Year: 2023

Full-duplex acoustic interaction system for cognitive experiments with cetaceans

Rychen, Jörg ; Semoroz, Julie ; Eckerle, Alexander ; Hahnloser, Richard H R ; Kleinberger, Rébecca

Abstract: Cetaceans show high cognitive abilities and strong social bonds. Their primary sensory modality to communicate and sense the environment is acoustics. Research on their echolocation and social vocalizations typically uses visual and tactile systems adapted from research on primates or birds. Such research would benefit from a purely acoustic communication system to better match their natural capabilities. We argue that a full duplex system, in which signals can flow in both directions simultaneously is essential for communication research. We designed and implemented a full duplex system to acoustically interact with cetaceans in the wild, featuring digital echo-suppression. We pilot tested the system in Arctic Norway and achieved an echo suppression of 18 dB. We discuss the limiting factors and how to improve the echo suppression further. The system enabled vocal interaction with the underwater acoustic scene by allowing experimenters to listen while producing sounds. We describe our motivations, then present our pilot deployment and give examples of initial explorative attempts to vocally interact with wild orcas and humpback whales.

DOI: https://doi.org/10.1075/is.22039.ryc

Originally published at:
Rychen, Jörg; Semoroz, Julie; Eckerle, Alexander; Hahnloser, Richard H R; Kleinberger, Rébecca (2023). Full-duplex acoustic interaction system for cognitive experiments with cetaceans. Interaction Studies : Social Behaviour and Communication in Biological and Artificial Systems, 24(1):66-86.
DOI: https://doi.org/10.1075/is.22039.ryc

Full-duplex acoustic interaction system for cognitive experiments with cetaceans

Jörg Rychen[1], Julie Semoroz[2], Alexander Eckerle[3], Richard HR Hahnloser[1], and Rébecca Kleinberger[4,5]
[1] University of Zurich and ETH Zurich | [2] Association Motonomy | [3] Ludwig-Maximilians-University of Munich | [4] MIT Media Lab | [5] Northeastern University

Cetaceans show high cognitive abilities and strong social bonds. Their primary sensory modality to communicate and sense the environment is acoustics. Research on their echolocation and social vocalizations typically uses visual and tactile systems adapted from research on primates or birds. Such research would benefit from a purely acoustic communication system to better match their natural capabilities. We argue that a full duplex system, in which signals can flow in both directions simultaneously is essential for communication research. We designed and implemented a full duplex system to acoustically interact with cetaceans in the wild, featuring digital echo-suppression. We pilot tested the system in Arctic Norway and achieved an echo suppression of 18 dB. We discuss the limiting factors and how to improve the echo suppression further. The system enabled vocal interaction with the underwater acoustic scene by allowing experimenters to listen while producing sounds. We describe our motivations, then present our pilot deployment and give examples of initial explorative attempts to vocally interact with wild orcas and humpback whales.

Keywords: human–dolphin communication, self-interference cancellation, echo cancellation, killer-whale, orca, humpback whale, inter-species communication, animal communication, bioacoustics, underwater communication, animal cognition

Introduction

Delphinidae have long been a subject of interest for researchers due to their outstanding cognitive abilities and communicative skills (Herzing & Johnson, 2015; Marino, 2002). Studying marine mammals through direct human inter-

https://doi.org/10.1075/is.22039.ryc |
Interaction Studies 24:1 (2023), pp. 66–87. ISSN 1572-0373 | E-ISSN 1572-0381

actions presents unique challenges, especially with wild animals. The aquatic environment is restrictive of sustained interactions, and the air/water barrier hampers acoustic communication. Early experiments in 1978 by John Lilly (Lilly, 1978) aimed to establish human–dolphin communication by co-housing a human experimenter and a dolphin in a specially designed house, but the scientific outcome remained limited. A few years later, experiments showed that dolphins can understand the syntax and semantics of visual and acoustic commands (Herman et al., 1984), suggesting complex auditory learning through reinforcement by food rewards. Vocal production learning and the ability to label objects have also been demonstrated (Richards et al., 1984). These observations led to research endeavors using custom-built keyboards and touchscreens for the animals. Such paradigms were often adapted from established techniques used for primates and birds. An extensive review of such systems by Denise Herzing (Herzing, 2016) provides various insights into the design of technologies for communicating with Delphinidae. Her analysis advocated for approaches that keep the human in the loop, rather than aiming for subjective cognitive assessment. She also recommends approaches based on sound, rather than other senses, as sound is a more natural modality for the animals. She highlighted the lack of tools adapted for underwater communication, emphasizing the importance of two-way audio.

Recent studies have since made stronger use of audio, for instance by using a hydrophone array within a large visual screen such that the dolphins could "point to" a location on the screen with an acoustic click (Amundin et al., 2008). Denise Herzing studied wild habituated dolphins by swimming with them and interacting with them through a device that could play back and detect a set of preprogrammed whistles (Kohlsdorf et al., 2013). These experiments addressed the ability of dolphins to use referential signals for objects or actions. By using a two-way acoustic interaction system, the researchers were able to conduct these experiments without reinforcement by food reward but were based solely on the curiosity and playfulness of the animals. However, fully featured acoustic communication requires a full-duplex system, where signals can flow in both directions simultaneously instead of alternatively, as explained below.

Acoustic communication implies for a given living or technological system to possess both an organ for emitting sound and one for perceiving sound. Inevitably the own emitted sound will be sensed by the sensory organ potentially leading to issues in distinguishing external from the self-produced sounds. Indeed, the self-produced sound is also typically quite loud due to physical proximity to the source, posing a challenge for the dynamic range of the sensory organ, the range of sound intensities that can be processed. For living creatures, this problem is addressed anatomically, for example by placing the ears behind the sound emitting organ and by directing the emitted sound forwards. Second, the own sound needs to be distinguished from the external sound by specific neural processes specifically involved in self-generated stimuli. The neurological underpinnings of this very complex problem are still under research (Haykin & Chen, 2005). In contrast to the biological solutions, technological acoustic communication systems sometimes use a trick to only let the signal flow one way at a time. In this approach, called half-duplex, although communication is possible in two ways, the direction is actively switched such that the signals flow only one way at a time. The walkie-talkie is a prominent example of this, whereby a push-to-talk button enables the sender and disables the receiver of the device. When communicating over a walkie talkie, both parties need to switch the signal direction and often a protocol is implemented, for example by saying «over» to signal to the other party to switch direction. However, when communicating with multiple individuals the protocol can become more complex. While half-duplex systems avoid both problems of dynamic range and separating the own voice, it hampers vocal communication significantly. Observed phenomena like turn taking and chorusing require to be able to listen while vocalizing to detect collisions or to synchronize and tune the pitch. Therefore, a full-duplex system where signals can flow in both directions simultaneously is needed to more precisely make sense of acoustic communication. For a full-duplex system it is essential to implement a circuit that cancels or at least attenuates the perception of the self-produced signal (Figure 1).

Echo suppression (also known as "echo canceling" or "self-interference suppression") is a signal-processing technique developed to separate both directions of a signal (Sondhi, 1967). It is a key technology for online conference applications and part of mobile phones. Echo suppression systems model the transfer of the 'own emission' to the 'own reception 'for example with an impulse response function (IRF), which is the response to a virtual, infinitely sharp impulse in the time domain (Antsaklis, 2021). The transferred signal can be estimated by a convolution of the emitted signal by the IRF (Figure 1) and subtracted from the mixture signal to yield the external signal more clearly. Such echo suppression systems are typically implemented by adaptive filtering, for example with the least-mean-squares (LMS) algorithm and can be tuned, for example, by playing white noise (Rychen et al., 2021).

Although full-duplex acoustic communication systems with echo suppression are widely used in underwater applications typically for long-range underwater data transmission (Shen et al., 2020), to the best of our knowledge, they have not yet been applied in research on either human–Delphinidae communication or cetaceans' cognitive abilities.

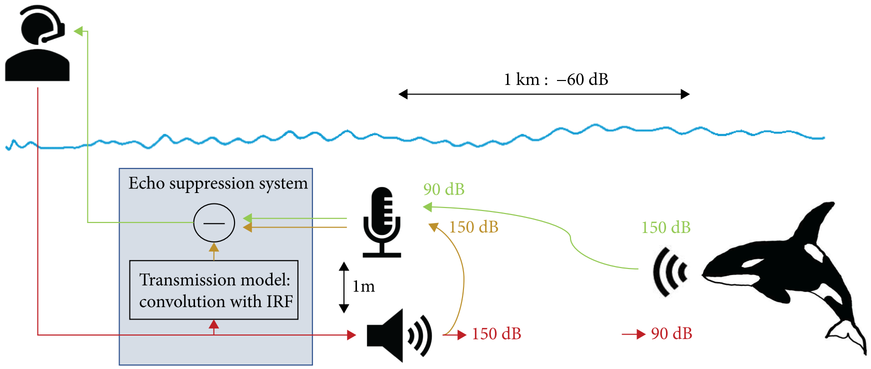


Figure 1. Full-duplex communication system with echo suppression. The signal that is broadcast over the speaker (red) is transferred acoustically to the hydrophone (yellow), where it is superimposed over the external signal of interest (green). Echo suppression models the transfer of the signal from speaker to microphone with an impulse response function (IRF). This modeled signal is subtracted from the raw microphone signal to yield a pure external signal (green). The situation shows an example to communicate with an animal 1 km away, where both parties emit at a level of 150 dB (re 1 µPa). The propagation attenuation is 60 dB and the own signal is therefore much stronger than the distant signal. The effect of the echo suppression system is to filter out as good as possible the own signal to allow to increase the gain for the signal of interest

We argue that full-duplex communication systems would enhance existing research protocols and enable new research applications with cetaceans. In the following sections, we review four classes of experiments that could benefit from full-duplex audio interactions.

Interactive playback experiments (Figure 2A)

Playback experiments are an important tool in research of animal communication and serve to validate hypotheses on the function or meaning of acoustic signals (Deecke, 2007). Although these experiments often require ethical considerations to avoid deception, they help to demonstrate the animal's ability to discriminate between classes of signals (Filatova et al., 2011; Curé, 2019). While these experiments have been principally one-way *interactive* playback experiments, as highlighted in a review article by Stefanie King (King, 2015), they promote research on communication and cognitive tasks. In this context, *interactive* generally means that the played-back signal is chosen dependent on the animal's behavior. For example, the experimenter may observe or listen to the animal and then decide which prerecorded or synthesized stimulus to play back and when to do it. A full-duplex system would be useful in such experimental approaches because they provide better access to turn-taking behavior, to choral behavior, and to simultaneous vocalization.

Phantom echo (Figure 2B)

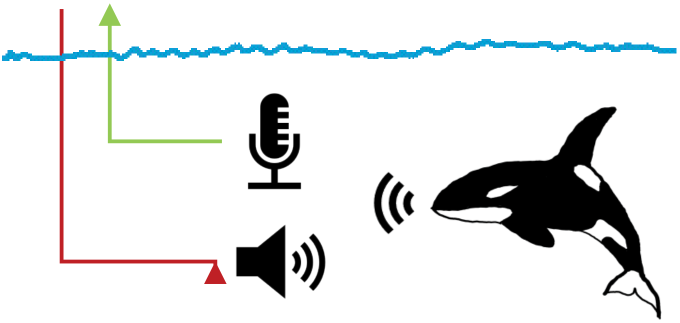
To study the echolocation abilities of dolphins, researchers have used the technique of phantom echo (or "artificial sonar target") (Aubauer & Au, 1998; Muller et al., 2007), whereat an incoming signal, i.e., a single click of the animal under investigation, is processed with low latency, and a simulated acoustic reflection of a virtual target is played back to the animal. This approach has been used, for example, to study how dolphins can discriminate between metallic cylinders whose walls differ in thickness only by a fraction of a millimeter (Branstetter, 2020). In their setup, to prevent repeated echoes, the transducer was placed much closer to the dolphin than the receiving hydrophone. The dolphin also had to remain at a precise position defined by a bite plate. In a related experiment, dolphins were able to detect jittering distance variations of an artificial target (Finneran, 2020). An experiment to test auditory stream segregation in actively swimming dolphins (Malinka, 2021, Chapter VI) used phantom echoes triggered by on-axis clicks toward the target. The system was implemented on a field-programmable gate array (FPGA) with low and precise latencies and used a "push-to-talk" technique, by blanking the received signal during the emission of the phantom echo, to avoid loops of repeated echoes. This limited the repetition rate of clicks for the dolphin. Full-duplex systems could enable phantom echo experiments with free moving animals and without restrictions on repetition rate, paving the way for more research on the echo location capabilities of dolphins, even in the wild.

Animal telecommunication (Figure 2C)

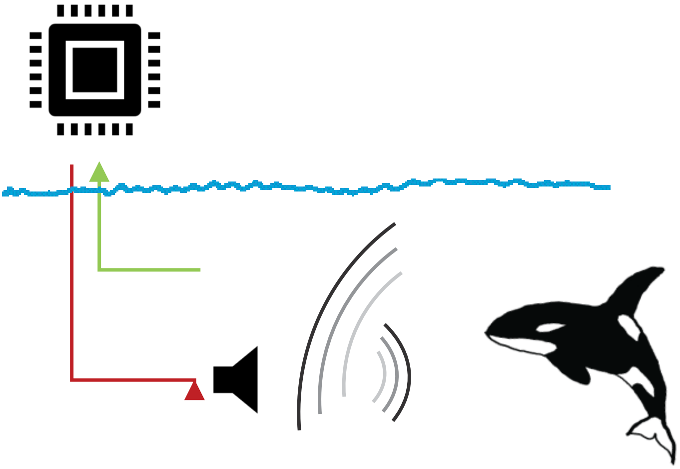
Dolphins use vocalizations and clicks to communicate between con-specifics. In 1965, an experiment recorded the communication between two dolphins kept in separate tanks, connected with an electronic communication system (Lang & Smith, 1965). This system had no echo-suppression, and the receiver and transmitter had to be placed far apart inside each pool to minimize the echo and prevent feedback oscillations. A squelch circuit that disables transmission when the signal is below a threshold was missing in this setup and from the background noise the dolphins could infer when the system was turned on. We argue that the technology of echo suppression presented in this paper could lead to significant improvements in experiments on the communication of dolphins. For instance, dolphins performing synchronous aerial jumps have been observed to coordinate their behavior with click trains (Marulanda et al., 2021). Could they synchronize their jumps in remote locations while connected with a telecommunication system? It has also been shown that two dolphins together could solve cooperative tasks (King et al., 2021). Could they collaborate remotely through a telecommunication channel? What kind of information would they be able to share with each other to solve increasingly complex problems? An experiment to test the use of referential labels in cooperating dolphins was proposed by Rossa et al. (2022). Full-duplex systems have provided some answers to such questions in a study of the influence of social learning on auditory discrimination task in songbirds (Narula et al., 2018). Also, animal telecommunication systems allow for modifying signals in real time to test hypotheses about the roles of auditory features for communication. The modifications could include adding noise, delays, or filtering out of individual links in communication networks (Rychen et al., 2021). In the wild, under water telecommunication systems could be a means to assist entrapped cetaceans (Jourdain, 2021) to find their way back to their conspecifics by establishing an acoustic link between the entrapment bay and the outside group members.

Human–Animal interaction (Figure 2D)

In the pilot deployment described below, a direct human–animal interaction was used to assess the performance of our system and to explore our ability to capture the animals' attention. In contrast to the interactive playback paradigm (Figure 2A), human–animal vocal interaction is technically simpler and more flexible. While a computer program offers a high degree of repeatability, a human in the loop shifts the research focus to softer factors of communication, such as curiosity, bonding, and surprise. In 2008, Rothenberg played clarinet in accompaniment with a singing male humpback whale and suggested interspecies music-making as a potential tool in helping understand the complex communication strategies of cetaceans (Rothenberg, 2008). Spontaneously mimicking of human vocalization has been observed in cetaceans and reports suggest mutual curiosity to be a driving factor (Ridgway et al., 2012). In addition, having the human reactivity in the loop could help with studying phenomena such as joint rhythm and turn taking (Pika et al., 2018) that are still difficult to handle with computer programs. A full-duplex communication system is needed to study these phenomena in a shared communication channel.



A.



B.

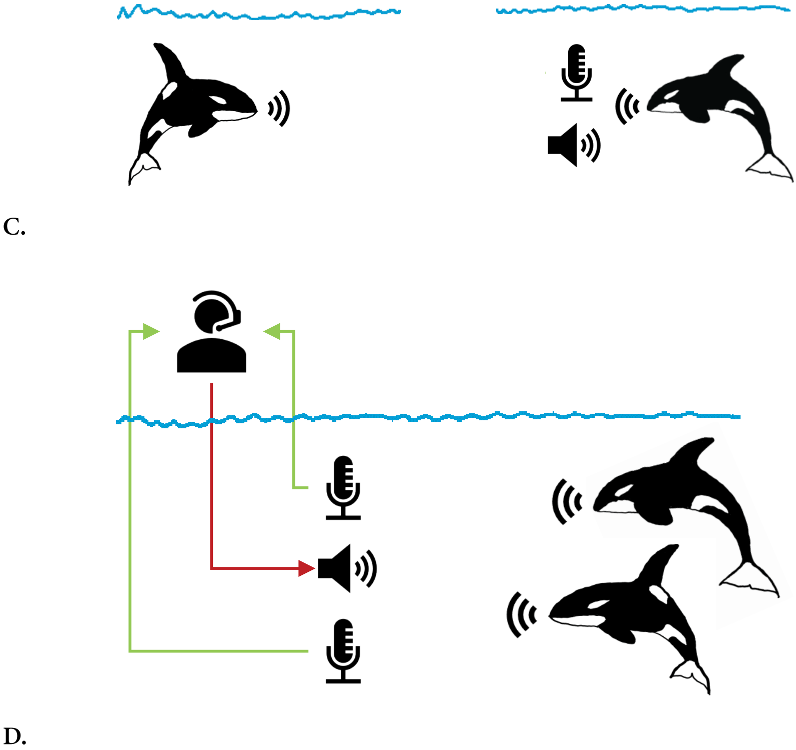


Figure 2. Four classes of acoustic interactive experiments with cetaceans that can benefit from a full-duplex system: (**A**) Computerized interactive playback experiments. Prerecorded acoustic stimuli are played back depending on the behavior of the animal. (**B**) Phantom echo (artificial sonar target). A low latency digital signal processing unit detects incoming echolocation clicks, and then emits a computed echo to simulate a virtual target. (**C**) Telecommunication between two animals to study their communication. This approach provides separated recordings of individuals, the signals can be modified or filtered in real-time, and additional signals can be injected. (**D**) Direct human – animal vocal interaction. This setup allows humans to interactively experience the underwater acoustic space

Based on the insights presented above, we conclude that researching the acoustic communications of cetaceans needs technology that allows for a full duplex signal flow. We therefore set out to implement a first full duplex system for underwater acoustic interactions, and to test it in the real scenario, using it from a boat to interact with killer whales and humpback whales in the fjords of northern Norway.

Pilot experiments

We designed, developed, and field-tested a system for acoustic full-duplex communication between a human experimenter and an underwater acoustic space (see Figure 2D). The system was first tested in 2020 and then deployed in 2021 in Norway in waters frequented by killer whales and humpback whales for the initial observation of their reactions and behaviors.

Through a stereo headphone, the experimenter could listen to the underwater sounds captured by two separate hydrophones in the water. The experimenter's vocalizations were broadcast over an underwater speaker located approximately six meters below the boat. Two echo-suppression filters reduced the signals of the direct sound transmission from the speaker to the hydrophones to attenuate the experimenter's own sounds, to allow for more gain and provide better perception of faint distant sounds, and reverberations in the fjords.

The goal of this initial pilot deployment was to collect insights regarding the following three objectives:

1. Assess the performance of our echo-suppression system for underwater acoustics by measuring the suppression of the direct echo in decibels (dB) and identify points for improvement.
2. Listen to wild animals' behaviors and attempt to react live to their vocalization. Observe potential initial signs of curiosity or reaction of the animals to human-produced sounds.
3. Subjectively report on human's active experience of the underwater acoustic scene.

System

The system comprised an underwater loudspeaker with a fixed array of four hydrophones (Figure 3B). The electronic system was mounted inside a transport case (Figure 3C), including a battery, analog amplifiers, and a digital signal processing system based on FPGAs (Figure 4). To ensure high dynamic range, low noise, and good linearity, we used instrument-grade analog to digital converters (ADC) and digital to analog converters (DAC) designed for audio applications.

The vocalizations of the human experimenter were either captured with a directional microphone or a contact microphone in the form of a piezo disk pressed to the throat.

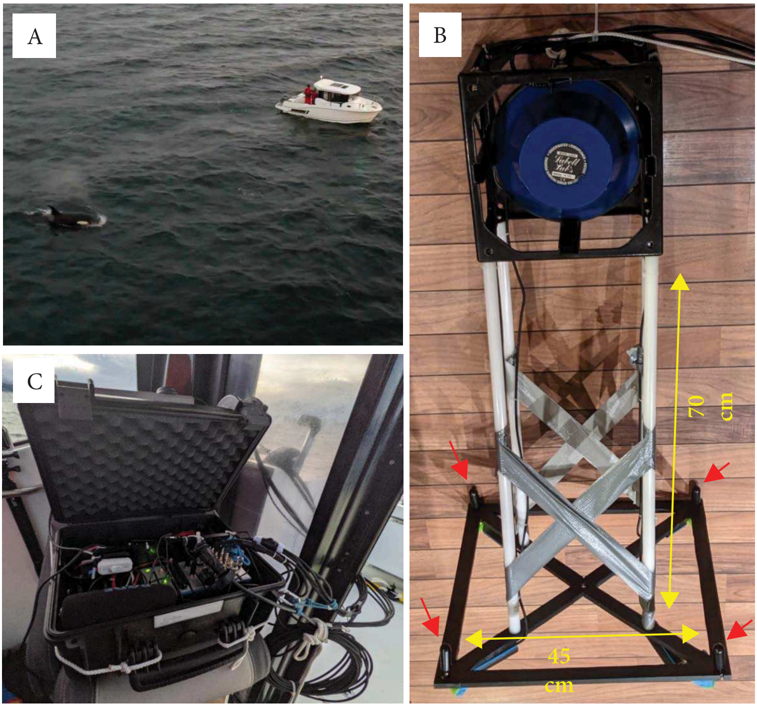


Figure 3. (**A**) We installed the full-duplex interaction system on a small motorboat. The experimenter wore a headset and interacted with the underwater sonic environment. (**B**) We mounted an underwater loudspeaker (LL916, Lubell Labs, USA) and four hydrophones (HTI-96, High Tech Inc., USA, red arrows) 70 cm below the speaker to a plastic frame made of polyethylene (PE-HD), that minimizes acoustic disturbances). The speaker was lowered overboard to about 6 m into the water. (**C**) To protect the electronics from saltwater splashes, they were placed inside the cabin of the motorboat

We implemented echo suppression with two adaptive LMS filters of 512 samples length, corresponding to 10.24 ms and to 15.36 m sound propagation distance. The filters were tuned by playing white noise for a few seconds at a source level of 145–155 dB (re 1 μPa rms at 1 m) in the 200 Hz–20 kHz band. The normalized tuning rate, a parameter that specifies how fast the adaptive filter converges, was typically 0.01, see (Rychen et al., 2021) for details. We calibrated the loudspeaker and its amplifier based on the frequency-averaged sound-pressure measurement of 184 dB re µPa/V at a distance of 1 m in the axial direction of the speaker in the band of 200 Hz–20 kHz.

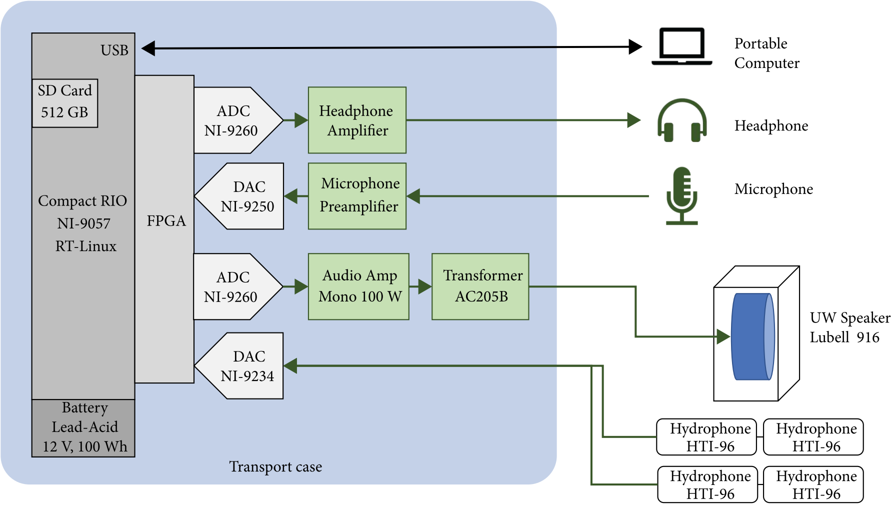


Figure 4. Overview of the full-duplex system architecture. We used a controller (CompactRio, NI-9057, National Instruments, USA) with a real-time Linux operating system. The integrated FPGA interfaces with input and output (IO) modules and runs the signal processing including the adaptive LMS filters for echo-suppression. Digital-to-analog converters (DAC) and analog-to-digital converters (ADC) with 24-bit sigma-delta converters were run at a sample rate of 51.2 kHz. All IO modules were driven by a common clock for synchronized sampling of all channels. The two output channels of the first NI-9260 module were amplified and connected to an analog stereo headphone. A directional condenser microphone was connected to a microphone preamplifier and then to the analog input module NI-9250. The underwater loudspeaker (Lubell Labs, USA) was driven by a mono 100-W class-D amplifier (TPA3116D2, Texas Instruments), followed by an impedance matching circuit (Transformer Model AC202B, Lubell Labs, USA). The signal was generated by one channel of the second NI-9260 module. We used four hydrophones (High Tech Inc., USA) with integrated ICP preamplifiers. These were connected directly to the four input channels of the NI-9234 module, which delivered an excitation current of 2 mA to each of the preamplifiers. We used the factory calibration measurement for hydrophones with sensitivities of −170 dB re 1 V/µPa (+/− 0.3 dB). The system used about 10 W and the 12 V, 100 Wh battery allowed for a run-time of at least 6 h. All signals on the FPGA were saved to a file on an SD-card. Live inspection of the signals and device control was made possible thanks to a user interface running on a portable computer. All programming was done in LabVIEW (National Instruments, USA)

Field test

We tested the system in December 2021 in fjords near the village of Skjervøy in northern Norway (N 70.0336, E 20.9880), during the season where many humpback whales (*Megaptera novaeangliae*) and killer whales (*Orcinus orca)* regularly feed on herring (*Clupea harengus*). Using a small motorboat (Figure 3A), we conducted daily excursions looking either for animals or a quiet place. Before recording, we stopped the boat and turned off the engines and the depth finder, lowered our hydroacoustic system into the water, and tuned the echo-suppression filters. Thereafter, wearing a stereo headphone, the experimenter could listen for underwater sounds, including the echoes from self-generated sounds broadcast through the speaker.

Results and discussion

We performed 18 boat excursions, lasting between 3 and 7 hours and deployed the system in total 91 times. From all deployments we could Extract 6 sessions of total 46 minutes in which animal vocalizations were clearly detectable and the ambient noise level from boats and waves were low enough such that the signals of interest could be analyzed by human listening and by graphic inspection of the spectrograms.

Echo suppression performance

We established an echo suppression of 18 dB, measured with white noise within the 200 Hz to 20 kHz band. The impulse response function is shown in Figure 5A. The impulse response function can be interpreted as the signal in response to a very short pulse with a defined intensity.

The impulse response function in Figure 5A is zero for the first 0.5 ms, which agrees well with the 70 cm distance between the hydrophones and the speaker (we corrected an offset of 1.388 ms to account for the specified delay of the ADC and DAC and an internal delay of four samples (80 μs) in our signal processing on the FPGA).

Figure 5B shows a hypothetical situation with an animal at a 1km distance, corresponding to an attenuation of 60 dB. Typically, orcas vocalize with a source level of 150–170 dB (Holt et al., 2011), yielding a sound pressure level of 90–110 dB at the hydrophone. The source level of our speaker was about 140–160 dB, arriving at the animal's location at a level of 80–100 dB. Our echo suppression of 18 dB reduced the level of our own signal to 122–142 dB and thus experimenters typi-

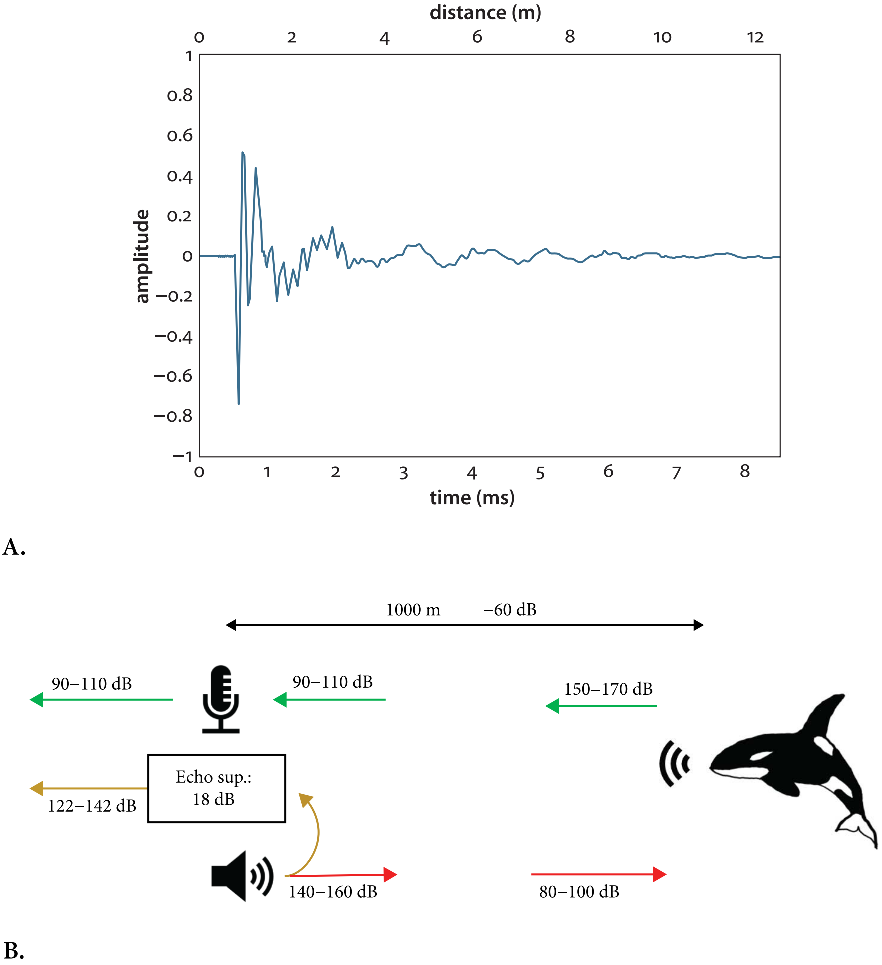


Figure 5. Experimental echo suppression performance. (**A**) The impulse response function after tuning the adaptive filter with white noise. The curve represents the 512 filter coefficients of the adaptive filter. The lower x-axis indicates the time delay, and the upper x-axis indicates the corresponding travel distance of the sound waves ($c = 1475$ m/s). (**B**) Diagram of sound levels (rms re 1 µPa) in a typical situation with an animal 1 km away. With an echo-suppression of 18 dB, the own sounds on the headphones are louder than those of the remote animal

cally heard their own voice louder than that of the animals. This limited the gain for the headphones to prevent pain in the ears and thus limited our ability to listen to distant animals.

The achieved echo suppression of 18 dB proved less effective than expected. The reason for the limitation in performance is unclear in this case. Typically, nonlinearities of the amplifier and the speaker cause harmonic distortions, which are the main limiting factors of echo suppression systems (Ling, 2014). A major improvement can be expected by sensing the speaker's signal with an additional analog-to-digital converter. This signal could be used in the regression of the adaptive filter, circumventing the non-linearity of the power amplifier and transformer. It has been shown that an echo-suppression of 69 dB can be achieved in this way (Shen et al., 2020). Another factor limiting the echo suppression is the time-varying impulse response function from a moving sea surface and was addressed in ref (Shen et al., 2020). Yet another improvement could be to use a more rigid frame for the loudspeaker and hydrophones (see Figure 3) to avoid mechanical resonances.

This pilot experiment allowed us to assess echo-suppression performance with a mobile system in a realistic experimental situation in a harsh environment.

Human experience of underwater soundscape

One motivation behind the development of the echo suppression system was to enable a new form of underwater acoustic perception for humans. The full-duplex system allows for experimenters to be vocalizing and clearly perceiving the echoes and reverberations of their voices within the fiords. The deep and granite-walled fjords of Arctic Norway form an unusual underwater acoustic scene. The echo-suppression filters attenuated echoes from reflections up to 6 m around the speaker (Figure 5A), but further echoes remained unchanged. By using two of the four hydrophones for the headphones, we obtained a stereo percept. The speed of sound being five times faster in water than in air, therefore, from the perspective of interaural time differences, placing the hydrophones 1 m apart in water corresponds to the roughly 20 cm distance of human ears. To obtain also interaural intensity differences, one would need a heavy and five times larger model of a human head with hydrophones in the ears (e.g., a concrete sculpture). However, computational methods exist to simulate head-related transfer functions, and underwater spatial audio has been demonstrated with a four-hydrophone array (Delikaris-Manias et al., 2018).

The limited echo-suppression impeded on the perception of distant echoes. The direct signal from the speaker to the hydrophone was strong, limiting the gain at which we could listen to the hydrophones. For a more immersive experience of the underwater world, further improvements to the echo-suppression system are needed, which seems feasible (Shen et al., 2020).

Making the underwater acoustic world perceivable to humans could be an important factor for raising awareness of the unique acoustical world inhabited by these animals, who evolved to use sound as their primary sense. Acoustic pollution is a major thread for killer whales and other cetaceans (Jourdain, 2019).

Vocal interactions with orcas and humpbacks

One important motivation of this research is to enable direct acoustic streaming between cetaceans and human experimenters. During the pilot deployment of the system, we seek initial interactions with wild animals.

We aimed to engage orcas or humpback whales in interaction by imitating their sounds and adapting to their vocal rhythms. The goal was to address an animal by providing a stimulus that was not random but that had a correlation with its own vocalizations. We expected to observe interest of the animals and playful exploration of the new sound source due to their natural curiosity.

Although this pilot experiment aimed at testing the new system in realistic conditions and was not designed to gain scientific insights into the animals' reaction, our initial observations may suggest hints of reaction of the animals to our sounds. Those initial observations are reported below.

Figure 6 shows four example spectrograms with animal vocalizations and human sounds in different colors. Subfigures A and C show interactions with humpback whales, and subfigures B and D show interactions with orcas.

The humpback whales in northern Norway become vocally very active by the end of December. They produce rhythmic sounds, but the sequences are not as stereotyped as those songs on the breeding grounds (Handel et al., 2012; Magnúsdóttir & Lim, 2019). They repeat characteristic vocal units, typically 2–6 times with gaps of about 2 s between each repeated unit. In some of our encounters, when we tried to produce imitations of their vocalizations during the gaps, they repeated the same unit more than 10 times.

For the orcas, we preferred to experiment with them immediately after a feeding event, when they engage in social and playful behaviors, and disperse into smaller groups, while vocalizing extensively. Their whistles are typically up-sweeps, with a few rare instances of reported down-sweeps calls. In a call-type catalog for Norwegian killer whales (Shapiro, 2008), we found only one call, the N64, with a clear down-sweep frequency contour. In one of our encounters, the experimenter was whistling with down-sweeps and the orcas engaged in a chorus of down-sweep whistles that lasted several minutes. A spectrogram is depicted in Figure 6B. Figure 6D shows an encounter with a juvenile and an adult orca, probably a mother-calf pair. The pair approached and circled our boat seemingly investigating our acoustic device in the water with direct click trains, often directly after we broadcast a vocalization. This may suggest that the orcas investigated the new unknown sound source in the water. More investigations based on scientific protocols are needed to validate these observations. Future work includes the identification of measurable factors and protocol designs using the system focused on call timing, acoustic characteristics, animal location and movements, etc.

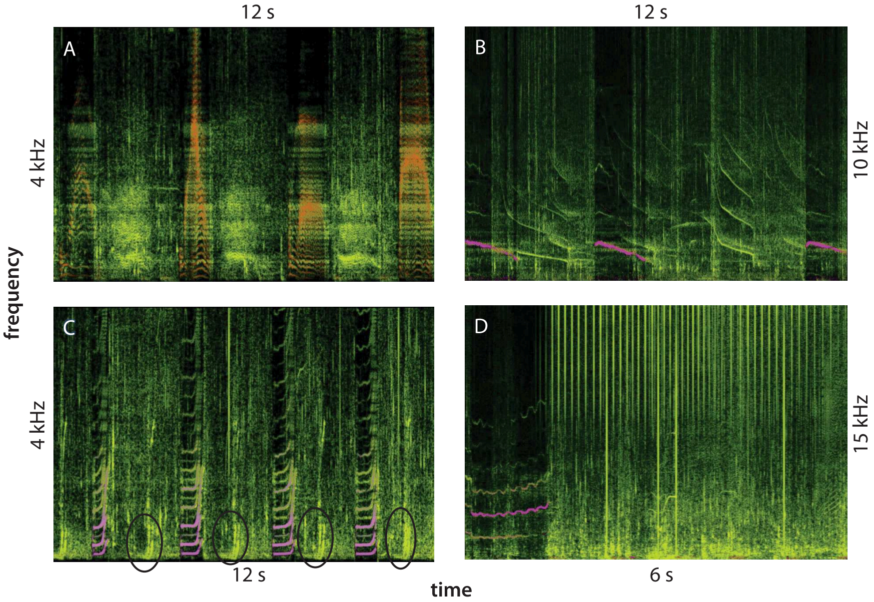


Figure 6. Four example spectrograms of interactions with humpback and killer whales. The horizontal axis is time, with extent indicated in seconds. The vertical axis is linear in frequency starting from zero with extent indicated in kHz. Green-yellow is the signal from the hydrophones and indicates all background sounds including animal vocalization. Overlaid in purple is the human signal broadcast through the speaker. (**A**) Humpback growls alternating with human quacks. (**B**) Down sweep chorus with orcas. **C**: Rhythmic humpback whups (encircled) alternating with human whup imitations. (**D**) Orca click trains presumably aimed at our acoustic system right after a human high-frequency sound. Some of the clicks overloaded the signal (vertical lines down to the lower edge). For audiovisual media: https://youtu.be/KWZFDMFLUfA

Conclusion

We presented the design and pilot deployment of a full-duplex communication system to research cetaceans' communication abilities. Our system provides a bidirectional acoustic link into the water. At the heart of the system is the echo-suppression that achieved a performance of 18 dB. We tested the system in northern Norway and made an initial exploratory attempt to vocally interact with wild orcas and humpbacks aiming to engage the animals by imitating their sounds and adapting their rhythm. Although all members of the expedition got the subjective impression that the animals did react to our sounds, this pilot experiment doesn't allow us to claim this with valid scientific evidence. However, this new technological approach aims to enable future designed study to significantly address the animal's curiosity towards human made sounds. Further engineering efforts are needed to improve the echo-suppression and adapt the system for specific experimental needs. Such a system may open the door to new research on interactive playbacks, artificial sonar targets, animal telecommunication, and human–animal communication.

Ethical statement

We received confirmation from the Norwegian Food Safety Authority that under Norwegian and European legislation related to animal research, formal approval and a license are not required (regulation of 18 June 2015 No 761 concerning the use of animals for scientific purposes § 2, f). The experiment was regarded as non-invasive. Nevertheless, we took technical measures to prevent startling the animals by limiting the rate of onset of acoustic emission levels. We limited the emission level to 170 dB re 1µPa rms. We limited our sonic interactions to periods of 15 minutes.

Funding

This work was supported by the Swiss National Science Foundation: Agreement No 31003A_182638, and the NCCR Evolving Language Agreement no. 51NF40_180888.
This article was made Open Access under a CC BY 4.0 license through payment of an APC by or on behalf of the authors.

Acknowledgments

We would like to thank Fiona Wüthrich, Elizabeth Ren, Berenice Fischer and Linus Rüttimann for their help during the expedition.

Author contributions

Conceptualization: JR, RK, RH, JS. Methodology: JR, RK. Investigation: JR, JS, AE. Data curation: JS, JR. Software & Hardware: JR, RH. Visualization: JR. Writing – original draft: JR, RK. Writing – review & editing: JR, RK, RH, AE, JS.

Conflict of interests

The authors declare no conflict of interest

References

Amundin, M., Starkhammar, J., Evander, M., Almqvist, M., Lindström, K., & Persson, H.W. (2008). “An echolocation visualization and interface system for dolphin research,” *J Acoust Soc Am*, vol. *123*, no. 2, pp. 1188–1194.

Antsaklis, P.J. (2021). “Linear Systems: Discrete-Time Impulse Response Descriptions,” in *Encyclopedia of Systems and Control*, Cham: Springer International Publishing, pp. 1144–1147.

Aubauer, R., & Au, W.W.L. (1998). “Phantom echo generation: A new technique for investigating dolphin echolocation,” *J Acoust Soc Am*, vol. *104*, no. 3, pp. 1165–1170.

Branstetter, B.K., et al. (2020). “Spectral cues and temporal integration during cylinder echo discrimination by bottlenose dolphins (Tursiops truncatus),” *J Acoust Soc Am*, vol. *148*, no. 2, pp. 614–626.

Curé, C., et al. (2019). “Evidence for discrimination between feeding sounds of familiar fish and unfamiliar mammal-eating killer whale ecotypes by long-finned pilot whales,” *Anim Cogn*, vol. *22*, no. 5, pp. 863–882.

Deecke, V.B. (2007). “Studying Marine Mammal Cognition in the Wild: A Review of Four Decades of Playback Experiments,” *Aquat Mamm*, vol. *32*, no. 4, pp. 461–482.

Delikaris-Manias, S., McCormack, L., Huhtakallio, I., & Pulkki, V. (2018). “Real-time underwater spatial audio: a feasibility study,” *144th Audio Engineering Society Convention 2018*, no. May.

Filatova, O.A., Fedutin, I.D., Burdin, A.M., & Hoyt, E. (2011). “Responses of Kamchatkan fish-eating killer whales to playbacks of conspecific calls,” *Mar Mamm Sci*, vol. *27*, no. 2, pp. 26–42.

Finneran, J.J., et al. (2020). “Dolphin echo-delay resolution measured with a jittered-echo paradigm,” *J Acoust Soc Am*, vol. *148*, no. 1, pp. 374–388.

Handel, S., Todd, S.K., & Zoidis, A.M. (2012). “Hierarchical and rhythmic organization in the songs of humpback whales (Megaptera novaeangliae),” *Bioacoustics*, vol. *21*, no. 2, pp. 141–156.

Haykin, S., & Chen, Z. (2005). “The cocktail party problem,” *Neural Computation*, vol. *17*, no. 9. MIT Press, pp. 1875–1902, Sep. 01.

Herman, L.M., Richards, D.G., & Wolz, J.P. (1984). “Comprehension of sentences by bottlenosed dolphins,” *Cognition*, vol. *16*, no. 2, pp. 129–219.

Herzing, D. (2016). “Interfaces and Keyboards For Human–Dolphin Communication: What Have We Learned?,” *Anim Behav Cogn*, vol. *3*, no. 4, pp. 243–254.

Herzing, D. L., & Johnson, C. M. (2015). *Dolphin communication and cognition : past, present, and future / edited by Denise L. Herzing and Christine M. Johnson*. The MIT Press.

Holt, M. M., Noren, D. P., & Emmons, C. K. (2011). “Effects of noise levels and call types on the source levels of killer whale calls,” *J Acoust Soc Am*, vol. *130*, no. 5, pp. 3100–3106.

Jourdain, E., et al. (2019). “North Atlantic killer whale Orcinus orca populations: a review of current knowledge and threats to conservation,” *Mamm Rev*, vol. *49*, no. 4, pp. 384–400.

Jourdain, E., et al. (2021). “Natural Entrapments of Killer Whales (Orcinus orca): A Review of Cases and Assessment of Intervention Techniques,” *Frontiers in Conservation Science*, vol. *2*, no. August, pp. 1–13.

King, S. L. (2015). “You talkin’ to me? Interactive playback is a powerful yet underused tool in animal communication research,” *Biol Lett*, vol. *11*, no. 7,

King, S. L., Guarino, E., Donegan, K., McMullen, C., & Jaakkola, K. (2021). “Evidence that bottlenose dolphins can communicate with vocal signals to solve a cooperative task,” *R Soc Open Sci*, vol. *8*, no. 3,

Kohlsdorf, D., Gilliland, S., Presti, P., Starner, T., & Herzing, D. (2013). “An underwater wearable computer for two way human–dolphin communication experimentation,” in *ISWC 2013 – Proceedings of the 2013 ACM International Symposium on Wearable Computers*, pp. 147–148.

Lang, T. G., & Smith, H. A. P. (1965). “Communication between dolphins in separate tanks by way of an electronic acoustic link,” *Science Science*, vol. *150*, no. 3705, pp. 1839–1844.

Lilly, J. C. (1978). “Communication between man and dolphin: The possibilities of talking with other species.,” *Crown Publishers, Inc., New York. 269pp. 1978*. Crown, New York, p. 269.

Ling, F. (2014). “Achievable performance and limiting factors of echo cancellation in wireless communications,” in *2014 Information Theory and Applications Workshop, ITA 2014 – Conference Proceedings*.

Magnúsdóttir, E. E., & Lim, R. (2019). “Subarctic singers: Humpback whale (Megaptera novaeangliae) song structure and progression from an Icelandic feeding ground during winter,” *PLoS One*, vol. *14*, no. 1,

Malinka, C. (2021). “Biosonar of narrow-band high-frequency toothed whales: Sampling a dynamic, multi-target world,” Aarhus University.

Marino, L. (2002). “Convergence of complex cognitive abilities in cetaceans and primates,” in *Brain, Behavior and Evolution*, vol. *59*, no. 1–2, pp. 21–32.

Marulanda, J. Lopez, et al. (Apr. 2021). “Acoustic behaviour of bottlenose dolphins under human care while performing synchronous aerial jumps,” *Behavioural Processes*, vol. *185*, p. 104357,

Muller, M. W., Au, W. W. L., Nachtigall, P. E., Allen, J. S., & Breese, M. (2007). “Phantom echo highlight amplitude and temporal difference resolutions of an echolocating dolphin, Tursiops truncatus,” *J Acoust Soc Am*, vol. *122*, no. 4, pp. 2255–2262.

Narula, G., Herbst, J. A., Rychen, J., & Hahnloser, R. H. R. (2018). “Learning auditory discriminations from observation is efficient but less robust than learning from experience,” *Nat Commun*, vol. *9*, no. 1, p. 3218, Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. 2018. “Taking turns: Bridging the gap between human and animal communication,” *Proceedings of the Royal Society B: Biological Sciences*, vol. *285*, no. 1880,

Richards, D. G., Wolz, J. P., & Herman, L. M. (1984). “Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, Tursiops truncatus.,” *J Comp Psychol*, vol. *98*, no. 1, pp. 10–28.

Ridgway, S., Carder, D., Jeffries, M., & Todd, M. (2012). “Spontaneous human speech mimicry by a cetacean,” *Current Biology*, vol. *22*, no. 20. Elsevier, pp. R860–R861.

Rossa, A., & Kimb, S. J. (Oct. 2021). “Behavioral experiments for gathering labeled animal vocalization data,” in *Proceedings of the 3rd International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots*, p. 35. Accessed: Oct. 21, 2022. [Online]. Available: http://vihar-2021.vihar.org/assets/vihar-2021-proceedings-v3.pdf

Rothenberg, D. (2008). “Whale music: Anatomy of an interspecies duet,” *Leonardo Music Journal*, vol. *18*, pp. 47–53.

Rychen, J., Rodrigues, D. I., Tomka, T., Rüttimann, L., Yamahachi, H., & Hahnloser, R. H. R. (2021). “A system for controlling vocal communication networks,” *Sci Rep*, vol. *11*, no. 1, pp. 1–15.

Shapiro, A. D. (2008). “Orchestration : the movement and vocal behavior of free-ranging Norwegian killer whales (Orcinus orca),” *Massachusetts Institute of Technology*.

Shen, L., Henson, B., Zakharov, Y., & Mitchell, P. (2020). “Digital Self-Interference Cancellation for Full-Duplex Underwater Acoustic Systems,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. *67*, no. 1, pp. 192–196.

Shen, L., Zakharov, Y., Henson, B., Morozs, N., & Mitchell, P. D. (2020). “Adaptive filtering for full-duplex UWA systems with time-varying self-interference channel,” *IEEE Access*, vol. *8*, pp. 187590–187604.

Sondhi, M. M. (1967). “An Adaptive Echo Canceller,” *Bell System Technical Journal*, vol. *46*, no. 3, pp. 497–511.

Address for correspondence

Jörg Rychen
Institute of Neuroinformatics
University of Zurich and ETH Zurich
Winterthurerstrasse 190
CH-8057 Zurich
Switzerland

jrychen@ethz.ch

Biographical notes

Jörg Rychen is employed at the institute of Neuroinformatics at ETH Zurich. Here, in the group of Prof. Hahnloser, he is responsible for the experimental technology for the research on the neural mechanisms of song learning in zebra finches. He is a physicist and got the PhD for experiments on the quantum hall effect in nanostructured semiconductors. After a short postdoc at the Swiss nanoscience center in Basel, he founded a company providing control systems for scanning probe microscopes. He is an expert in data acquisition and control systems with real-time signal processing on FPGAs. He sold the company in 2008 and since then his interest focuses on bioacoustics and on animal communication. Besides the work on songbirds, he is interested in the acoustic communication system of cetaceans. He develops data acquisition systems and conducts field expeditions to Norway to record wild orcas and humpback whales.

https://orcid.org/0000-0002-3302-0193

Julie Semoroz's political posture in itself proposes a misalignment with post-industrial society and the capitalist system. She proposes views and provokes reflection on urban space, ecology, utopia, living together with an anthropological approach mixed with art. She works on the invisible and researches information, non-verbal communication and sound frequencies in collaboration with neuroscientists. She experiments with noise-like and experimental sound forms. In 2020, she received an award from the Swiss National Science Foundation in collaboration with the Swiss Center for Affective Sciences (CISA) at the University of Geneva for her project TWELVE THOUSAND TWENTY and collaborate with NCCR Evolving Language.

info@juliesemoroz.ch

Alexander Eckerle studies biology at the Ludwig-Maximilians-University (LMU) in Munich, Germany. The work in the fields of his interest in marine ecosystems and cetaceans is supported by his former education as mechanical engineer at the Munich University of Applied Sciences (MUAS) and the skill to drive boats. Furthermore he is an experienced scuba and free diver. He works with different cetaceans in various locations in the field.

https://orcid.org/0000-0002-7839-9459

eckerle.alex@googlemail.com

Richard Hahnloser studied Physics and is Professor and Co-director of the Institute of Neuroinformatics, jointly affiliated with the University of Zurich and ETH Zurich. He has more than 20 years of experience in bioacoustics. He studies vocal production and vocal learning in songbirds and he is interested in neuroscience problems that can be characterized by a computational goal encompassing both sensory inputs and motor outputs. He is affiliated with the Neuroscience Center Zurich (ZHZ) and with the Swiss National Centre of Competence in Research (NCCR) Evolving Language.

https://orcid.org/0000-0002-4039-7773

rich@ini.ethz.ch

hrichard@ethz.ch

Rébecca Kleinberger has over twelve years of experience researching and creating new voice technologies and technologically-mediated interspecies interventions. She is an Assistant Professor at Northeastern University and a Research Affiliate at the MIT Media Lab and the McGovern Institute for Brain Science at MIT. As part of her research on voice technology beyond words she has developed and deployed sonic and vocal interventions for animals at the San Diego Zoo, and technological systems for companion animals.

https://orcid.org/0000-0002-5667-5598
rebklein@mit.edu

Publication history

Date received: 29 April 2022
Date accepted: 29 April 2022