Aquatic Mammals 2017, *43*(2), 162-176, DOI 10.1578/AM.43.2.2017.162

First Results of an Underwater 360° HD Audio-Video Device for Etho-Acoustical Studies on Bottlenose Dolphins (*Tursiops truncatus*)

Juliana Lopez-Marulanda,[1] Olivier Adam,[1,2] Torea Blanchard,[1,3] Marie Vallée,[4] Dorian Cazau,[5] and Fabienne Delfour[3,6]

[1]*Institute of Neurosciences Paris Saclay, CNRS UMR 9197, University Paris Sud, Orsay, France*
E-mail: juliana.lopez-marulanda@u-psud.fr
[2]*Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7190, Institut Jean Le Rond d'Alembert, F-75005 Paris, France*
[3]*Parc Astérix, Plailly, France*
[4]*Association Abyss, La Réunion, France*
[5]*ENSTA Bretagne Lab-STICC, CNRS UMR 6285, Université Européenne de Bretagne, Brest, France*
[6]*Laboratoire d'Ethologie Expérimentale et Comparée EA 4443 (LEEC), Université Paris 13, Sorbonne Paris Cité, F-93430 Villetaneuse, France*

Abstract

Bottlenose dolphins (*Tursiops truncatus*) are highly social odontocetes that live in a fission-fusion society and demonstrate production of a varied sound repertoire, including clicks, whistles, and burst-pulsed sounds, as well as a diverse behavioral repertoire. To better understand the species' behavior, it is necessary to compare visual and acoustic observations and link vocalizations to individuals and their specific actions. However, the task of linking sounds to individual dolphins is challenging for human observers because dolphins do not always display specific visual cues when producing a sound, and also because human hearing is not naturally adapted to locate underwater sound sources. To respond to these challenges, a new underwater 360° HD audio-video device, the BaBeL, was designed and built. This device consists of a five-hydrophone array attached to two wide-angle video cameras that together cover a 360° field of vision. Acoustic recordings were analyzed with a customized program to detect and localize sound sources and to identify individual vocalizing dolphins. Data from a population of bottlenose dolphins were collected during 14 boat surveys along the northwest coast of Reunion Island (France) by following a strict pre-established protocol to standardize data collection. A total of 21 min of audio-video were recorded when dolphins were present, and 42 click trains and 42 whistles were detected from these data. Dolphins identified as vocalizers were also present for 17% ($n = 7$) of emitted click trains and 33% ($n = 14$) of emitted whistles on the videos. Therefore, an analysis of three video sequences as examples of the scope of this methodology is presented. The results show that when the observers stayed ahead and avoided the direct path of groups of five to nine dolphins, only one animal emitted click trains while swimming towards the observers or after turning its rostrum in the humans' direction, and this dolphin was never the one leading the group. The benefits of using this audio-video device for underwater observations of dolphins in clear water with good visibility are discussed.

Key Words: behavior, acoustics, hydrophone array, acoustic localization, bottlenose dolphin, *Tursiops truncatus*

Introduction

Bottlenose dolphins (*Tursiops truncatus*) are highly social odontocetes with a fission-fusion social structure (Connor et al., 2000; Mann et al., 2000; Gibson & Mann, 2008; Tsai & Mann, 2013). Group members may travel very short or very long distances within a habitat of limited visibility (Connor et al., 1998). As such, communication via acoustic signals is the most effective strategy for sharing information under water (Janik, 1999). Bottlenose dolphins display a complex and extensive repertoire of sounds such as clicks or pulsed sounds (Au, 1993; Au & Fay, 2012), burst-pulsed sounds (Lopez & Bernal-Shirai, 2009), and whistles or tonal sounds (reviewed in Janik, 2009).

Dolphin sound production is enriched by non-acoustic communication signals during social interactions when individuals are within visual range of one another. These animals display various body postures (Pryor, 1990), contacts (Sakai et al., 2006; Dudzinski et al., 2009), and bubble emissions (Marten et al., 1996). Several dolphin vocalizations are associated with behavioral contexts; for example, burst-pulsed sounds (squawks and whines) have been associated with agonistic behaviors (Herzing, 1996), and low-frequency bray calls are related to feeding (Janik, 2000; King & Janik, 2015).

Studies of dolphin social behavior and communication rely on simultaneous descriptions of both visual and acoustic signals (Thomas et al., 2002). However, the main obstacle associated with completing these descriptions is the difficulty in identifying which dolphin in a group is the vocalizer. This challenge is caused by two factors: (1) human hearing is not adapted to localize to sound sources underwater; and (2) dolphins do not show visible, regular signs when emitting sounds, like opening their mouths or displaying external clues (Janik, 2009). To overcome these obstacles, several methodologies have been developed. Animals have been isolated (Caldwell et al., 1990; Sayigh et al., 1990) or tagged (Tyack, 1991; Nowacek et al., 1998); however, these approaches can be considered invasive and might lead to modification of the subjects' behaviors and vocalizations. Emission of bubble streams concurrent with vocalizations has been used to identify a vocal animal because sometimes dolphins emit bubbles while whistling (McCowan & Reiss, 1995; Herzing, 1996); however, whistles with bubble streams are not representative of the entire whistle repertoire of bottlenose dolphins (Fripp, 2005, 2006).

As a non-intrusive method to identify the vocalizing animal, different hydrophone arrays have been designed. These arrays allow for processing of the differences in time of arrival of the sound to each hydrophone to determine where the call originated. The position of the sound source is linked to the video recordings to confirm which animal is in the same position as the sound source, thereby identifying the vocalizer. Fixed arrays using two (López-Rivas & Bazuá-Durán, 2010), three (Watkins & Schevill, 1974), four (Brensing et al., 2001; Quick et al., 2008), and eight (Thomas et al., 2002) hydrophones have been used to link audio recordings to behavioral observations or video recordings; however, fixed arrays are not well adapted to study highly mobile, free-ranging dolphins, and the video recordings were often obtained from a fixed point at the surface. The main problem with acquiring behavioral information from the surface is that the documented behaviors could possibly represent only a very small percentage of an animal's behavioral activity (Janik, 2009). Moreover, an array with two hydrophones (Lopez-Rivas & Bazuá-Durán, 2010) allows data to be obtained only on the angle of arrival and not the real position of the emitting dolphin. Four hydrophones are needed to localize a moving dolphin in 3D (Watkins & Schevill, 1972; Wahlberg et al., 2001)

Mobile arrays of two (Dudzinski et al., 1995), four (Au & Herzing, 2003; Schotten et al., 2004), and 16 (Ball & Buck, 2005) hydrophones have been used to study dolphin vocalizations and their associated underwater behaviors, but they, too, presented several disadvantages. Dudzinski et al.'s (1995) system did not allow localization in the vertical axis, and the systems with four hydrophones were used only for localization of click emitters (Au & Herzing, 2003; Schotten et al., 2004). The 16-hydrophone array had elements separated by 3.2 cm (Ball & Buck, 2005) but did not allow the confirmation of the emitter's identity if animals were located outside of the narrow angle of the video camera.

As part of this study, an audio-video system that was non-intrusive and compact enough to be deployed from a small boat was developed. This underwater device includes five hydrophones and a 360° HD video recording system with a limited blind spot that allows localization of sounds to free-swimming, vocalizing dolphins coming from almost every direction. In this article, details about this system's design and the software developed to localize to sounds and to link them to individually identified dolphins are provided. Three audio-video sequences of free-ranging bottlenose dolphins have been analyzed to illustrate the benefits of this system in dolphin ethological and acoustical research.

Methods

Recording Device

Simultaneous audio and video recordings were collected using a waterproof audio-video system named BaBeL (BioAcoustique, Bien Être et Langage) (Figure 1). The acoustic set-up was comprised of five calibrated Aquarian H2a-XLR hydrophones connected and synchronized to a ZOOM H6 digital audio recorder. Audio recordings were made at a 96-kHz sampling frequency and coded on 24 bits. The recorder was placed in a waterproof housing rated to 60 m depth. The architectural design of the hydrophone array was a compromise between a large aperture between hydrophones and maneuverability since the system needed to be deployed from small boats with limited space and to be controllable by one observer when submerged. The synchronized hydrophones were positioned to obtain the time delay of arrival to provide the 3D estimations of dolphin positions.

The video portion of the BaBel system was comprised of two Kodak SP360 video cameras

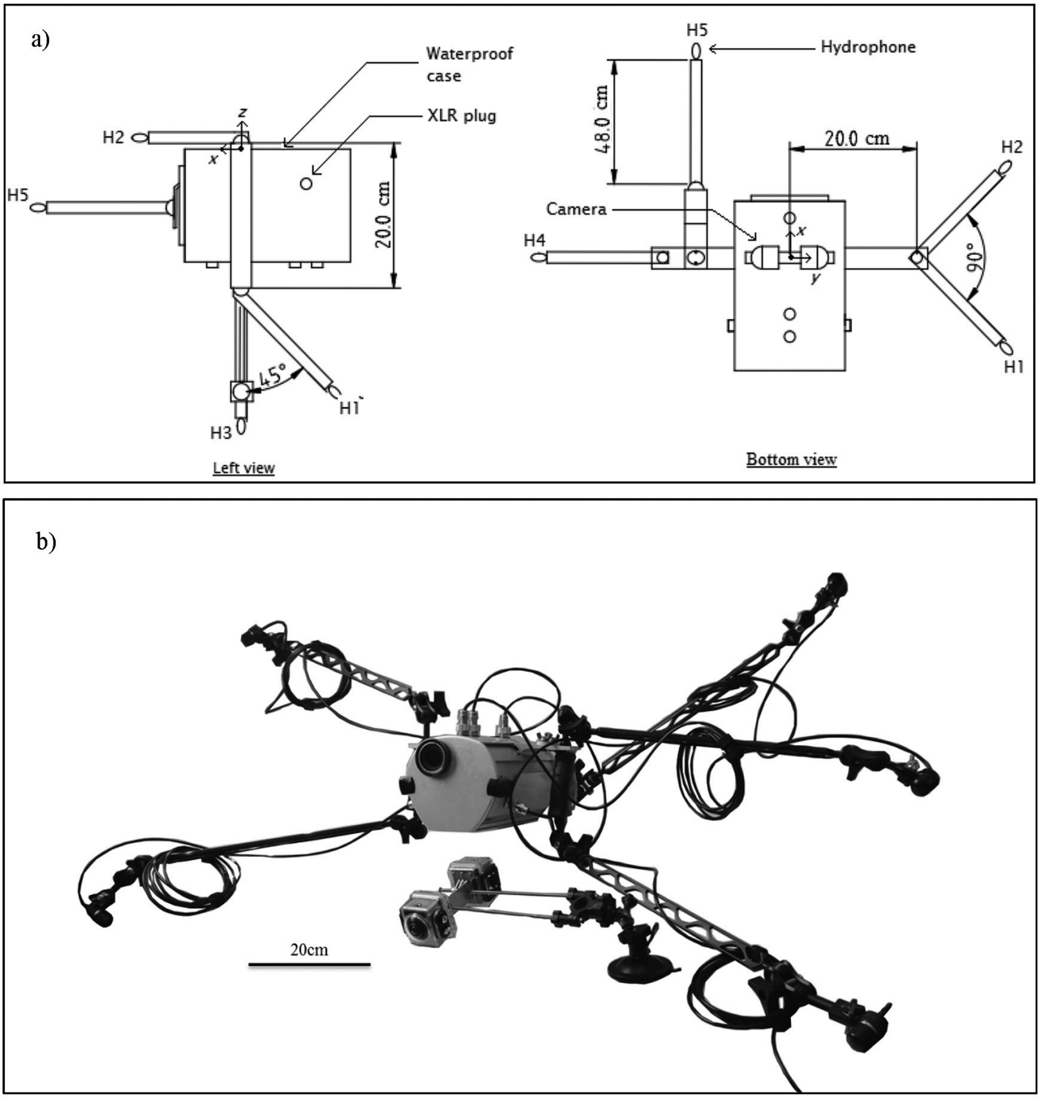

Figure 1. BaBeL (BioAcoustique, Bien Être et Langage) device: (a) Diagram and orientation of hydrophones (H); and (b) picture of the device with 360º cameras unattached—five hydrophones installed on five deployable arms and plugged to a ZOOM H6 (inside the adapted waterproof case).

both with a wide-angle field of view (214°); the cameras were placed opposite of each other to the left and the right to allow for 360° field of vision for the system. These cameras were positioned below the waterproof housing of the acoustic recorder (Figure 2). Video and audio files were stored for *a posteriori* analysis.

Custom-Made Program for Data Analysis

A geometrical localization method was used to estimate the position of an acoustic source. This method used the spatial distribution of hydrophones, the acoustic properties of the source (e.g., propagation speed and spherical propagation model), and the measure of the time differences of arrival (TDOA) of the acoustic wave from the source to the different hydrophones (Alameda-Pineda & Horaud, 2014). The aim was to estimate the differences in time of arrival of emitted sounds; the cross-correlation function method for whistle detection (Van Lancker, 2001) and the threshold time energy for click detection

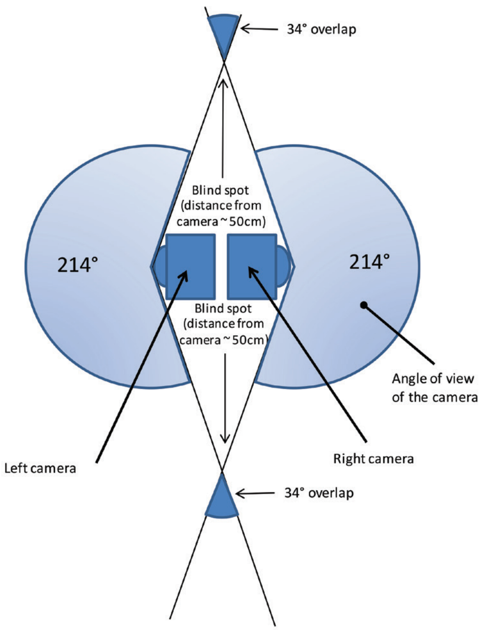

Figure 2. Disposition of two Kodak SP360 video cameras. Each camera has a 360° (N-S-E-W) plus 214° angle of view. As both cameras are placed opposite to each other, there is a 34° overlap in the images and a ≈ 50-cm blind spot between the cameras.

(Blanchard, 2015) were used. To display the estimated position of the acoustic source in the video image, a conversion position-pixel that took into account the deformations of the image because of the spherical curved lens of the Kodak SP360 video cameras was used. With these considerations, a customized program to analyze data obtained with the BaBeL system was created in *MATLAB*®, Version 2013a (Mathworks, Natick, MA, USA) to synchronize the video and audio recordings and then to estimate the localization(s) of each vocalizing dolphin (Blanchard, 2015) (Figure 3). After identifying the location of the vocalizing dolphin, video analysis allowed for the identification of the dolphin based on recognizable scars and marks.

Tests with Artificial Sounds

Two simulated whistles to test this approach with different signal-to-noise ratios (SNRs) (Figure 4) were created. The objective of this test was to confirm performance of the time correlation for acoustic signals deteriorated by underwater acoustic propagation or when ambient noise is present in the marine area. To verify our localization method, the system was tested in a 3.1 × 8.2 m^2 rectangular freshwater swimming pool. The BaBeL was immersed in the center of the

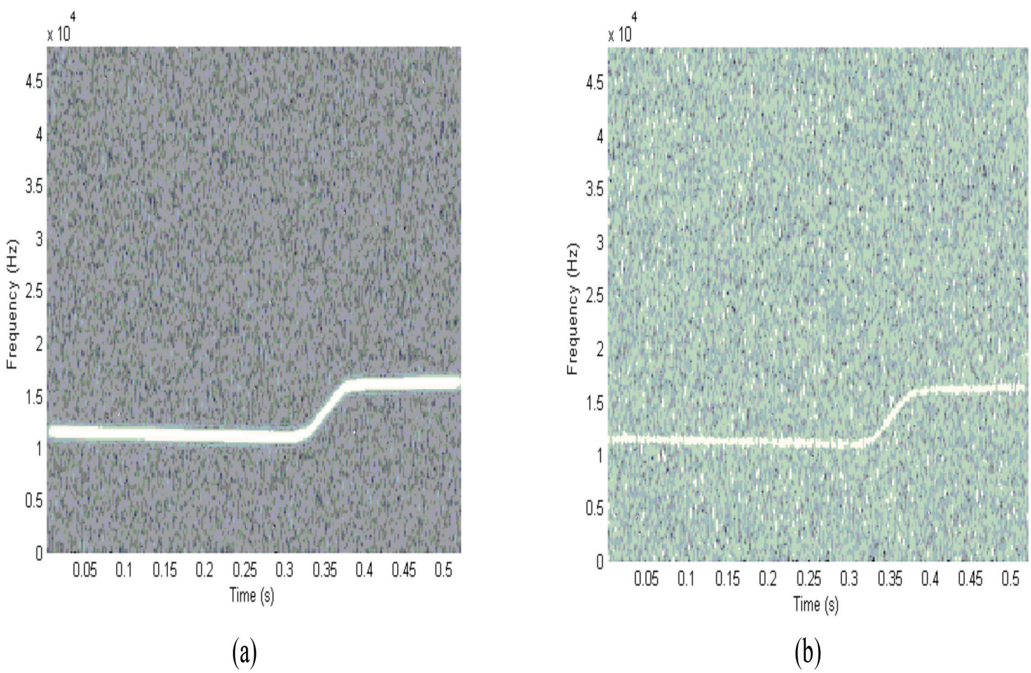

Figure 4. Simulated whistles with two signal-to-noise ratios (SNRs): (a) +20 dB and (b) -10 dB.

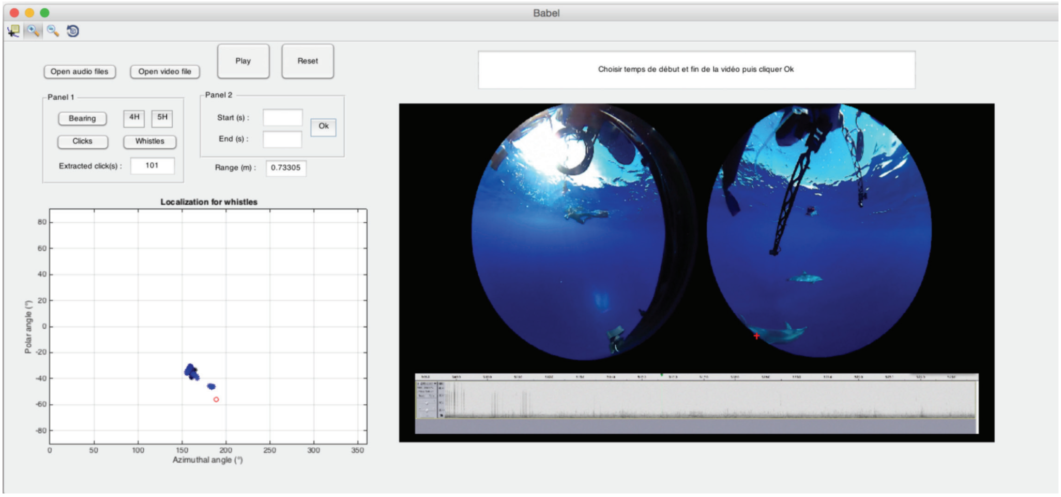

Figure 4. Screen display to track dolphins by videos and passive acoustics. On the bottom left, estimations of the angles from the successive clicks (in blue) and the whistle (in red). On the right, the red cross points to the emitter dolphin.

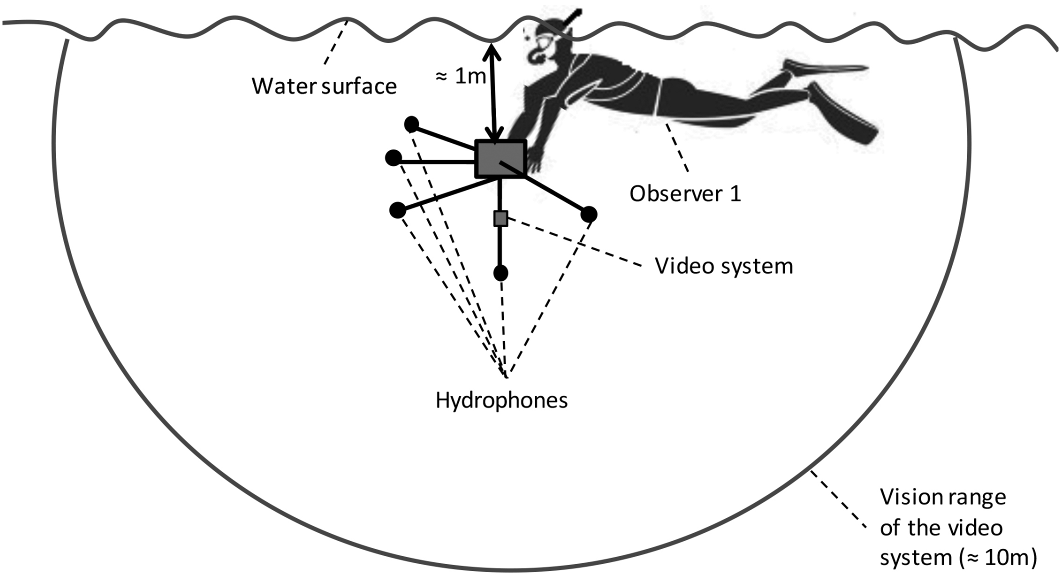

Figure 5. Disposition of BaBeL device in the water; the whole system is controlled by Observer 1. The vision range of the camera depends on water clarity. For the Reunion Island, it is ≈10 m. The real scales are not represented in the figure.

pool at 2.5 m from the edge. Percussive sounds were generated by knocking together two steel bars from nine different known places in the horizontal plane of the device's gravity center. Using the position-pixel conversion, the position of each percussive sound source in the video image was estimated and compared to the location estimated by the custom program.

Bottlenose Dolphin Data Collection

Acoustic and video data were collected on free-ranging bottlenose dolphins along the northwest coast off Reunion Island, a French territory in the Mascarene Islands in the Southwest Indian Ocean. The species is observed in this location throughout the year in groups of 10 to 100 individuals (48 individuals on average) and occurs in deeper water

(425.6 m on average) and further offshore (1.2 to 6 km from the coast) than other cetacean species in this area (Dulau-Drouot et al., 2008).

Fourteen boat surveys were conducted from 21-29 May 2015 and from 6-18 June 2015 to search for bottlenose dolphins and collect etho-acoustical data. When a group of dolphins was sighted, a strict pre-established protocol was followed (see *Agreement on the Conservation of Cetaceans in the Black Sea Mediterranean Sea and Contiguous Atlantic Area* [ACCOBAMS], Resolution 4.18) to decide if observers would enter the water to start a recording session. First, the boat was positioned parallel to the animals' travel direction at a distance of more than 50 m. The behavioral response of the dolphin group was recorded into one of three categories: (1) "avoidance," (2) "indifference," and (3) "oncoming" (see ACCOBAMS for definitions). If the behavioral response was cataloged as "indifference" or "oncoming," the boat was slowly positioned 100 m ahead of the first animal of the group, never interfering with the travel direction of the animals. Once in this position, two observers slipped into the water.

Procedure in Water—One observer swam with the BaBeL device submerged below the sea surface (≈1 m under the surface) (Figure 5), while the other observer recorded the animals on a backup SONY HDR-GW66 video camera. Date and time on all video cameras were synchronized for *a posteriori* analysis. Since BaBeL was being operated for the first time, the intent was to document all the events. The backup video sequences might be used later to confirm what was observed on the BaBeL wide-angle cameras, and the recorded sequences might be replayed to document the BaBeL operator's position and behavior in the water.

At the beginning of the recording session, two successive claps were made—one in front of each camera in order to synchronize both videos with audio recordings during the *a posteriori* analysis. Both observers remained floating at the surface with their bodies oriented perpendicular to the group's travel direction, avoiding the direct path of the dolphins and letting the animals choose at what speed and distance they approached. When dolphins slowly moved along the observers, they swam calmly in parallel with the animals. Depending on whether the dolphins stayed around the observers or departed, recording sessions were repeated several times on the same group by carefully re-orienting the boat and by informing the observers each time the dolphins swam by. A recording session finished when dolphins were not visible for 5 min or when weather conditions prevented continued observations.

Audio-Video Analysis

The claps at the beginning of each recording session were used to manually synchronize acoustic and video data with video editing software (*Final Cut Pro X*, Version 10.1.3©, Apple Inc.). A single video file was created displaying the videos of the two Kodak SP360 video cameras in the same window, as well as one of the five audio tracks and its corresponding turning spectrogram (FFT size: 1,024, overlap 50%, Hanning window) obtained with *Audacity*, Version 2.0.6 (GNU General Public License). We chose only one track in the video as a referent since the five audio tracks were used for our custom-made acoustical analysis in *MATLAB*® (Blanchard, 2015).

The location of the vocalizing dolphin was noted as "visible" when our program was able to point out one of the dolphins in the video, "ambiguous" when the program pointed out two dolphins that were close to each other or in the same direction, and "not visible" when the program pointed to another direction indicating that the emitter dolphin was far outside the range of vision of the video cameras, estimated at further than 10 m away in all directions but also dependent on the wide-angle lens (reduces the size of objects) and water clarity.

To conduct our etho-analysis, the sequences in which we could locate with no doubt at least one vocalizing dolphin were selected (see Appendix 2, Figure 1); a focal-animal sampling technique was used to note occurrence and duration of body postures, tactile contacts, and other behaviors displayed during intraspecific interactions and during interactions towards humans in video sequences (Altmann, 1974; Mann, 1999). Since all sightings were mainly "swim by" wherein the dolphins did not remain near the observers for long, individual dolphins in each sequence were listed in order of appearance in the video. The "all occurrences" recording sampling method focused on frequencies, and durations of occurring behaviors was used (Martin et al., 1993). The analyzed sequences allowed the researchers to create a behavioral catalog (Tinbergen, 1963), which included nonsocial and social (intraspecific and human-dolphin interactions) behaviors and sounds produced (Table 1).

Results

Tests with Artificial Sounds

As this study is dedicated to the analysis of behaviors, only situations when at least one dolphin was visible in the videos, at a distance of less than 10 m, were taken into account. If this dolphin emitted clicks and/or whistles during a period with no underwater noise, then the SNR ratio

Table 1. Behavioral catalog of the dolphins observed while swimming by observers and documented with the BaBeL device

Behaviors	Code	Definition
Pectoral rubbing	PR	The dolphin touches another dolphin.
Synchronized swimming	SyS	Dolphins swim in synchronous manner within one body length of another dolphin, showing parallel movements and body axes.
Swim upside down	SUD	The dolphin swims with its belly turned up.
Swim upside down underneath	SUDU	The dolphin swims with its belly turned up underneath a conspecific.
Side swimming	SS	The dolphin swims with its belly turned to the right or the left next to a conspecific.
Approach	APP	The dolphin approaches the observers by leaving the direction axis of its group.
Swim towards observers	STO	The dolphin swims towards the BaBeL device and the observers.
Turn rostrum	TR	The dolphin turns its rostrum in the direction of an observer.
Leave	L	The dolphin stops swimming towards the observers and starts to move away.
Whistle	W	The dolphin whistles.
Click train	C	The dolphin emits a click train.

Table 2. Accuracy of the time correlation on simulated signals

Simulated signal	Features	SNR = 20 dB	SNR = -10 dB
Whistle #1	Duration: 0.1 s Fundamental: 11 kHz	ΔTDOA = 0	ΔTDOA = 1.8 ms
Whistle #2	Duration: 0.5 s Fundamental frequency 11 kHz switch to 16 kHz at 0.35 s	ΔTDOA = 0	ΔTDOA = 0

was higher than 20 dB. If the dolphin vocalizes further away, SNR decreased and could be negative. We performed our approach for positive and negative SNR (Table 2). The time differences of arrival (TDOAs) were still correctly estimated for SNR larger than -10 dB, which is acceptable for our study because underwater noise was low compared to dolphin sounds. (SNR was always positive in our acoustic recordings.)

Results of the first test comparing estimations of positions in video and audio show that differences in estimation for azimuthal localizations are less than 12° except for in positions 3 and 4 (Table 3). For elevation localizations, the difference is less than 10°, except for in position 8. Positions 3, 4, and 8 can only be seen right on the edge of the image, making estimations more difficult due to image compression. Taking into account that the maximal vision range of the BaBeL is estimated to be 10 m depending on water clarity, a 10° difference in estimations from video and audio means that localization at 10 m from the BaBeL can have a maximum difference of 1.7 m from the position of the source in the video, which is less than one bottlenose dolphin body length. For distances from the device, the error of the custom program was 1.1 to 3.9 m (Table 3).

Data Description

During 14 boat surveys, dolphins were sighted four times, allowing collection of 21.03 min of 360° HD audio-video data with dolphins present. Recordings allowed the detection of 42 click trains and 42 whistles. The vocalizing dolphin was localized and visible on the video for seven click trains (17%). The vocalizer was not visible for 25 click trains (59%); and for 10 click trains (24%), localization of the vocalizer was ambiguous. For whistles, localization analysis was not possible for five whistles (12%) because of a low SNR ratio. The vocalizing dolphin was visible on the video for 14 whistles (33%), the vocalizer was not visible on the video for 18 whistles (43%), and the localization of the vocalizer was ambiguous for five whistles (12%).

Three recording sessions (24 May at 0937 and 0949 h, and 27 May at 1316 h) were chosen during which it was possible to localize the

Table 3. Localization performance of our custom-made program using the sound produced by two bars of steel during tests in a pool

	Position estimated from the videos			Position estimated from the acoustic recordings			Difference in estimation		
Position	Azimuth (°)	Elevation (°)	Distance (m)	Azimuth (°)	Elevation (°)	Distance (m)	Azimuth (°)	Elevation (°)	Distance (m)
1	33.0	5.3	4.7	21.4	3.4	6.3	11.6	1.9	-1.6
2	5.7	1.2	4.5	6.2	-2.6	0.8	-0.5	3.8	3.7
3	328.8	3.5	4.7	349.6	1.1	0.8	-20.8	2.4	3.9
4	315	3.1	3.3	349.3	-3.1	1.5	-34.3	6.2	1.8
5	8.0	-6.1	3.0	2.9	-7.9	0.8	5.1	1.8	2.2
6	43.6	2.3	3.3	36.3	4.2	1.3	7.3	-1.9	2.0
7	61.6	5.6	2.1	55.4	-2.8	0.7	6.2	8.4	1.4
8	6.8	0.7	1.5	0.6	12.5	0.4	6.2	-11.8	1.1
9	304.4	-0.7	2.1	303.1	-7.7	0.5	1.3	7.0	1.6
						Mean	-2.0	2.0	1.8
						SD	15.2	6.0	1.6

dolphin vocalizing to facilitate completion of detailed analyses of dolphin behavior according to the behavioral catalog (Table 1; see Appendix 1). Results show that the first animal of the group to approach the observers did not produce click trains. A click train was made after an approach and/or movement of the rostrum towards the device: in the first observation, the click train was emitted by the second individual after it turned its rostrum towards the device (see Appendix 1, Figure 1). In the second observation, the click train was emitted by the last individual after it approached the device (see Appendix 1, Figure 2; video available on the *Aquatic Mammals* website: www.aquaticmammalsjournal.org/index.php?option=com_content&view=article&id=10&Itemid=147). In the third observation, the click train was emitted by the second individual after it approached the device (see Appendix 1, Figure 3). The same individual produced the four whistles presented in our second sequence; the first whistle was emitted before the approach, and the three others were emitted after leaving. This animal produced no whistles while swimming towards the observers.

Discussion

The results with the BaBeL system are promising for the study of dolphin behavior. Its accuracy using simulated underwater sounds in a pool was validated. The BaBeL design during field testing was verified: the device achieved neutral buoyancy, maneuverability, and simplicity of deployment simultaneously. The BaBeL is relatively easy to deploy from small boats to record behavior and acoustic data on free-ranging dolphins and can also be used with delphinids under human care. Contrary to other hydrophone arrays, the BaBeL system can be used to detect, locate, and track dolphins emitting sounds in a 3D space. The hydrophone arrays of Au & Herzing (2003), Schotten et al. (2004), and Ball & Buck (2005) all present hydrophones in the same plane, making it impossible to discriminate from the audio recording if the emitter dolphin was in front of or behind the device. The design of our system places hydrophones in different planes, allowing us to determine the position of the vocalizing dolphin regardless of its direction of approach to the observers, and the wide-angle HD 360° video cameras provide information to localize to an identified vocalizing dolphin visually. When animals are in the visual range of the camera, this 360° audio-video system could greatly increase the number of vocalizations that can be attributed to an individual dolphin.

Simultaneous visual and acoustic recordings are necessary for localizing to a vocalizing dolphin. This system is mainly limited by visual detection, which depends not only on water clarity but also on the wide-angle video cameras. Wide-angle lenses affect the perspective by exaggerating the distance between objects. They make subjects at moderate and far distances seem further away than they really are. Consequently, only dolphins vocalizing near BaBeL (within 10 m) were visually and acoustically detected. As previously demonstrated by Watkins & Shevill (1974), the accuracy of this acoustic localization system should decrease as the distance of dolphins from the device increases. To improve accuracy, the distance between the hydrophones can be increased, but this would reduce system maneuverability. Therefore, it is recommended that this device only be used in clear water and preferably with dolphin populations habituated to the presence of human swimmers, or with dolphins under human care.

The five possible situations that observers might encounter while using BaBeL are summarized in Appendix 2. Contrary to using a regular camera, an observer using such a 360° video system increases the possibility of capturing ongoing behaviors regardless of his position with respect to the animals and his concentration level, thus reducing human error. In 59% of the detected click trains and in 43% of detected whistles, this device could acoustically and visually detect dolphins, but the customized program did not point to one of the dolphins present on the video, meaning that the dolphin emitting the sound was out of our range or vision. Therefore, in these cases, assigning the recorded sound to any of the dolphins present on the video would have been a mistake, demonstrating the necessity of the use of hydrophone arrays to aid in interpreting dolphins' vocal behavior location to avoid erroneous assumptions about the identity of the vocalizer. When dolphins were detected visually and acoustically, and the customized program pointed to one individual dolphin present in the video, it was found that the echolocating dolphin was never the first one of the group. Moreover, in one case, it was found that one dolphin emitted a click train while its conspecific swimming behind turned its head to observers. This particular dolphin might have been eavesdropping on the returning echoes of its echolocating conspecific (Gregg et al., 2007). Finally, echolocating dolphins swam dorsal side up or ventral side up, suggesting that the animals use different body postures while producing acoustic signals. Further investigations of the body postures dolphins display while vocalizing would enable better understanding of their sound production and their communication system.

The BaBeL system offers a method of data collection to conduct an etho-acoustical analysis of bottlenose dolphin sound emissions potentially to be associated with individual dolphins and their underwater behaviors. The BaBel and the associated software algorithms for data analysis represent an improved tool for ethologists to record and collect data on all dolphins present in a 360° space via focal and group follows.

Acknowledgments

FD's contribution was partially funded by Focused on Nature. Parc Asterix funded previous preliminary fieldwork of FD and JLM and funded TB's work. Materials and fieldwork were financed by the CERECAR project from Abyss Foundation (La Réunion, France). We are very grateful to the Parc Asterix Dolphinarium curator, Birgitta Mercera, and her trainers for their help in testing previous versions of our device within their group of dolphins. We would like to thank Abyss Foundation volunteers and Emmanuel Antongiorgi for their support in collecting data, and Isabelle Charrier and the Bioacoustic Communication Team of NeuroPSI for her constructive comments during the development of this work. We thank Isabella Clegg for revising the English of this manuscript. Finally, we would like to give special thanks to editor Dr. Kathleen Dudzinski for her advice and encouragement to improve this manuscript.

Literature Cited

Agreement on the Conservation of Cetaceans in the Black Sea Mediterranean Sea and Contiguous Atlantic Area (ACCOBAMS), Resolution 4.18. (n.d.). *Guidelines on the granting of exceptions to Article II, paragraph 1, for the purpose of non-lethal in situ research in the agreement area*. Retrieved from www.accobams.org

Alameda-Pineda, X., & Horaud, R. (2014). A geometric approach to sound source localization from time-delay estimates. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *22*(6), 1082-1095. https://doi.org/10.1109/TASLP.2014.2317989

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*, 227-265. https://doi.org/10.1163/156853974X00534

Au, W. W. L. (1993). *The sonar of dolphins*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4612-4356-4

Au, W. W. L., & Fay, R. R. (Eds.). (2012). *Hearing by whales and dolphins* (Vol. 12). New York: Springer Science & Business Media.

Au, W. W. L., & Herzing, D. L. (2003). Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*). *The Journal of the Acoustical Society of America*, *113*(1), 598-604. https://doi.org/10.1121/1.1518980

Ball, K. R., & Buck, J. R. (2005). A beamforming video recorder for integrated observations of dolphin behavior and vocalizations. *The Journal of the Acoustical Society of America*, *117*(3, Pt 1), 1005-1008. https://doi.org/10.1121/1.1831284

Blanchard, T. (2015). *Innovation audio-visuel pour le suivi des grands dauphins* [Audio-visual innovation for monitoring bottlenose dolphins] (Master's thesis). Ecole

Nationale Supérieure d'Ingenieurs du Mans, Le Mans, France.

Brensing, K., Linke, K., & Todt, D. (2001). Sound source location by difference of phase, on a hydrophone array with small dimensions. *The Journal of the Acoustical Society of America*, *109*(1), 430-433. https://doi.org/10.1121/1.1333421

Caldwell, M. C., Caldwell, D. K., & Tyack, P. L. (1990). Review of the signature whistle hypothesis for the Atlantic bottlenose dolphin. In S. Leatherwood & R. R. Reeves (Eds.), *The bottlenose dolphin: Recent progress in research* (pp. 199-234). San Diego: Academic Press.

Connor, R. C., Mann, J., & Watson-Capps, J. (2006). A sex-specific affiliative contact behavior in Indian Ocean bottlenose dolphins, *Tursiops* sp. *Ethology*, *112*(7), 631-638. https://doi.org/10.1111/j.1439-0310.2006.01203.x

Connor, R. C., Mann, J., Tyack, P. L., & Whitehead, H. (1998). Social evolution in toothed whales. *Trends in Ecology and Evolution*, *13*(6), 228-232. https://doi.org/10.1016/S0169-5347(98)01326-3

Connor, R. C., Wells, R. S., Mann, J., & Read, A. J. (2000). The bottlenose dolphin: Social relationships in a fission-fusion society. In J. Mann, R. C. Connor, P. L. Tyack, & H. Whitehead (Eds.), *Cetacean societies: Field studies of dolphins and whales* (pp. 91-126). Chicago: University of Chicago Press.

Dulau-Drouot, V., Boucaud, V., & Rota, B. (2008). Cetacean diversity off La Réunion Island (France). *Journal of the Marine Biological Association of the United Kingdom*, *88*(6), 1263-1272. https://doi.org/10.1017/S0025315408001069

Dudzinski, K. A., Clark, C. W., & Würsig, B. (1995). A mobile video/acoustic system for simultaneous underwater recording of dolphin interactions. *Aquatic Mammals*, *21*(3), 187-194.

Dudzinski, K. M., Gregg, J. D., Ribic, C. A., & Kuczaj II, S. A. (2009). A comparison of pectoral fin contact between two different wild dolphin populations. *Behavioural Processes*, *80*(2), 182-190. https://doi.org/10.1016/j.beproc.2008.11.011

Fellner, W., Bauer, G. B., Stamper, S. A., Losch, B. A., & Dahood, A. (2013). The development of synchronous movement by bottlenose dolphins (*Tursiops truncatus*). *Marine Mammal Science*, *29*(3), E203-E225. https://doi.org/10.1111/j.1748-7692.2012.00609.x

Fripp, D. (2005). Bubblestream whistles are not representative of a bottlenose dolphin's vocal repertoire. *Marine Mammal Science*, *21*, 29-44. https://doi.org/10.1111/j.1748-7692.2005.tb01206.x

Fripp, D. (2006). Bubblestream whistles are not representative of bottlenose dolphin whistle repertoires reply to McCowan. *Marine Mammal Science*, *22*, 496-501. https://doi.org/10.1111/j.1748-7692.2006.00028.x

Gibson, Q. A., & Mann, J. (2008). The size, composition and function of wild bottlenose dolphin (*Tursiops* sp.) mother-calf groups in Shark Bay, Australia. *Animal Behaviour*, *76*, 389-405. https://doi.org/10.1016/j.anbehav.2008.01.022

Gregg, J. D., Dudzinski, K. M., & Smith, H. V. (2007). Do dolphins eavesdrop on the echolocation signals of conspecifics? *International Journal of Comparative Psychology*, *20*(1), 65-88.

Herzing, D. L. (1996). Vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*. *Aquatic Mammals*, *22*(2), 61-79.

Janik, V. M. (1999). Origins and implications of vocal learning in bottlenose dolphins. In H. O. Box & K. R. Gibson (Eds.), *Mammalian social learning: Comparative and ecological perspectives* (pp. 308-326). Cambridge, UK: Cambridge University Press.

Janik, V. M. (2000). Food-related bray calls in wild bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society B: Biological Sciences*, *267*(1446), 923-927. https://doi.org/10.1098/rspb.2000.1091

Janik, V. M. (2009). Acoustic communication in delphinids. In M. Naguib & V. M. Janik (Eds.), *Advances in the study of behavior* (1st ed., Vol. 40, pp. 123-157). New York: Academic Press. https://doi.org/10.1016/S0065-3454(09)40004-4

King, S. L., & Janik, V. M. (2015). Come dine with me: Food-associated social signaling in wild bottlenose dolphins (*Tursiops truncatus*). *Animal Cognition*, *18*(4), 969-974. https://doi.org/10.1007/s10071-015-0851-7

López, B. D., & Bernal-Shirai, J. A. B. (2009). Mediterranean common bottlenose dolphin's repertoire and communication use. In *Dolphins: Anatomy, behavior, and threats* (pp. 129-148). New York: Nova Science Publishers, Inc.

López-Rivas, R. M., & Bazúa-Durán, C. (2010). Who is whistling? Localizing and identifying phonating dolphins in captivity. *Applied Acoustics*, *71*(11), 1057-1062. https://doi.org/10.1016/j.apacoust.2010.05.006

Mann, J. (1999). Behavioral sampling methods for cetaceans: A review and critique. *Marine Mammal Science*, *15*(1), 102-122. https://doi.org/10.1111/j.1748-7692.1999.tb00784.x

Mann, J., Connor, R. C., Barre, L. M., & Heithaus, M. R. (2000). Female reproductive success in bottlenose dolphins (*Tursiops* sp.): Life history, habitat, provisioning, and group-size effects. *Behavioral Ecology*, *11*, 210-219. https://doi.org/10.1093/beheco/11.2.210

Marten, K., Shariff, K., Psarakos, S., & White, D. J. (1996). Ring bubbles of dolphins. *Scientific American*, *275*(2), 82-87. https://doi.org/10.1038/scientificamerican0896-82

Martin, P., Bateson, P. P. G., & Bateson, P. (1993). *Measuring behaviour: An introductory guide*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9781139168342

McCowan, B., & Reiss, D. (1995). Maternal aggressive contact vocalizations in captive bottlenose dolphins (*Tursiops truncatus*): Wide-band, low-frequency signals during mother/aunt-infant interactions. *Zoo Biology*, *14*(4), 293-309. https://doi.org/10.1002/zoo.14301404022

Nowacek, D. P., Tyack, P. L., Wells, R. S., & Johnson, M. P. (1998). An onboard acoustic data logger to record biosonar of free-ranging bottlenose dolphins. *The Journal of the Acoustical Society of America*, *103*(5), 2908. https://doi.org/10.1121/1.422063

Pryor, K. W. (1990). Non-acoustic communication in small cetaceans: Glance, touch, position, gesture, and bubbles. In J. A. Thomas & R. A. Kastelein (Eds.), *Sensory abilities of cetaceans* (pp. 537-544). New York: Springer. https://doi.org/10.1007/978-1-4899-0858-2_37

Quick, N. J., Rendell, L. E., & Janik, V. M. (2008). A mobile acoustic localization system for the study of free-ranging dolphins during focal follows. *Marine Mammal Science*, *24*(4), 979-989. https://doi.org/10.1111/j.1748-7692.2008.00231.x

Sakai, M., Hishii, T., Takeda, S., & Konshima, S. (2006). Flipper rubbing behaviors in wild bottlenose dolphins (*Tursiops aduncus*). *Marine Mammal Science*, *22*(4), 966-978. https://doi.org/10.1111/j.1748-7692.2006.00082.x

Sakai, M., Morisaka, T., Kogi, K., Hishii, T., & Kohshima, S. (2010). Fine-scale analysis of synchronous breathing in wild Indo-Pacific bottlenose dolphins (*Tursiops aduncus*). *Behavioural Processes*, *83*(1), 48-53. https://doi.org/10.1016/j.beproc.2009.10.001

Sayigh, L., Tyack, P., Wells, R., & Scott, M. (1990). Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: Stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, *26*(4), 247-260. https://doi.org/10.1007/BF00178318

Schotten, M., Au, W. W. L., Lammers, M. O., & Aubauer, R. (2004). Echolocation recordings and localization of wild spinner dolphins (*Stenella longirostris*) and pantropical spotted dolphins (*S. attenuata*) using a four-hydrophone array. In J. A. Thomas, C. F. Moss, & M. Vater (Eds.), *Echolocation in bats and dolphins* (pp. 393-399). Chicago: The University of Chicago Press.

Thomas, R. E., Fristrup, K. M., & Tyack, P. L. (2002). Linking the sounds of dolphins to their locations and behavior using video and multichannel acoustic recordings. *The Journal of the Acoustical Society of America*, *112*(4), 1692-1701. https://doi.org/10.1121/1.1494805

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410-433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Tsai, Y-J. J., & Mann, J. (2013). Dispersal, philopatry, and the role of fission-fusion dynamics in bottlenose dolphins. *Marine Mammal Science*, *29*(2), 261-279. https://doi.org/10.1111/j.1748-7692.2011.00559.x

Tyack, P. L. (1991). Use of a telemetry device to identify which dolphin produces a sound. In K. Pryor & K. S. Norris (Eds.), *Dolphin societies* (pp. 319-344). Berkeley: University of California Press.

Van Lancker, E. (2002). *Acoustic goniometry: A spatio-temporal approach* (Doctoral dissertation). Faculté Polytechnique de Mons.

Wahlberg, M., Mohl, B., & Madsen, P. (2001). Estimating source position accuracy of a large aperture hydrophone array for bioacoustics. *The Journal of the Acoustical Society of America*, *109*, 397-406. https://doi.org/10.1121/1.1329619

Watkins, W. A., & Schevill, W. (1972). Sound source location with a three-dimensional hydrophone array. *Deep-Sea Research*, *19*, 691-706. https://doi.org/10.1016/0011-7471(72)90061-7

Watkins, W. A., & Schevill, W. E. (1974). Listening to Hawaiian spinner porpoises, *Stenella cf. longirostris*, with a three-dimensional hydrophone array. *Journal of Mammalogy*, *55*(2), 319-328. https://doi.org/10.2307/1379001

Appendix 1

Etho-Acoustical Description

First Observation—In this sequence (34 s), we analyzed the behaviors of nine adult dolphins (Ind1 to Ind9) and one calf (Ind10) (Figure 1). The first dolphin (Ind1) approached and swam towards the BaBeL for 15 s; and at 7 and 9 s, Ind1 turned its rostrum towards the BaBel and continued swimming towards the observers for 6 s more before leaving. Ind2 approached and swam towards the device for 11 s before leaving, emitting a click train at 4 s. Ind3 swam by the device. Ind4 swam upside down underneath Ind5 for 15 s then moved to a synchronized swimming position above Ind 5. Ind5 swam synchronously for 14 s above Ind4 and then moved to a side swimming posture for 1 s before returning to synchronized swimming below Ind4. Ind6 swam synchronously below Ind7 for 14 s; at 11 s, Ind6 turned its head to observers and continued swimming synchronously below Ind7. At 19 s, Ind6 swam upside down underneath Ind7; at 20 s, Ind6 conducted pectoral fin-to-pectoral fin rubbing to Ind7 for 7 s. At 27 s, Ind6 continued swimming synchronously below Ind7, and Ind7 swam synchronously with Ind6 and touched Ind6 with its pectoral fin on two occasions for 2 s at 14 s on the body and then for 1 s at 26 s on the belly. At 27 s, Ind6 and Ind7 stopped body contact and started synchronized swimming next to each other. Ind8 swam by the observers. Finally, Ind9 and Ind10 swam synchronously by the BaBel next to each other.

Second Observation—In the second sequence (40 s), five adult dolphins passed in front of the BaBeL system. The first and second dolphin (Ind1 and Ind2) approached synchronously. Ind1 swam towards the observers for 4 s, pointed its rostrum towards the recording system with no sound being detected, and left. Ind2 also swam towards the observers for 5 s, pointed its rostrum towards the observers with no sound being detected, and left. Following behind the first two dolphins, Ind3 whistled, swam towards the observers for 2 s, and left. After leaving, Ind3 emitted three more

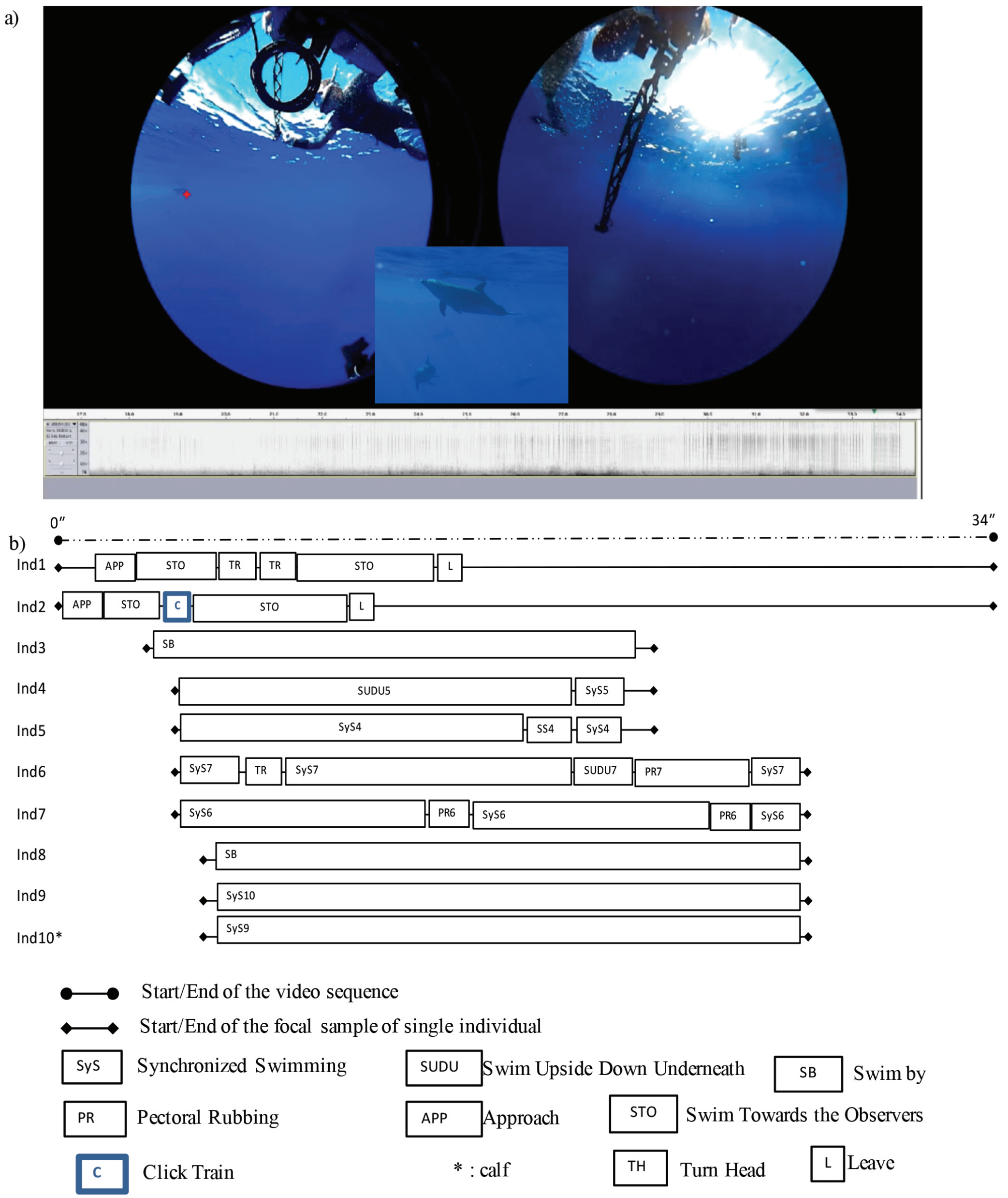

Figure 1. *First observation:* (a) Screenshot of the location of the clicking dolphin (Ind2) in 360° video and backup video with the red cross pointing to the source of the sound emission; and (b) timelines for ten individuals.

whistles before disappearing out of the range of vision of the video system. Ind4 swam by the BaBel, while Ind5 approached and swam towards the observers for 4 s, emitted a click train, and continued swimming towards the observers for 5 s before leaving (Figure 2).

Third Observation—In this 22 s sequence, six adult dolphins passed from left to right in front of the BaBeL (Figure 8). Ind1 swam less than 50 cm distance above Ind2. Ind2 swam upside down underneath Ind1 for 16 s; at 3 s, Ind2 turned its rostrum towards the observers and, at 8 s, emitted

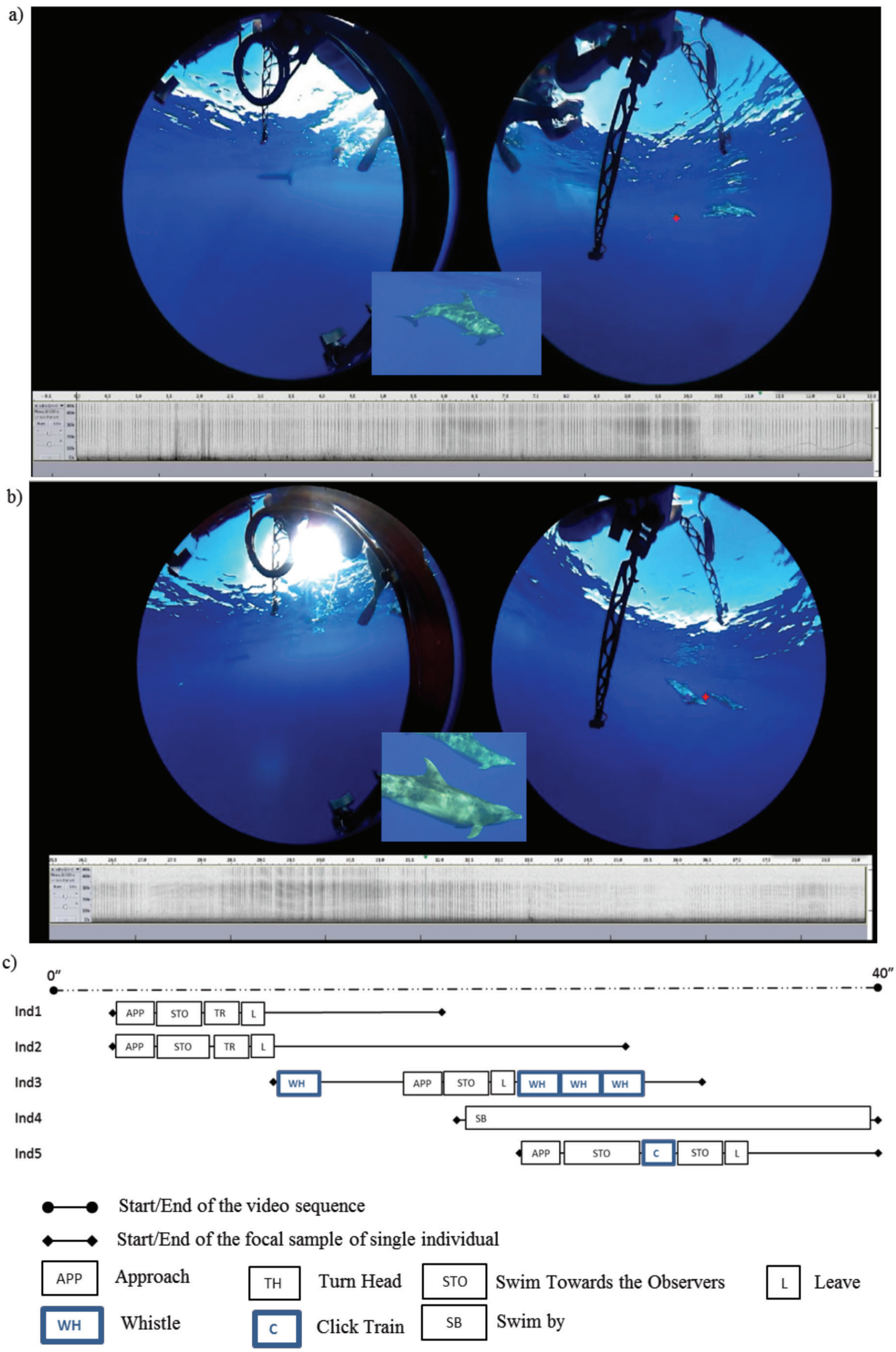

Figure 2. *Second observation:* (a) Screenshot of the location of the whistling dolphin (Ind3) in 360° video and backup video with the red cross pointing to the source of the sound emission; (b) screenshot of the location of the clicking dolphin (Ind5) in 360° video and backup video; and (c) timelines for five individuals (Ind1 to Ind5).

a click train. At 11 s, Ind2 rubbed Ind1's belly with its pectoral fin for 8 s. At 19 s, Ind2 stopped its contact with Ind1 and swam synchronously above it until the end of the sequence.

Ind3 swam by the BaBel, and Ind4 synchronously swam next to Ind5 and Ind6. Ind5 swam upside down underneath Ind4 for 4 s and then continued swimming synchronously with Ind4. Ind4 pectoral fin rubbed Ind5's belly for 1 s at 3 s and then continued swimming synchronously with Ind5 and Ind6 (Figure 3).

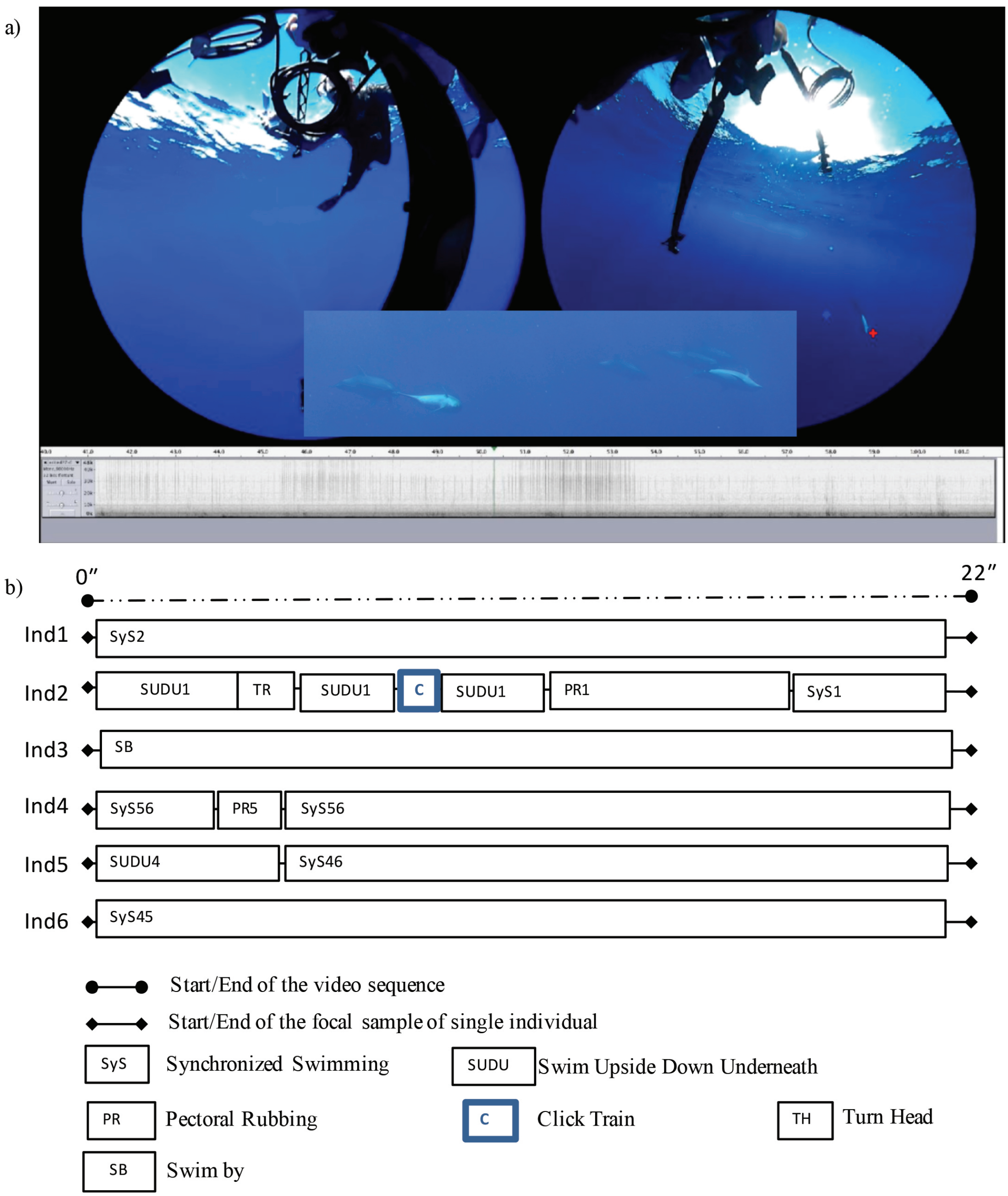

Figure 3. *Third observation:* (a) Screenshot of the location of the clicking dolphin (Ind3) in 360° video and backup video with the red cross pointing to the source of the sound emission; and (b) timelines for six individuals.

Appendix 2

The five possible situations observers could encounter while operating the BaBeL.

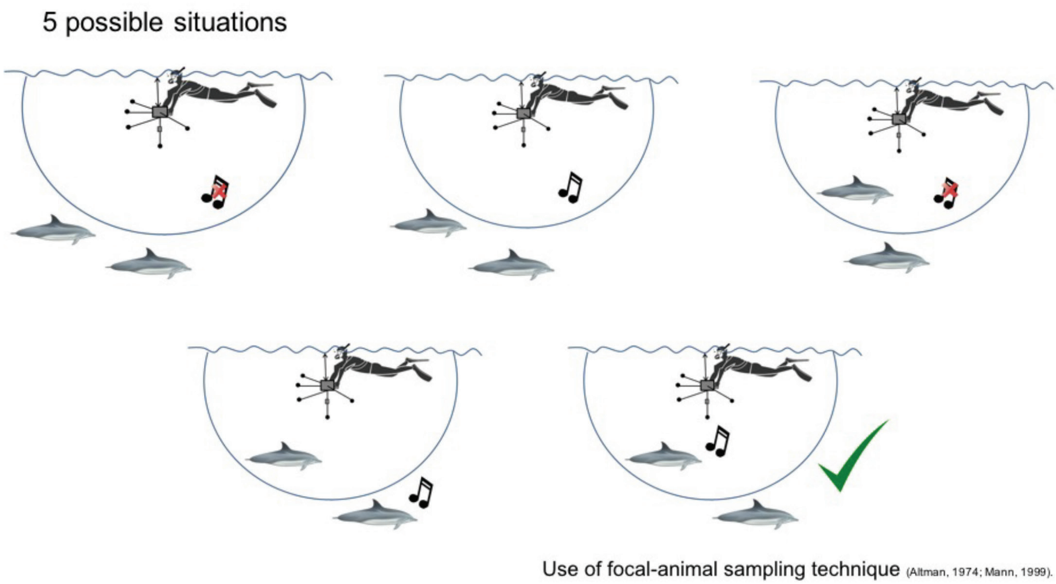

Figure 1. Use of focal-animal sampling technique. *Top left:* No sound detected and dolphins out of the camera's visual range; *Top middle:* Sound detected but dolphins out of the camera's visual range; *Top right:* No sound detected but dolphins in the visual range of the camera; *Bottom left:* Dolphin in the visual range of the camera and sound detected, but the dolphin vocalizing is not present in the video; and *Bottom right:* Dolphin in the visual range of the camera, with the sound detected and the vocalizing dolphin present in the video.