715.
- 595 Hunter, J., and Banks, J. (1787). “Observations on the structure and oeconomy of whales,”  
596 *Philosophical Transactions of the Royal Society of London* **77**, 371–450.
- 597 Ing, R. K., Quieffin, N., Catheline, S., and Fink, M. (2005). “In solid localization of finger  
598 impacts using acoustic time-reversal process,” *Applied Physics Letters* **87**(20), 204104.
- 599 Keller, C. H., Hartung, K., and Takahashi, T. T. (1998). “Head-related transfer functions  
600 of the barn owl: measurement and neural responses,” *Hearing research* **118**(1), 13–34.
- 601 Ketten, D. (1994). “Functional analyses of whale ears: adaptations for underwater hear-  
602 ing,” in *OCEANS’94. ’Oceans Engineering for Today’s Technology and Tomorrow’s Preser-  
603 vation. ’Proceedings*, IEEE, Vol. 1, pp. I–264.
- 604 Ketten, D. R. (1992). “The marine mammal ear: specializations for aquatic audition and  
605 echolocation,” in *The evolutionary biology of hearing* (Springer), pp. 717–750.
- 606 Ketten, D. R. (1997). “Structure and function in whale ears,” *Bioacoustics* **8**(1-2), 103–135.
- 607 Ketten, D. R. (2000). “Cetacean ears,” in *Hearing by Whales and Dolphins* (Springer New  
608 York, New York, NY), pp. 43–108.

- 609 Kim, S., Kuperman, W., Hodgkiss, W., Song, H., Edelmann, G., and Akal, T. (2003).  
610 “Robust time reversal focusing in the ocean,” *The Journal of the Acoustical Society of*  
611 *America* **114**(1), 145–157.
- 612 Krysl, P., and Cranford, T. W. (2016). “Directional hearing and head-related transfer func-  
613 tion in odontocete cetaceans,” in *The Effects of Noise on Aquatic Life II* (Springer), pp.  
614 583–587.
- 615 Macpherson, E. A., and Sabin, A. T. (2013). “Vertical-plane sound localization with dis-  
616 torted spectral cues,” *Hearing research* **306**, 76–92.
- 617 McCormick, J. G., Wever, E., Palin, J., and Ridgway, S. (1970). “Sound conduction in the  
618 dolphin ear,” *The Journal of the Acoustical Society of America* **48**(6B), 1418–1428.
- 619 Møhl, B., Au, W., Pawloski, J., and Nachtigall, P. (1999). “Dolphin hearing: relative  
620 sensitivity as a function of point of application of a contact sound source in the jaw and  
621 head region,” *The Journal of the Acoustical Society of America* **105**(6), 3421–3424.
- 622 Mooney, T. A., Yamato, M., and Branstetter, B. K. (2012). “Hearing in cetaceans: from  
623 natural history to experimental biology,” *Adv. Mar. Biol* **63**(197-246).
- 624 Moore, P. W., and Au, W. W. (1975). “Underwater localization of pulsed pure tones by  
625 the california sea lion (*z alophuscalifornianus*),” *The Journal of the Acoustical Society of*  
626 *America* **58**(3), 721–727.
- 627 Moore, P. W., Pawloski, D. A., and Dankiewicz, L. (1995). “Interaural time and intensity  
628 difference thresholds in the bottlenose dolphin (*tursiops truncatus*),” *Sensory Systems of*  
629 *Aquatic Mammals* 11–23.

- 630 Mulsow, J., Finneran, J. J., and Houser, D. S. (2014). “Interaural differences in the bot-  
631 tlenose dolphin (*tursiops truncatus*) auditory nerve response to jawphone click stimuli,”  
632 The Journal of the Acoustical Society of America **136**(3), 1402–1409.
- 633 Nachtigall, P. E. (2016). “Biosonar and sound localization in dolphins,” [http://](http://neuroscience.oxfordre.com/view/10.1093/acrefore/9780190264086.001.0001/acrefore-9780190264086-e-103)  
634 [neuroscience.oxfordre.com/view/10.1093/acrefore/9780190264086.001.0001/](http://neuroscience.oxfordre.com/view/10.1093/acrefore/9780190264086.001.0001/acrefore-9780190264086-e-103)  
635 [acrefore-9780190264086-e-103](http://neuroscience.oxfordre.com/view/10.1093/acrefore/9780190264086.001.0001/acrefore-9780190264086-e-103), oxford Research Encyclopedia of Neuroscience.
- 636 Norris, K. S. (1964). “Some problems of echolocation in cetaceans,” in *Marine Bioacoustics*,  
637 edited by W. N. Tavolga (Pergamon Press, New York), pp. 316–336.
- 638 Norris, K. S. (1968a). “The echolocation of marine mammals.,” in *The biology of marine*  
639 *mammals*, edited by H. T. Andersen (Academic Press, New York), pp. 391–423.
- 640 Norris, K. S. (1968b). “The evolution of acoustic mechanisms in odontocete cetaceans,” in  
641 *Evolution and Environment*, edited by E. T. Drake (Yale University Press, New York), pp.  
642 297–324.
- 643 Norris, K. S., and Harvey, G. W. (1974). “Sound transmission in the porpoise head,” The  
644 Journal of the Acoustical Society of America **56**(2), 659–664.
- 645 Popov, V. V., and Supin, A. Y. (1990). “Localization of the acoustic window at the dolphins  
646 head,” in *Sensory Abilities of Cetaceans* (Springer), pp. 417–426.
- 647 Renaud, D. L., and Popper, A. N. (1975). “Sound localization by the bottlenose porpoise  
648 *tursiops truncatus*,” Journal of Experimental Biology **63**(3), 569–585.
- 649 Reysenbach de Haan, F. W. (1957). “Hearing in whales,” Acta Oto-Laryngologica. Supple-  
650 mentum **134**, 1.

- 651 Richardson, W. J., Greene Jr, C. R., Malme, C. I., and Thomson, D. H. (2013). *Marine*  
652 *mammals and noise* (Academic press).
- 653 Ryabov, V. (2010). “Role of the mental foramens in dolphin hearing,” *Natural Science* **2**(6),  
654 646–653.
- 655 Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and  
656 Roch, M. A. (2008). “Classification of rissos and pacific white-sided dolphins using spectral  
657 properties of echolocation clicks,” *The Journal of the Acoustical Society of America* **124**(1),  
658 609–624.
- 659 Soldevilla, M. S., McKenna, M. F., Wiggins, S. M., Shadwick, R. E., Cranford, T. W.,  
660 and Hildebrand, J. A. (2005). “Cuvier’s beaked whale (*ziphius cavirostris*) head tissues:  
661 physical properties and ct imaging,” *Journal of experimental biology* **208**(12), 2319–2332.
- 662 Song, Z., Zhang, Y., Thornton, S. W., Li, S., and Dong, J. (2017). “The influence of air-  
663 filled structures on wave propagation and beam formation of a pygmy sperm whale (*kogia*  
664 *breviceps*) in horizontal and vertical planes,” *The Journal of the Acoustical Society of*  
665 *America* **142**(4), 2443–2453.
- 666 Song, Z., Zhang, Y., Wei, C., and Wang, X. (2016). “Inducing rostrum interfacial waves  
667 by fluid-solid coupling in a chinese river dolphin (*lipotes vexillifer*),” *Physical Review E*  
668 **93**(1), 012411.
- 669 Supin, A. Y., and Popov, V. V. (1993). “Direction-dependent spectral sensitivity and in-  
670 teraural spectral difference in a dolphin: Evoked potential study,” *The Journal of the*  
671 *Acoustical Society of America* **93**(6), 3490–3495.



- 672 Van Opstal, J. (2016). *The auditory system and human sound-localization behavior* (Aca-  
673 demic Press).
- 674 Wei, C., Au, W. W., Song, Z., and Zhang, Y. (2016). “The role of various structures in the  
675 head on the formation of the biosonar beam of the baiji (*lipotes vexillifer*),” *The Journal*  
676 *of the Acoustical Society of America* **139**(2), 875–880.
- 677 Wei, C., Song, Z., Au, W. W., Zhang, Y., and Wang, D. (2018). “A numerical evidence of  
678 biosonar beam formation of a neonate yangtze finless porpoise (*neophocaena asiaeorien-*  
679 *talis*),” *Journal of Theoretical and Computational Acoustics* **26**(02), 1850009.
- 680 Youssef, K., Argentieri, S., and Zarader, J.-L. (2012). “Towards a systematic study of  
681 binaural cues,” in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International*  
682 *Conference on*, IEEE, pp. 1004–1009.