Recording vocalizations with Bluetooth technology

**Andrés Gaona-González · Ana María Santillán-Doherty ·
Rita Virginia Arenas-Rosas · Jairo Muñoz-Delgado ·
Miguel Angel Aguillón-Pantaleón · José Domingo Ordoñez-Gómez ·
Alejandra Márquez-Arias**

Abstract We propose a method for capturing vocalizations that is designed to avoid some of the limiting factors found in traditional bioacoustical methods, such as the impossibility of obtaining continuous long-term registers or analyzing amplitude due to the continuous change of distance between the subject and the position of the recording system. Using Bluetooth technology, vocalizations are captured and transmitted wirelessly into a receiving system without affecting the quality of the signal. The recordings of the proposed system were compared to those obtained as a reference, which were based on the coding of the signal with the so-called *pulse-code modulation technique* in WAV audio format without any compressing process. The evaluation showed $p < .05$ for the measured quantitative and qualitative parameters. We also describe how the transmitting system is encapsulated and fixed on the animal and a way to video record a spider monkey's behavior simultaneously with the audio recordings.

Keywords Bluetooth · Captivity · *Ateles geoffroyi* · Vocalizations

A. Gaona-González · A. M. Santillán-Doherty (✉) ·
R. V. Arenas-Rosas · J. Muñoz-Delgado ·
M. A. Aguillón-Pantaleón
Dirección de Neurociencias,
Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz,
Calzada México Xochimilco 101,
Col. San Lorenzo Huipulco 14370 Tlalpan,
México, DF, México
e-mail: santild@imp.edu.mx

J. D. Ordoñez-Gómez · A. Márquez-Arias
Facultad de Ciencias,
Universidad Nacional Autónoma de México,
Mexico City, Mexico

Vocalizations play a significant role in the sociobiology of animal species, such as in mate acquisition (Fischer, Kitchen, Seyfarth, & Cheney, 2004), reproductive synchronization (Geissmann, 2002; Mitani, 1985), group alarm calling (Whitehead, 1987), and the coordination of daily life activities in general (Peng-Fei, Wen, Sheng, & Xue-Long, 2009; Ramos-Fernández, 2005; Sugiura, 2007). The objective of this study was to propose a new method for recording vocalizations based on Bluetooth technology that could provide certain advantages over other configurations that have been used in primatological studies over the years. Table 1 summarizes relevant information from some investigations of primates' vocalizations and the advantages and disadvantages of the various types of equipment that have been used. As can be observed, the majority of the studies listed present one or more of the following characteristics: The objective was to obtain vocalizations of animals in their natural habitat; most of the recordings were done from a distance (average of 10–30 m); there were long periods of discontinuous recording, with a maximum period of continuous sampling of 100 min; and finally, almost all of the studies reported the use of directional microphones.

In order to obtain an appropriate vocalization recording, it is not only necessary to select an adequate high-quality system, but also to apply a good recording technique that results in the reception of a high-fidelity signal. The total continuous recording time and the distance between the recorder and the subject, as well as its change of position during the registering session, are limiting factors that directly affect the potential use in a wider context of the species. For example, in their study of female baboons' vocalizations, Fischer et al. (2004) reported that they were unable to measure the vocalization amplitude in the field due to the constantly changing distance between the subjects and the recording system, in spite of using a high-

Table 1 Condensed data from primate vocalization studies

Reference	Species	Procedure	Equipment
Mendez, Randrianambinina, Rabesandratana, Rasoloharijaona, & Zimmermann, 2008	Lepilemur (21 males)	Focal individual recording, 196 vocalizations processed, natural setting	–Professional Sony WMD6C walkman +Directional Senmheiser ME65/K6
Sabatini & Ruiz, 2008	Leontopithecus rosalia (23 groups)	150 individuals, reproduction of phrases, recording distance: 20, 40, 80, & 120 m	–Professional Sony DAT TCD-D8 +Directional Senmheiser ME66
Sugiura, 2007	Macaca fuscata (23 females)	Focal individual recording (67 h), 280 vocalizations processed, natural setting	–Digital Sony TCD-D8 +Directional Sony ECM-672 & Nondirectional Sony ECM-261
Gros, 2006	Cebus capucinus (24 and 31 individuals)	10 min of focal recording, 2,078 h total. Recording distance: 3–8 m, natural setting	–Professional Sony WMD6C analog walkman +Senmheiser ME80 & ME66, BW: 50 Hz–15 kHz
Ramos-Fernández, 2005	Ateles geoffroyi (17 individuals)	Focal group sampling using a video camera	–Sony Walkman Pro recorder, laptop ES 1688 drive & Wave SE-II program +Directional Senmheiser K3U-ME88
Mendes & Ades, 2004	Brachyteles arachnoides hypoxanthus (43 individuals)	_____	–Marantz PM D 430 Cassette player, RTA/Signal program version 3.1 +Directional Senmheiser ME88
Range & Fischer, 2004	Cercocebus torquatus atys (132 individuals)	15 min record – 60 min rest – 15 min record, for a period of 16 months, recording distance: 5–20 m	–Sony DAT PCM-M1 +Directional Senmheiser ME68
Gil-da-Costa, Palleroni, Hauser, Touchton, & Kelley, 2003	Alouatta palliata palliata (7 individuals)	Focal individual recording, recording distance: 30 m	–Sony DAF TCD-D8 recorder +Sony SRS2001, Cambridge Sound works
Rukstalis, Fite, & French, 2003	Callithrix kahlli (16 individuals)	Recording distance: 1.5 m, 30 min, three times a week	–Marantz PMD201 cassette player, Canary 1.2.4 program +Directional Senmheiser ME80, frequency range: 50–1500 Hz
Fischer & Hammerschmidt, 2002	Macaca sylvanus (92 individuals)	Recording distance: 1–10 m	–Sony WMDC6, Sony DAT TCD-D3 & D100 +Senmheiser ME66
Ghazanfar, Flombaum, Miller, & Hauser, 2001	Saguinus oedipus (11 individuals)	Focal individual using videocamera, vocalization recordings analyzed with Sound Designer program	–Alesis RA-100 amplifier +Directional Senmheiser ME-60
Gouzoules & Gouzoules, 2000	Macaca mulatta, Macaca nigra, Macaca arctoides	Focal individual, microphone inside cage	–Uher 400 monitor, Sony TCM-5000 tape recorder +Directional Senmheiser
Rendall, Seyfarth, Cheney, & Owren, 1999	Papio cynocephalus ursinus	Focal 1 h, recording distance: 5 m	–Pro walkman WM D6C. Sony, type IV cassette +Directional Senmheiser ME80
Teixidor & Byrne, 1999	Ateles geoffroyi (32 individuals)	100 min per individual, recording distance: 10–30 m	–Marantz CP 430 cassette placer +Directional Senmheiser (MKH816T)
Fischer, 1998	Barbary macaques (100 individuals)	_____	–Sony WM DC6 cassette player +Directional Senmheiser ME80

Reference	Species	Recording conditions	Equipment
Hauser, 1998	*Macaca mulatta* (900 individuals)	Recording distance: 1–3 m	−Sony TCD-d7 cassette player +Directional Sennheiser MKH 816
Jorgensen & French, 1998	*Callithrix kuhli* (7 individuals)	Vocalization recordings	−Marantz PMD 201 tape recorder, spectogram analysis using the Canary 1.2 program +Directional Realistic Condenser
Cheney & Seyfarth, 1997	*Papio cynocephalus ursinus* (17 individuals)	Focal individual recording	−Nagra DH placer +Not specified
Brown, Gomez, & Waser, 1995	*Cercopithecus mitis, Cercocebus albigena, Cercopithecus aethiops, Papio cynocephalus*	Recording distance: 12.5 to 100 m	−Nagra IV-L tape recorder, 85-W Sony XM 700[a] amplifier, Nagra IV-S tape recorder +Directional AKG
Aich, Moos-Heilen, & Zimmermann, 1990	*Theropithecus gelada* (29 individuals)	Recording distance: 1–20 m	−UHER 53 Report stereo +Directional Sennheiser MKH 816, UHER M646 & Sennheiser K3N/ME88
Whitehead, 1987	*Alouatta palliata palliata*	Focal group, recording distance: 55 m, vocalizations analyzed with Angioscan audiospectrum program	−Nagra IVD tape recorder +Directional Sennheiser MKH 815
Baldwin & Baldwin, 1976	*Alouatta palliata* (151 individuals)	Focal group vocalization recording, vocalizations analyzed with the 6061B Sona-Graph program	−Uher 4000 tape recorder +Nondirectional Uher M514

fidelity system (DAT Sony Walkman TCD-100 and a directional Sennheiser ME66 microphone). On the other hand, the possibility of continuous recording enhances not only the possibility of analyzing characteristics of the animal's vocalizations, but also the intervals between single vocalizations.

An ideal recorder is one that registers sounds without alterations by matching at least their dynamic and frequency range and by preserving all of their features. Traditional analog tape recorders, as described in Table 1, are not ideal, nor are compact cassette or reel-to-reel recorders. They degrade the signals they record by adding hiss, distortion, frequency response alterations, speed variations, print-through effects, and drop-outs. In contrast, digital recorders are not affected by the same problems. Within the dynamic range and the frequency limits (due to the number of bits and sampling frequency), they record and reproduce signals with great accuracy, low noise, flat frequency response, and no speed variations. Moreover, the inherent noise inserted due to the process of quantization during the analog/digital conversion can only be minimized by applying a higher resolution in the digital code corresponding to 16 bits or more, depending on the particular situation.

In spite of the arguments above, and although some contemporary primate studies do use digital systems, others have kept using analog systems, which means that the recording procedure has to be limited by the type of equipment used, even though it is not the optimal procedure. Arguments that have justified the use of this type of equipment are the necessity to record high-fidelity bioacoustic signals and the importance of recording animals in their natural habitat. Nevertheless, the sampling conditions shown in the referred studies imply inconvenient factors that are not always avoidable. For example, the size and the weight of the equipment, the energy consumption, the vulnerability of the system to the environmental conditions (e.g., heat, humidity), and also the equipment's deterioration, which gets progressively worse.

The method proposed herein is based on Bluetooth technology, with which vocalizations are captured and transmitted wirelessly to a receiving system without affecting the quality of the recorded bioacoustic signal.

Method

In order to develop an adequate system, the bandwidth (BW) and the general registering conditions of the vocalizations had to be taken into consideration:

Recording conditions

The need to obtain continuous long-term and high-fidelity vocal recordings on free-moving animals reduces the number

of recording systems that can be used. In this study, the recording system had to meet the following requirements:

Regarding the species The spider monkey (*Ateles geoffroyi*) is medium-sized and slender. This precludes the possibility of affixing equipment of even medium dimensions and weight on its body. The complete equipment had to be as small as possible, so that the animal was comfortable and got used to it soon. Also, it had to be out of its reach.

Regarding the audio quality The selected system had to register and record the vocalizations in high fidelity, so it was important that the complete bandwidth of the vocalizations be recorded, because integration of the high-frequency harmonics participates in the definition of a vocalization's timbre, and might make the difference in certain circumstances in relation to the behavior of the focal animal. Two captive male adult spider monkeys had their

vocalizations recorded using a portable high-fidelity system that included an omnidirectional Sony microphone (Model ECM-ZS90) and a Sony Hi-MiniDisc (HMD; Model MZ-NH1) for recording the audio signal in the linear pulse-code modulation (PCM) mode, covered the audio bandwidth from 20 Hz to 20 kHz, sampling at 44.1 kHz with 16 bits per sample and per channel and without any compression process. In order to analyze the audio frequency response, a PCI Sound Blaster Live 24-bit audio card was used, for a sampling rate of 44.1 kHz and 16 bits per sample. Sound Forge version 8.0 was used for signal processing and frequency response analysis, as shown in Fig. 1.

These preliminary tests indicated that the harmonics of the species' vocalizations can reach over 19 kHz, which meant that it was essential to ensure that the recording system covered the total audio bandwidth (20 Hz–20 kHz).

Regarding the environment The animals live in an outdoor enclosure (National Institute of Psychiatry Ramon de la
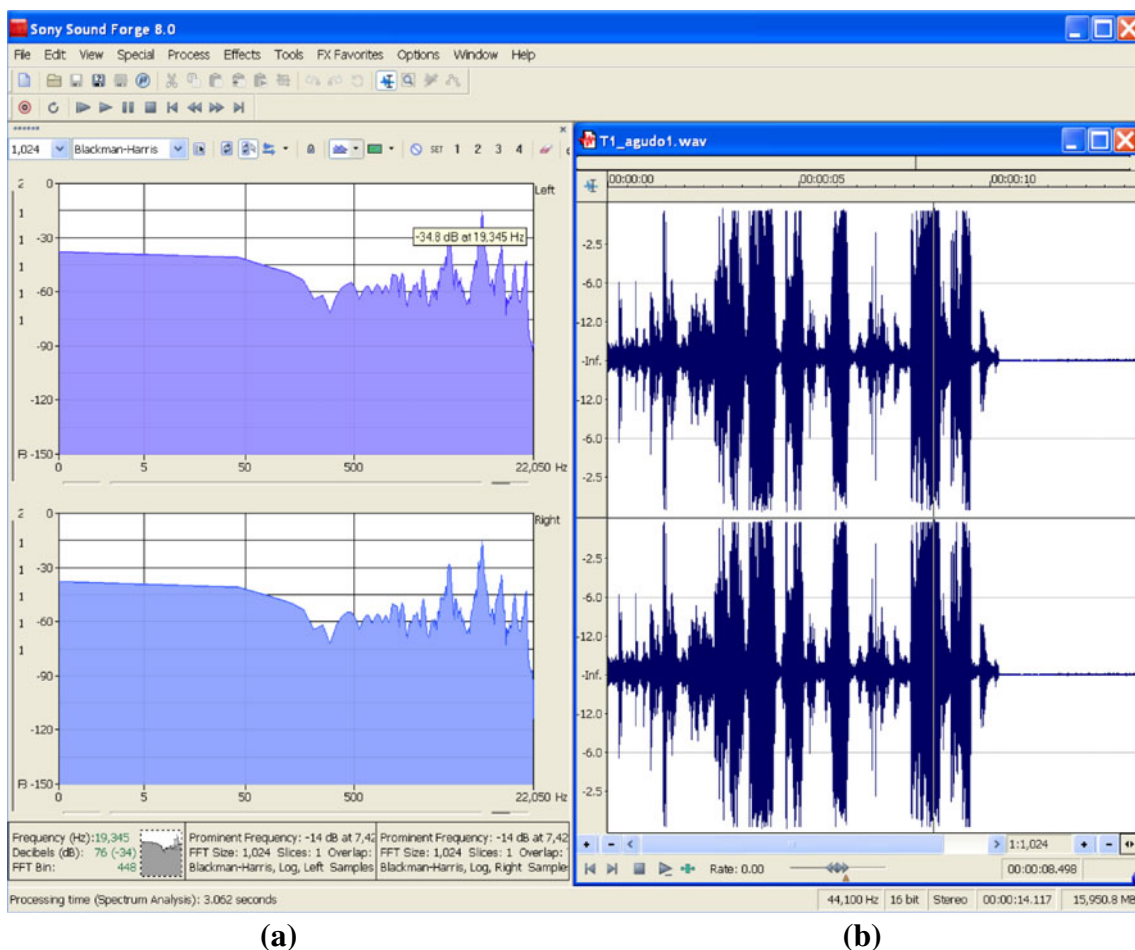

(a)   **(b)**

Fig. 1 Short-distance vocalizations from one awake spider monkey on a 14.1-s time segment: (**a**) Frequency response approximately 3 s around the cursor position (white line in panel b). A fast Fourier transform algorithm (FFT) was used. An evident harmonic of 19.345 kHz can be seen. (**b**) Time domain view

Fuente Muñiz, Mexico City). They could move within a space 7 m in height and with a radius of action of approximately 3 m, with a wire fence roof, so that they were living in a natural day/night cycle (exposed to light, rain, and natural temperatures). It was very important that the system be protected primarily from the humidity. Also, because the animal's free movement conditions required the use of a wireless system, the maximum distance between the animal and the receiving system in the cage had to be considered (in this case, the distance was of approximately 7 m).

Regarding the energy consumption of the system The more time the power supply is working, the longer are the periods of continuous recording. In this system, we proposed that the animal carry the equipment for the audio registering with it, improving the value of the study, since not only could the characteristics of the animal's vocalizations be analyzed, but also the intervals of single vocalizations, as well as the intervals of sessions of vocalization bouts. Therefore, we included a rechargeable energy source, which provided a maximum quantity of energy with minimum consumption and physical weight.

Selection of the main components of the system for sensing and recording

Microphone Unidirectional microphones, like those listed in Table 1, have to be constantly adjusted by the researcher in order to capture the sounds of interest in a moving target. Therefore, we proposed the use of a light miniature microphone, which required minimum conditions of polarization, so that the animal could carry it through the recording session (fixed on a collar 5–10 cm from the mouth), which facilitated a continuous (e.g., a complete day/night cycle) and clear recording, without worrying about the background noise. An Electret unidirectional condenser microphone (Model UCM 97100) with the following specifications was chosen. Dimensions: 0.97 cm × 1.0 cm; sensitivity: –47 dB, 0 dB = 1 V/Pa, 1 kHz; tolerance: ±4 dB; bandwidth: 20 Hz–20 kHz, with a flat response in the complete range, except around 16 kHz; maximum operation voltage: 10 V; standard operation voltage: 1.5 V; resistance: 680 V; maximum power consumption: 0.5 mA; and noise/signal relation (NSR): 58 dB.

Voice signal coding and transmission technology Nowadays, there are some definitions and specifications for wireless technology that may have better performance than the Bluetooth technology—for example, ultra wideband (UWB), or worldwide interoperability for microwave access and IEEE 802.16 (Wi-MAX) (www.bluetooth.com/bluetooth/). However, these are not available yet, whereas

Bluetooth is already available with devices and specific applications, including high-quality audio transmission.

The Bluetooth technology was developed during the last 10 years and was conceived to be an open system, with an architecture similar to present computer networks, so that new specific applications that guarantee the communication between devices could be developed, such as high-fidelity voice and music transmission. Some of its main advantages, in reference to this project, were that it can penetrate solid objects, and that the Class 2 type operates omnidirectionally up to 10 m and at up to 1 Mbps using minimal transmission power (2.5 mW) with low-cost transceivers (transmitter/receiver). Also, Bluetooth technology adds frequency switching in order to avoid interference (FHSS: frequency hopping spread spectrum), uses three security modes, and can get point–multipoint links. Bluetooth is based on the open system interconnect (OSI) model of the International Standard Organization (ISO) for the interconnection of open systems. The Bluetooth specification uses an architecture of protocols that divides the different network functions in a system of five levels or layers (deBont, Groenewegen, & Oomen, 1995; see www.bluetooth.com/bluetooth/). Bluetooth has defined two profile generations for communication applications, one of which consists of a specific profile A2DP (Advanced Audio Distribution Profile) that was designed for high-quality audio transmission. This set of essential characteristics fulfills the requirements for the wireless collection of spider monkey vocalizations based on the A2DP profile of the Bluetooth technology.

Audio coding and decoding in the A2DP profile The audio signal coding and decoding is a very important aspect with regard to the possibility of adapting the technique and for a reliable analysis of the vocalizations, since any kind of modification could cause alterations or limitations that might make the technique useless for research purposes. In order to transmit the audio signal and conserve the high-fidelity characteristics, Bluetooth applies its own version of the codification technique called *subband coding* (SBC), in which an algorithm of low complexity is defined using four or eight subbands, as well as an algorithm of localization of adaptive bit and block-adaptive PCM quantizers. This version is based on Bluetooth specification 1.2–2007 (official document ISO/IEC 11172-3_1993: "Information technology—Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s"; see deBont et al., 1995). The SBC technique has been known for rejecting masked frequencies, although the system tried to restore them during the decoding process. Of course it is not beneficial if a signal is manipulated and altered by eliminating a part of it and/or introducing other components to it as part of the registering or recording

process. Nevertheless, the following aspects have to be considered:

The Bluetooth specification in the A2DP profile guarantees a high-quality signal in the complete audio bandwidth, using high sampling rates in the analog-to-digital (A/D) conversion for the coding (44.1 kHz and 48 kHz) in the SBC and high band rates on the coded digital signal (512 kbps) in the receiver or sink devices.

Also, the A2DP profile is designed for music transmission, which implies the transmission of an audio signal rich in frequency distribution due to the sum of frequency responses of the musical instruments (including voice). This means that, in order to implement the SBC of the audio signal, it should be able to maximize the content of all of these frequencies, especially when considering the masked frequencies, since their elimination optimizes the space in the bandwidth in order to transmit online to a possible receptor with the rate limitations that the transmission channel offers. Nevertheless, the official document that defines the A2DP profile does not specify a numeric index of the possible loss or alteration of information due to its own definition of the SBC coding of the audio signal in case of music transmission, and even less so in the case of voice transmission.

The generic audio coding method is based on human audio perception (psycho–acoustic model), so that masked frequencies can be eliminated without being perceived, although in the case of music this does affect the characteristics of the coded signal. In the case of voice coding, however, the phenomenon of masked frequencies tends toward zero due to the focal recording. Although there are very elaborate vocalizations in the distribution of the bandwidth in the most complex case, according to the bandwidth analysis of the preliminary registers that we have carried out, we expected that there would be a fundamental frequency near 4 kHz and four or five harmonics maximum (for high-pitched vocalizations), which do not belong to the group of masked frequencies, since the rest of the auditory information consists of the background sound. Even though a vocalization could be quite elaborate in its frequency distribution, it is important to keep in mind that it is only one voice—that is, an audio signal with fewer spectral frequency components than would be expected for a high-fidelity music transmission. On the other hand, taking into account that the preliminary analysis of the bandwidth of *A. geoffroyi*'s vocalizations showed that those with the highest fundamental frequencies were of 4 kHz and had four or five harmonics, we can be quite sure that these sounds will be well enclosed within the 20 kHz considered as the frequency analysis limit in the present work, and that they will not form part of the frequencies that can be masked. Therefore, it is not possible to adopt or reject the application of this technology without

having compared the results to those of the reference recordings, obtained by HMD and a high-fidelity microphone, after being transmitted through the available Bluetooth devices and complementary circuits proposed. This analysis can be carried out by using sonograms of signal segments of these vocalizations from both of these sources and the corresponding comparative description.

In order to validate the herein-proposed Bluetooth technology method, a comparative analysis of the spectrograms presented in the Results section was performed for the signal obtained through the HMD Sony equipment (Mod. MZ-NH1) and the high-fidelity Sony microphone (Mod. 96 ECM-ZS90), as well as for the signal obtained through the method herein proposed, based in the Bluetooth technology.

Description of the Bluetooth system A set of Philips devices called a stereo Bluetooth headset receiver and transmitter (Model SHB2000/77) with the following important specifications was acquired: Bluetooth audio devices support via A2DP; operation frequency band 2.4–2.48 GHz, which fulfills the operation norm of Bluetooth V1.2 Class 2; stereo audio in high quality; low energy consumption in standby mode (0.6 mA); battery changing by USB standard (5 V at 200 mA); weight 26.3 g (receiver) and 13 g (transmitter); continuous use up to 10 h, or 250 h in standby mode; and linking time <4 s.

Modifications made to the equipment

Preamplification circuit for the Electret microphone In order to accomplish the objective, it was necessary to substitute for the Bluetooth transmitter's connection to a music source an Electret microphone (UCM 97100) and its own preamplification circuit with the required impedance coupling. The design of this stage consisted of the three elements: microphone, preamplifier, and outgoing amplifier.

The microphone's output signal was too weak as an amplifier input, so the amplitude had to be increased by a preamplifier in order to be able to control it—that is, the preamplifier was used to provide the level of amplitude of the audio signal required by the Bluetooth transmitter. An outgoing amplifier was connected at the end of the circuit. A National Semiconductor LM386N-1 integrated circuit was chosen, as the number of external components that could be added to the integrated circuit was reduced, which was important in order to optimize the need for electric power for its operation. It was also important because fewer components allowed the circuit to be compact, optimizing weight and size. The constructed circuit guaranteed an amplification range (G) between 20 and 200, with low distortion and a wide range in the voltage of polarization,

so that it could be powered by the same voltage required by the Bluetooth transmitter (3.7 unipolar volts).

Because the circuit consumes a total of 5 mA, a rechargeable Lithium Sony battery (Model LIS2106) with 1 A-h was selected as the energy source, sufficient for powering the circuit for 200 continuous hours (8.3 days).

The evaluation of this circuit together with the microphone was carried out by using a Hewlett Packard function generator (Model 3310B) in order to generate an audio frequency sine wave. A Creative subwoofer (Model T300), with low-frequency response from 40 Hz on, and a high-fidelity BM Boschmann tweeter (Model MM-8NEO) with a response range between 2.5 and 29 kHz were used as devices for the audio generation. The generated sound, captured by the microphone and preamplified by the present circuit, was linked directly to the Bluetooth transmitter, and the receiver was connected to an analog audio input to the computer where the bandwidth was monitored, carrying out the same procedure as in the evaluation of the Bluetooth system described below. In this stage, a band response between 40 Hz and 21.6 kHz was identified.

Audio transmission and reception system (Bluetooth) A first evaluation was carried out using a music player audio source. Concerning this element, the objective was to determine the energy consumption and the maximum transmission distance. The following practical results were obtained:

a) Current consumption: 25 mA
b) Total time of continuous usage: 6.5 h. For our purposes, it was necessary to substitute for the battery that comes with the system another battery, labeled as GCR18650D of Li-ion, of 2,200 mA-h, that was also rechargeable and provides up to 88 h of continuous use (3.6 days).
c) Maximum range in open spaces (line of sight): 20 m.

In a second evaluation of the bandwidth response, an audio generator (NHC tone generator version 2.01, trial mode) was used in a computer to generate a sine wave at the required frequency (one single tone at a time). This signal was sent to the audio output (Sound Blaster PCI adapter, Model Live 24 bits) where the Bluetooth transmitter was connected. On the other hand, the Bluetooth receiver was connected to the analog audio input of another sound card (same model) in another computer, in which the Soundforge program version 9 monitored this audio signal in real time, visualizing the frequency response that provided the fast Fourier transform function and tested the signal transmission in the complete audio bandwidth. The frequency range that could be transmitted went from 20 Hz up to 21.247 kHz, including the bandwidth registered previously in the monkey's vocal emissions.

Fixing the equipment to the animal's body The Bluetooth transmitter circuit was put inside its case, followed by the preamplification and conditioning circuits, which were placed into the empty space of the original battery. This circuit gives the signal the proper level for the next stage, and eliminates any noise or the direct-current components that the signal could have acquired. It was also important at this stage to provide the signal with the proper impedance coupling in order for it to maintain the same features as in previous stages in order to deliver them in the same fashion to the receiver. The battery that powered the two circuits was joined with adhesive and placed outside the case. This complete unit—that is, the conditioning circuit, transmission device, and batteries—was covered by a waterproof polyethylene material and put inside a plastic box in order to cushion any kind of impact. The box was fastened by a harness that was placed between the neck and the armpits of the spider monkey. It included a clip and material that facilitated adequate adjustment to the animal's body and was placed on the upper part of its back. For the design of the harness, both the anatomy of the species and its habits of arboreal locomotion (brachiation) were taken into consideration. The microphone was placed on the neck-collar attached to the harness, near the animal's mouth but outside its field of vision. From there, the microphone's cable led to the conditioning circuit, hidden with adhesive underneath the harness shoulder strap, and into the plastic box that carried the circuit and the transmitter. Before the equipment was placed on the monkey's body, the animal received general anesthesia with 0.01 mg/kg of Zoletil 50, which is enough to allow the manipulation of the animal safely. The equipment remained switched on so that it could transmit the vocalizations continuously 24 h a day during the next 3.6 days from the moment of placement. The system could be carried by the animal for about 18 days if the Bluetooth transceiver was switched to standby mode. The total weight of the harness was 94 g, while the weight of the transmitting system including batteries was 92.1 g, which gives a total weight of 186.1 g. In order to have the best possible line of vision, the Bluetooth receiver was placed on the inner side-wall of the monkey's cage at a medium height, covered by a dark acrylic material so it would avoid attracting the monkey's attention. From there, the analog audio signal was led through a 3-m cable connected to the working computer inside the laboratory. Figure 2 shows a schematic summary of the equipment and procedure.

This research was done in compliance with the Mexican Official Norm NOM-062-ZOO-1999: Technical Guidelines for the Production and Use of Laboratory Animals. Our study was also submitted to and approved by the Ethical and Research Committees (Project No. 3330-B) of the Instituto Nacional de Psiquiatría "Ramón de la Fuente Muñiz."
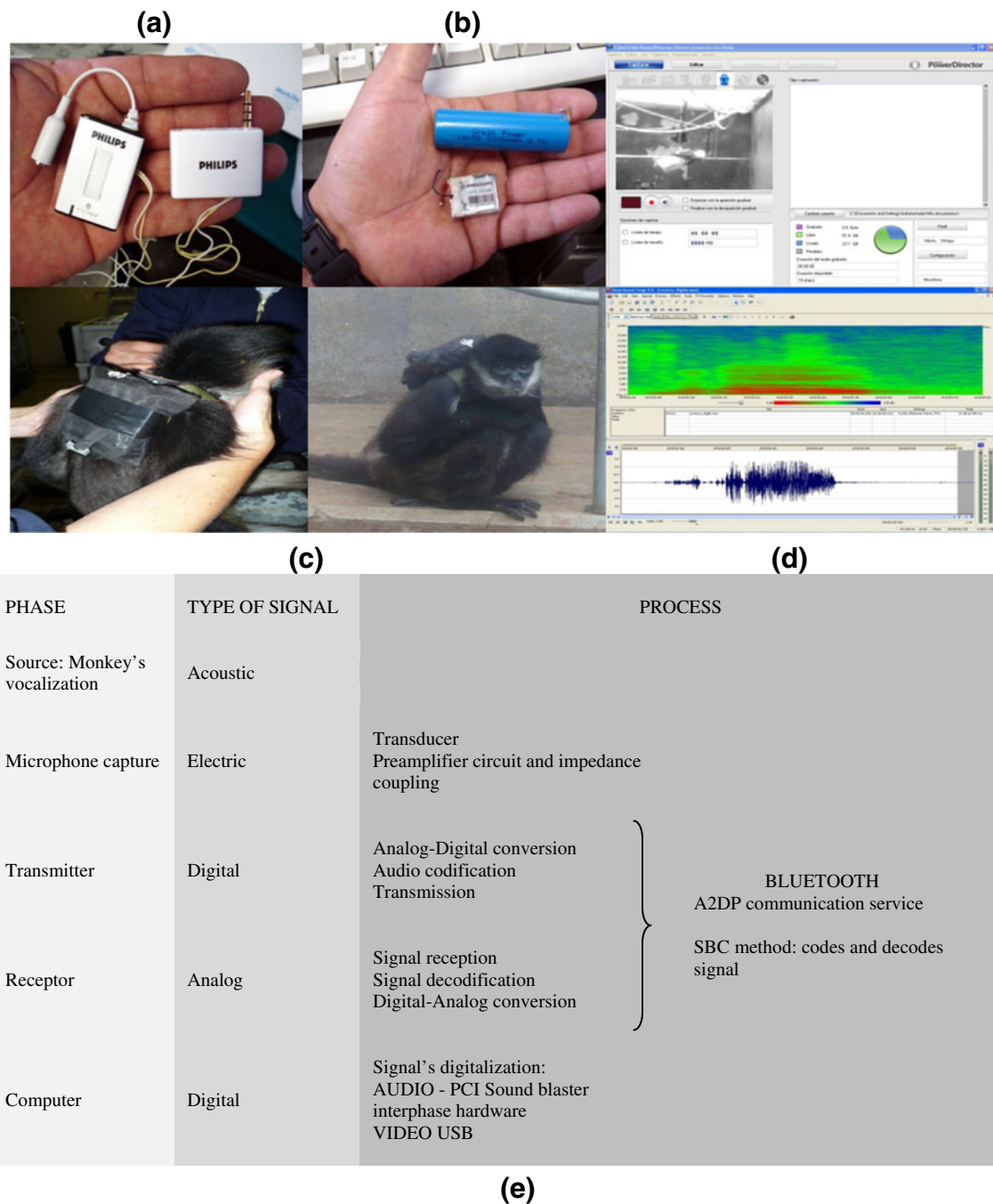
Fig. 2 **a** Bluetooth device: Transmitter–receptor. **b** Battery modifications. **c** Monkey carrying the equipment. **d** View of audio and video signals in computer

Contextual analysis of the vocalizations Although adequate bioacoustic high-fidelity signals are useful for the classification and description of the vocal repertoire of a species, disciplines such as ethology and systematics will highly profit by including the recording of behavior, which allows for the contextualization of a vocal emission, so that researchers may then understand its function. In order to achieve a contextual analysis of the vocalizations, both recordings (vocal and behavioral) were synchronized

through the Power Director ver. 7 (of Cyberlink) program, and with the aid of video codecs (Megacodecs ver. 1.0), the high fidelity required for the audio and the low resolution required for the video could be configured separately, generating small files that could be handled by audio analysis programs. Both signals went into the computer through USB gates. After the video signal was acquired using a conventional video camera with night shot (Sony Handycam), the analog output entered, through a Kworld

digitalizing card connected to the USB gate, into the host computer. On the other hand, the analog audio channel that came from the receptor equipment was connected to the input of an external digitalizing module that guaranteed coverage of the range of audio frequencies up to 20 kHz, and the output of this module entered the host computer through another USB port, version 2.0.

Sonogram analysis using the A2DP Bluetooth profile as transmission media Since there were no indications that the information was altered by the use of these devices for voice coding, it was necessary to evaluate this technology by a quantitative and qualitative comparative analysis of the time frequency behavior. We compared the sonograms of one vocalization obtained using a high-fidelity recording technique—that is, without any information compression processing—as a reference, and those obtained for the same vocalization, transmitted by the A2DP Bluetooth profile version 1.2.

For this purpose, in a typical descriptive report, the quantitative variables consisted of the fundamental frequency in Hertz (Hz) and the harmonics, the amplitude of each in decibels (dB), and the duration in seconds (s). For the qualitative characteristics, all of the frequency changes during a vocal emission, including vibratos and guttural sounds, were taken into account.

Results

Figure 3a and b represent the spectrogram of the same vocalization registered with the HMD Model MZ-NH1 and an omnidirectional Sony microphone Model EM-ZS90 (Fig. 3a) and the selected Bluetooth system (Fig. 3b). According to the results obtained from the assessment done using both methods, we can see that a favorable balance was achieved and that the method herein proposed is reliable and can be used to increase knowledge of the vocal repertoire of captive spider monkeys. Visual inspection indicates similar sonogram patterns, so, for the objective of obtaining the vocalization repertoire, quantitative and qualitative parameters can be obtained from either of the two graphics. Both signals were recorded in WAV format. The sonogram pattern was characterized by high bandwidth and lack of frequency distribution.

Table 2 illustrates an exercise that measures the parameters in the time and frequency domains of the spider monkey vocalizations. Based on the calculation of the
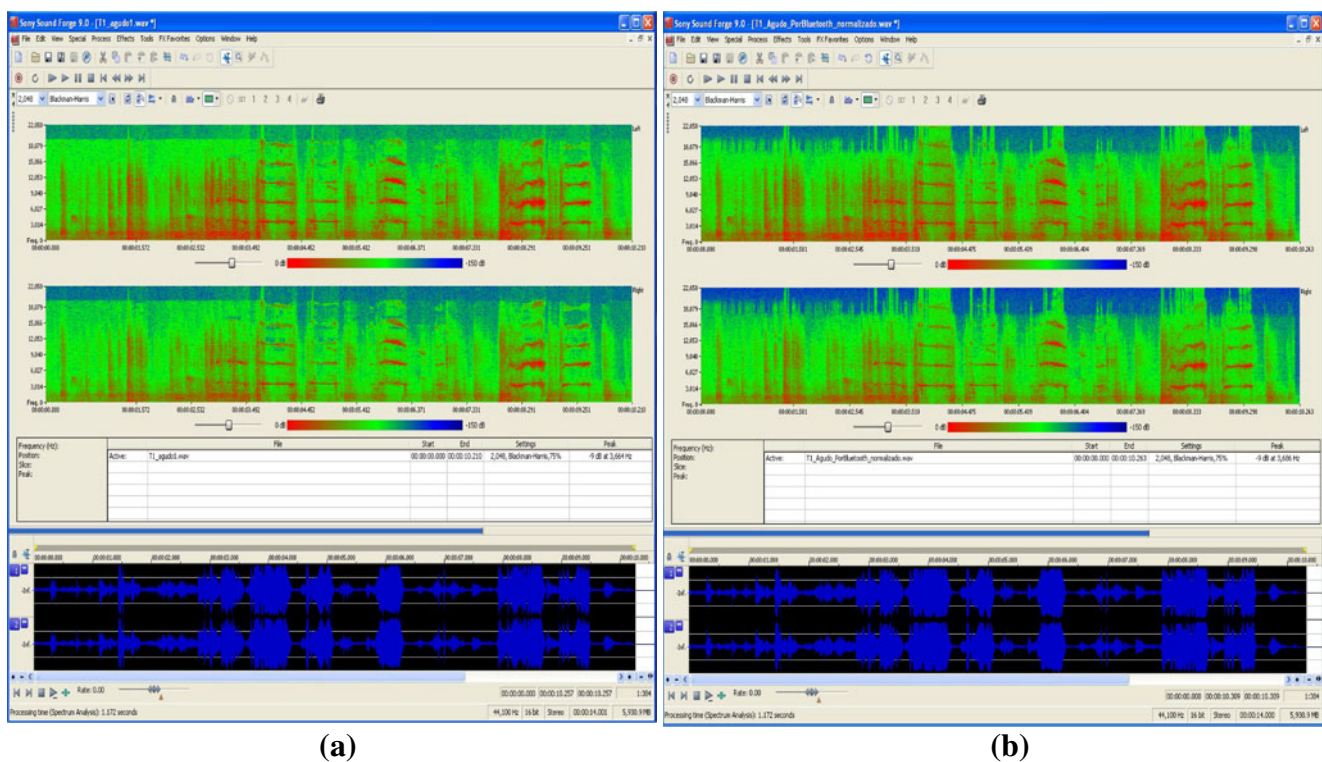

(a)   (b)

Fig. 3 Time and frequency comparative analysis using a short-distance vocalization from a captive spider monkey (reference uncompressed PCM coding vs. the proposed SBC on A2DP ver. 1.2 coding). Sonograms from the same audio time signal segment were obtained for **a** a Sony HMD recorder and ECM-ZS90 omnidirectional microphone (reference technique) and **b** the Philips SHB2000/77 Bluetooth equipment and auxiliary electronics we designed (proposed technique)

Table 2 Comparative values for quantitative and qualitative parameters measured from the graphical results obtained for Fig. 2

$F(0)$ [Hz] max	$F(0)$ [Hz] average	$F(0)$ [Hz] min	Harmonic 1 [Hz]	Harmonic 2 [Hz]	Harmonic 3 [Hz]	Harmonic 4 [Hz]	Duration [s]	Melodic features
PCM Coding With the HMD								
4040	3793	3546	7941	11873	15727	————————	0,059	
4163	3816	3469	7863	11795	15727	————————	0,327	
4086	3739	3392	7555	11102	14725	18580	0,624	
4086	3777,5	3469	7555	11256	15034	18657	0,475	
4009	3739	3469	7324	11025	14571	18195	0,054	
4317	3931,5	3546	7286	11795	15419	19197	0,463	3
5242	4933,5	4625	9714	14802	————————	————————	0,122	1
3931	3738,5	3546	7329	10687	14494	18272	0,147	2
4086	3469	2852	7632	11256	14956	18580	0,634	3
4052	3515,5	2979	7389	10846	14660	18474	0,475	3
SBC Coding With the A2DP Bluetooth Profile								
3978	3685	3392	7941	11873	15805	————————	0,054	
4240	4008,5	3777	7863	11718	15573	————————	0,329	
4086	3777,5	3469	7478	11102	14725	18479	0,62	
4086	3816	3546	7555	11256	14956	18743	0,475	
4009	3739	3469	7324	10947	14571	18118	0,057	
4240	3970	3700	7709	11718	15573	19082	0,46	3
5165	4972,5	4780	9714	14829	————————	————————	0,122	1
3931	3661,5	3392	7247	10793	14494	18167	0,149	2
4009	3430,5	2852	7555	11256	15111	18654	0,641	3
4114	3570,5	3027	7298	10947	14578	18367	0,473	3

2a: Parameters obtained for the reference coding technique (uncompressed PCM coding). 2b: Parameters obtained from the Bluetooth transceiver and auxiliary incorporated designs (SBC codec on A2DP profile). 1, Melodic features are present in the vocalization; 2, Melodic characteristics are more evident in harmonic components; 3, *Vibratos* are present in harmonics with an oscillation of 100 Hz.

sonograms that correspond to Fig. 3, the parameters were measured directly on the basis of the graphic with the previous adjustment of the adequate scale for a precise measurement in each particular case. The amplitude measurements were omitted, because both signals presented different gain factors. There was a coincidence in the qualitative characteristics observed in both cases. Concordance coefficients for all quantitative parameters had $p < .05$. For qualitative characteristics, we obtained the same warnings for *vibratos* and other melodic characteristics at the fundamental and harmonic components in both cases.

As can be observed in Fig. 3, the pattern that described the time frequency behavior of a voice sequence (sonogram) was very similar in both cases, using Fig. 3a as a reference, so that the following points could be highlighted:

- There was no alteration either in the number of harmonics or in their high- or low-frequency range. This characteristic proved to be highly relevant with respect to the possibility that the SBC codification could alter or eliminate some of these components.
- There were no frequency displacements or shifts.

- There were neither phantom frequencies, nor overlapping of frequencies, nor any kind of noise generated by the subsampling.
- Reliable results could be obtained in the measurements of the quantitative parameters of the vocalizations (fundamental frequencies, harmonic frequencies, amplitudes, and duration).
- The qualitative parameters, such as *vibratos* or any change of frequency or guttural sound, could also be measured on the basis of the representation in Fig. 3b.
- The amplitude analysis was not described because there were different gain factors between the Bluetooth system and the reference equipment (HMD with the omnidirectional microphone).

Behavior of the bioacoustic signals with the possible presence of masked frequencies

In order to evaluate the conditions in which the minimum probability of masked frequencies would occur, the same analysis carried out in Fig. 3 was carried out using a long-

distance vocalization of a spider monkey in its natural habitat (Mennill, 2005). The audio consists of three "whinny" vocalizations (two in the foreground and one in the background) mixed with a bird whistle. This recording is a good example for masked frequencies in accordance with the design philosophy of the generic technique of SBC audio codification on which the A2DP Bluetooth profile is based, as the sonograms reveal a rich content in frequencies for the "whinny" vocalization in which the whistle bird appears intermixed in the same frequency and in neighboring frequencies. Figure 4 shows the comparative graphic of this signal segment using both codification techniques (PCM and SBC in Bluetooth), where the singing of the bird can be observed especially during the first "whinny." These results not only offered more certainty that the proposed system for the audio register of the spider monkey's calls is reliable with respect to the required audio quality, but also reduced the expectation that the audio might alter or eliminate important information for its application in bioacoustic studies of the spider monkey. As in our evaluation of Fig. 3, the amplitudes were not considered due to unequal amplitude gains. Nevertheless, it could be noted again that both patterns were identical, so that a reliable measurement of quantitative and qualitative parameters was expected. It is also important to observe that the frequency distribution is different from the one in Fig. 3.

Discussion

A must in vocal behavior research is updating a species' bioacoustic database by taking the opportunities offered by improvements in recording techniques and by new technologies that present great advantages and versatility, both in their application and in the recording quality.

In the digital world, the final word has not yet been said about the quality of the signal that a coding algorithm can offer. Since the techniques of audio coding consist of algorithms that are programmed in software, the evolution of their optimization is accelerated, with the objective of conserving the high fidelity of signals at a higher response speed, especially when the techniques do not require the elimination of the content of the signal, due to the efficiency of their coding–decoding technique and/or to the improvement of the transmission speed. A good example is a coding algorithm (Massey & Wray, 2007) in recent development (enhanced apt-X), which handles four subbands with a fixed compression factor of 4:1 and a delay or synchronization time of only 3 ms. The evaluation of the
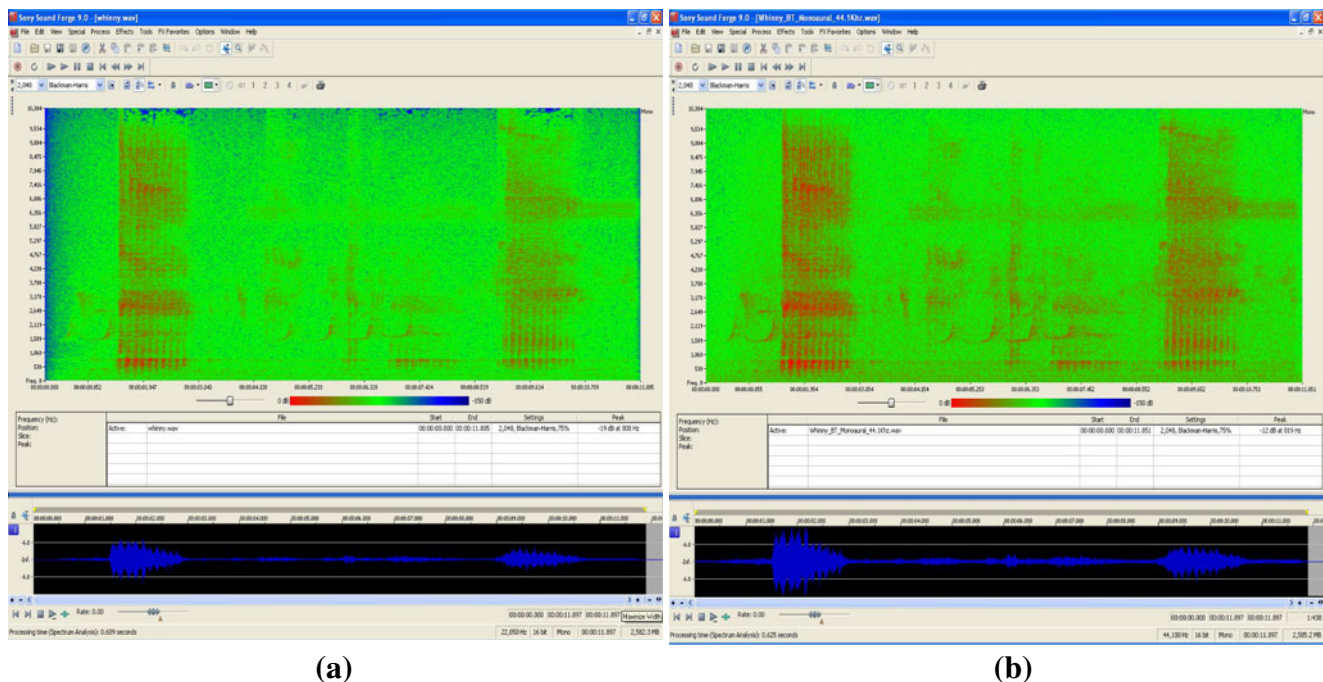

(a)   **(b)**

Fig. 4 Same comparative procedure as in Fig. 3, but now for a long-distance vocalization recording from a spider monkey in its natural habitat. Recording includes a whinny vocalization mixed with a bird whistle. The latter sound could be masked by the whinny sound. This constitutes a didactic example to check whether the generic SBC codec could eliminate or modify the masked audio signal, as indicated by its psycho–acoustic model design philosophy. Visual inspection indicates high coincidence of the patterns. The bird whistle pattern visually remains in the output of the Bluetooth transceiver. The sonogram is characterized by a high-frequency distribution in a low-bandwidth behavior. Differences in color saturation indicate the unequal amplitude gains for each case

frequency response for the algorithm herein presented indicated a pattern that was very similar to that of the PCM technique used as reference in this study across almost the whole of the frequency range. Although the response to high frequencies is not as good as expected yet, this technique can be considered as a good candidate for being adopted in the immediate future, as an optional codec within the A2DP profile of the Bluetooth technology, for example. Enhanced apt-X also allows the coding and transmission of auxiliary information with an additional signal on the same channel, without any loss of quality in the audio signal that is being coded.

In this work, a favorable balance was achieved according to the results obtained in the assessments of the proposed method and equipment, thus showing it to be a reliable option in the study of the vocal repertoire of the spider monkey. We selected and adapted a device that guarantees good quality of audio capturing and transmission, low weight, small size, and minimum power consumption. The evaluation of the components, both individually and as a complete system, showed an adequate range of functionality for the bandwidth of the spider monkey's vocalizations. However, besides discussing the technological aspects of this new device, we consider it important to evaluate its applicability as well.

The main disadvantage of the Class 2 system we used is the short range it covers (from 10 to 20 m). However, the massive availability of Class 1 Bluetooth version 2.0 equipment is expected soon. This equipment will be able to transmit a signal at 3 Mbps that could be used in semicaptive environments with a 100-m radius. Recently, the company AIRcable (www.noticias3d.com/noticia.asp?idnoticia=19570) developed a Bluetooth adapter with an antenna that facilitates communication in a radius of up to 30 km (Host XR). Such a device would be of a great value in bioacoustic studies of animals in their natural habitat, although some considerations must be taken into account.

Two of the main problems that have to be faced in primate vocalization studies, particularly when recording in the field, are the impossibility of obtaining long-term continuous recordings and of measuring the amplitude of vocalizations due to variability in the distance between the subject of study and the recording equipment.

The fact that the monkey carries with it the element that captures, amplifies, and transmits the audio, provided with a high-power battery that supports the low consumption of the complete circuit (30 mA), opens up the possibility of resolving these two problems. When the animal carries the device, not only is the distance between the mouth and the microphone very short (5–10 cm), which means there is no loss in energy of the acoustic signal due to the distance, allowing for a more precise assessment of acoustic parameters, but also the amplitude (intensity) of the signal

can be considered, since the distance between mouth and microphone remains constant. This is important, as it has been reported that this parameter is associated with the individual's level of motivational or emotional arousal (Scherer, 1989). The work done with directional microphones has been able to sample loud vocalizations such as aggressive or long-distance calls, but the short distance between the equipment and the sound source also enables the researcher to study new, very-low-intensity vocalizations. We are currently analyzing the affiliation and play vocalizations of spider monkeys, some of which cannot be heard (or recorded) beyond 2 m. In the same way, this system provides a recording of the complete frequency range of a vocalization, which sometimes is lost when the microphone is placed at a longer distance.

Through our study, we were also able to identify several disadvantages of the proposed system, in particular when referring to its application to free-ranging populations. In the first place, it is an intrusive method, since it requires capturing, anesthetizing, and manipulating the animal twice, once to put on the equipment, the other to remove it. This procedure can modify the individual's behavior as well as expose it to social disadvantages while carrying the equipment. On the other hand, there is a high risk of not recovering the equipment at all! In this respect, our intention is to try out the higher-ranging Class 1 system with animals living in a natural semi-free environment, in order to resolve these issues. Also, we are working on reducing the weight and augmenting the energy supply in order to be able to increase the continuous recording time. The emphasis on securing long-term continuous recordings has to do with the need of studying the chronobiology of vocal behavior and also with the applicability of the method for the study of nocturnal primates.

Up until now, the method we have proposed is mostly useful for the specific conditions of captivity; however, we are currently working on developing the system and hope to solve the discussed disadvantages.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of gelada baboon (*Theropithecus gelada*): Acoustic structure and behavioral context. *Folia Primatologica, 55*(3–4), 109–132.

Baldwin, J., & Baldwin, I. (1976). Vocalizations of howler monkeys *Alouatta palliata* in Southwestern Panama. *Folia Primatologica, 26*, 81–108.

Brown, C., Gomez, R., & Waser, P. (1995). Old World monkey vocalizations: Adaptation to the local habitat. *Animal Behavior, 50*, 945–961.

Cheney, D., & Seyfarth, R. (1997). Reconciliatory grunts by dominant female baboons influence victims behaviour. *Animal Behavior, 54*, 409–418. doi:10.1006/anbe.1996.0438

deBont, F., Groenewegen, M., & Oomen, W. (1995). A high quality audio-coding system at 128 kb/s. *Memories of the 98th Audio Engineering Society Convention* (pp. 25–28). Paris, France.

Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behavior, 55*, 799–807. doi:10.1006/anbe.1997.0663

Fischer, J., & Hammerschmidt, K. (2002). An overview of the Barbary macaque, *Macaca sylvanus*, vocal repertoire. *Folia Primatologica*. 73,32–45.

Fischer, J., Kitchen, D., Seyfarth, R., & Cheney, D. (2004). Baboon loud calls advertise male quality: Acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology, 56*, 140–148. doi:10.1007/s00265-003-0739-4

Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological Reviews, 77*, 57–76. doi:10.1017/S1464793101005826

Ghazanfar, A. A., Flombaum, H. I., Miller, C. T., & Hauser, M. D. (2001). The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): Playback experiments with long calls. *Journal of Comparative Physiology, 187*, 27–35. doi:10.1007/s003590000173

Gil-da-Costa, R., Palleroni, A., Hauser, M., Touchton, J., & Kelley, P. (2003). Rapid acquisition of fan alarm response by a neotropical primate to a newly introduced avian predator. *Proceedings of the Royal Society B, 270*, 605–610. doi:10.1098/rspb.2002.2281

Gouzoules, H., & Gouzoules, S. (2000). Agonistic screams differ among four species of macaques: The significance of motivation-structural rules. *Animal Behavior, 59*, 501–512. doi:10.1006/anbe.1999.1318

Gros, J. (2006). Acoustic analysis and contextual description of food-associated calls in white-faced capuchin monkeys (*Cebus capucinus*). *International Journal of Primatology, 27*, 273–294. doi:10.1007/s10764-005-9012-1

Hauser, M. D. (1998). Functional referents and acoustics similarity: Field playback experiments with rhesus monkeys. *Animal Behavior, 55*, 1647–1658. doi:10.1007/s100710100093

Jorgensen, D., & French, J. (1998). Individuality but not stability in marmoset long calls. *Ethology, 104*, 729–742. doi:10.1111/j.1439-0310.1998.tb00107.x

Massey, G., & Wray, S. (2007). Audio coding for wireless applications. *EE Times*. Available at www.eetimes.com/design/audio-design/4015814/Audio-coding-for-wireless-applications

Mendes, F., & Ades, C. (2004). Vocal sequential exchanges and intragroup spacing in the Northern Muriqui *Brachyteles arachnoids hypoxanthus*. *Anales de la Academia Brasileira de Cienias, 76*, 399–404. doi:10.1590/S0001-37652004000200032

Mendez, M., Randrianambinina, B., Rabesandratana, A., Rasoloharijaona, S., & Zimmermann, E. (2008). Geographic variation in loud calls of sportive lemurs (*Lepilemur* spp.) and their implications for conservation. *American Journal of Primatology, 70*, 1–11. doi:10.1002/ajp.20554

Mennill, D. (2005). Other animal sounds of Santa Rosa: Spider monkey (*Ateles geoffroyi*), Retrieved June 15, 2009, from http://web2.uwindsor.ca/courses/biology/dmennill/CostaRica/CROthers.html

Mitani, J. C. (1985). Gibbon song duets and intergroup spacing. *Behaviour, 92*, 59–95.

Peng-Fei, F., Wen, X., Sheng, H., & Xue-Long, J. (2009). Singing behavior and singing functions of black-crested gibbons (*Nomascus concolor jingdongensis*) at Mt. Wuliang, Central Yunnan, China. *American Journal of Primatology, 71*, 539–547. doi:10.1002/ajp.20686

Ramos-Fernández, G. (2005). Vocal communication in a fission–fusion society: Do spider monkeys stay in touch with close associates? *International Journal of Primatology, 26*, 1077–1092. doi:10.1007/s10764-005-6459-z

Range, F., & Fischer, J. (2004). Vocal repertoire of sooty mangabeys (*Cercocebus torquatus atys*) in the Taï National Park. *Ethology, 110*, 301–321. doi:10.1111/j.1439-0310.2004.00973.x

Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. L. (1999). The meaning of grunt variants in baboons. *Animal Behavior, 57*, 583–592. doi:10.1006/anbe.1998.1031

Rukstalis, M., Fite, J. E., & French, J. A. (2003). Social change affects vocal structure in a Callitrichid primate (*Callithrix kuhlii*). *Ethology, 109*, 327–340. doi:10.1046/j.1439-0310.2003.00875.x

Sabatini, V., & Ruiz, C. (2008). Acoustical aspects of the propagation of long calls of wild *Leontopithecus rosalia*. *International Journal of Primatology, 29*, 207–223. doi:10.1007/s10764-008-9240-2

Scherer, K. R. (1989). Vocal correlates of emotional arousal and affective disturbance. In *Handbook of psychophysiology: Emotion and social behavior* (pp. 165–197). London: Wiley.

Sugiura, H. (2007). Effect of proximity and behavioral context on acoustic variation in the coo calls of Japanese macaques. *American Journal of Primatology, 69*, 1412–1424. doi:10.1002/ajp.20447

Teixidor, P., & Byrne, R. W. (1999). The "whinny" of spider monkeys: Individual recognition before situational meaning. *Behaviour, 136*, 279–308.

Whitehead, J. (1987). Vocally mediated reciprocity between neighbouring groups of mantled howling monkeys, *Alouatta palliata palliata*. *Animal Behavior, 35*, 1615–1627.