

Key issues in the study of primate acoustic signals

Marco Gamba & Cristina Giacoma

Dipartimento di Biologia Animale e dell'Uomo, Università di Torino, Via Accademia Albertina 13, 10100 Torino, Italy; e-mail: marco.gamba@unito.it.

Summary – *Fast computers and relatively inexpensive software have caused a revolution in the field of digital signal processing and this will undoubtedly affect the study of human and primate bioacoustics. These new or updated tools, together with cutting edge techniques in sound analysis, are reviewed and introduced to readers, and examples are provided to give a better understanding of the topics. This paper looks at the study of non-human primate vocalisations. Language capacity in primates is now investigated more often than vocal communication in its acoustical form. However, several areas in the study of primate vocal communication are still poorly known and several effective techniques are still applied only rarely to the study of primate calls. We will review some recent papers dealing with primate vocal communication, discuss results, techniques and the implications for future studies. The topics range from The Source-Filter Theory of sound production to vocal registers, from formant analysis to non-linear phenomena and the implications for primate phylogeny.*

Keywords – *Language, vocalisation, evolution, primate behaviour.*

The Source-Filter Theory of sound production

A widely accepted description of the mechanism underlying speech production in humans is the Source-Filter Theory of vowel production. This theory states that speech output signals are the result of two main processes, one related to vocal folds vibration, the other intervening during resonance in the vocal tract. Humans, other primates and most mammals have in common the basic structure of the larynx, but both human and non-human primate systems also share another component of the vocal apparatus: the supralaryngeal vocal tract. Although Gunnar Fant (1960) described the so-called Source-Filter Theory of speech production, the pioneering studies of Von Kempelen (1791) and Muller (1848) contained insights into the phonatory processes. Muller (1848) was able to show empirically that the sounds coming directly from the larynx differed from usual speech sounds. Sounds much more similar to speech phonemes could be achieved when a tube, or a system of tubes, was placed over the vocal folds,

simulating the passage of the vocal signal through the airways.

As postulated by Fant (1960), speech output is the combination of a source of sound energy (the larynx) and of the transfer function determined by the supralaryngeal vocal tract resonance (filter). The supralaryngeal vocal tract serves as a variable acoustic filter, the shape of which determines the phonetic quality of the sound.

Both in human and non-human primates, the sound generation process starts with the adduction of the two vocal cords. This event blocks the flow of air from the lungs, resulting in an increased subglottal pressure. When the pressure becomes greater than the resistance provided by the vocal folds, they open again. If the pressure conditions are steady, vocal cords will continue to open and close in a quasiperiodic way.

The frequency of puffs of air passing through the glottis determines the fundamental frequency (F0) of the laryngeal source. The column of air then vibrates into the supralaryngeal vocal tract, which consists of both the oral and nasal airways. The vocal tract serves as a filter, which suppresses the passage of sound energy at certain frequencies,

leaving others unchanged. Those frequencies at which the energy remains at its maximum, sustained by the supralaryngeal vocal tract, are called formants. The frequency value of formants is determined in part by the length, the shape and the volume of the vocal tract (Fig. 1). The vocal tract can be altered to a certain degree, into a variety of shapes, by means of the articulators. This means that, potentially, a wide range of sounds can be produced. In humans, for instance, each vowel sound is characterised by a specific configuration of the vocal tract (Fant, 1960; Flanagan, 1965; Lieberman & Blumstein, 1988). Thus, the filter transfer function is determined by this acoustically resonant system and it applies to the sound independent of its fundamental frequency.

Vocal tract elongation is thought to be associated with threatening vocalizations used in aggressive interactions in many non-human primate species. As demonstrated by Fitch (1997), lip protrusion or lowering of the larynx may play an important role in modifying the formant patterns in the vocalisations of macaques.

The fundamental frequency and its harmonics are the result of laryngeal vibration. However, both the amplitude at the vocal folds and the filter function determine the amplitude of each single harmonic in the output signal.

Several phenomena may cause the Source-Filter Theory idealization not to be achieved in either human or non-human primate vocal production. Both the open or closed quotients (respectively, the percentage of time in each cycle during which the folds are open or closed) can undergo relatively subtle changes that may result in a much more complex spectral pattern than the one ideally described above. It is, for instance, the case of oscillation patterns that, in the power spectrum, even lack harmonics or produce a non-parallel series of harmonics interleaved with the multiples of the fundamental frequency. In these cases, the source filter theory still applies, but the source characteristics do not match the expectations generated in the ideal model.

One of the issues that may arise at this point is whether or not a wide application of the Source-Filter Theory is legitimate in the case of non-human primates. The answer, following the reasoning above, is bound to be positive. We will later discuss applications of the Source-Filter model

to the study of non-human primate communication.

Recording a vocal signal

Audio recording is the most valuable basic tool for voice and vocalization analyses. The recording of the human voice for speech analysis is an intricate task (Dejonckere *et al.*, 2001), but researchers dealing with the study of the acoustic behaviour of captive and wild animals are faced with an even more difficult problem. As well as being challenged by economic, logistical, environmental and technical barriers, they need reliable and reasonably priced equipment, which may also need to be water resistant and energy efficient.

High-quality field and captive recordings can be obtained by using Digital Audio Tape (DAT) recorders, solid-state recorders, hard disk recording systems, professional tape and reel-to-reel recorders. No embedded filtering and manual recording volume control are must-have features. Nowadays, professional audio tape recorders and reel-to-reel machines are scarce on the market, mainly because of the digital revolution. A few models (e. g. Marantz CP430 or Nagra IV-S) can still be purchased new or second hand, but they often need extensive maintenance (e.g. for the tape-speed control, cassette deck mechanism, etc.). DAT recorders (e.g. Sony TCD-D100) definitely provide high quality recordings and affordable solutions, but these too are disappearing from the market. Both DAT recorders and, indeed, other high quality recorders, may be somewhat unreliable when used for a long time in extreme conditions, which means that there can be considerable problems with using DAT recorders in field situations. DAT recorders are very sensitive to humidity and will frequently develop operating problems when exposed to severe conditions for an extended period of time. Functionality can sometimes, but not always, be restored after 12 hours in a dry place.

Hard disk recorders use a high-capacity hard disk to record digital audio (or digital video). Even though they are relatively new systems, they appear to be a good alternative to the more traditional equipment as they provide several additional capabilities that may really speed up the work (e.g. immediate file access, editing, etc.). These systems

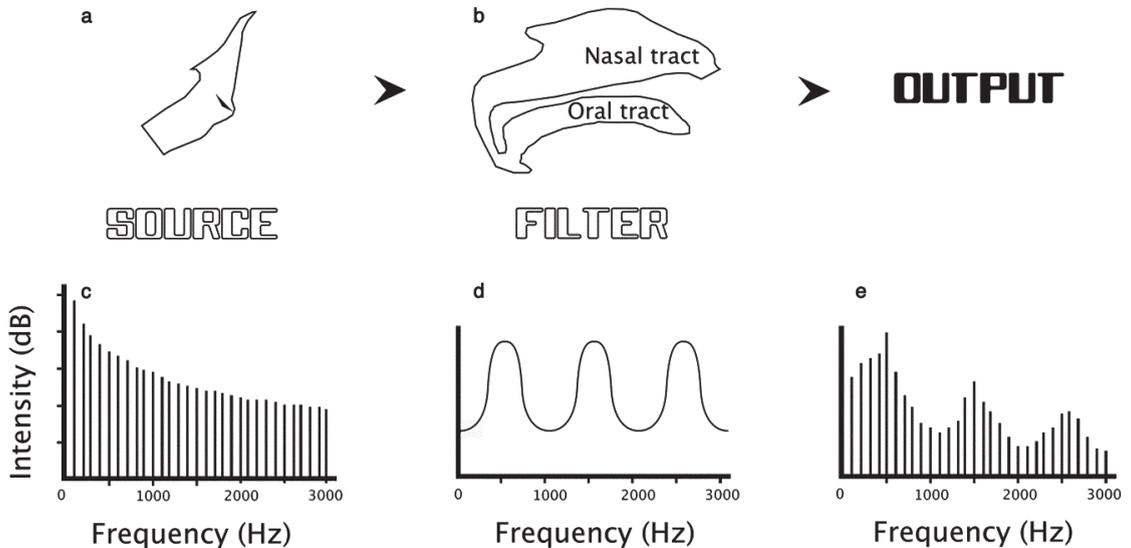


Fig. 1 – The Source-Filter Theory of speech production. Sound is generated by vocal fold vibration (a) and then passes through the supralaryngeal tracts (b), being modified by resonances in the cavities. The interaction between the glottal source spectrum (c) and the vocal tract transfer function (d) results in an output signal (e).

can be stand alone or computer-based. However, they still need to be tried out extensively in the field before they can be considered a reliable tool for use in recording the sounds of wild animals.

Solid-state recorders (e.g. Marantz PMD671) are probably the best news for wildlife sound investigators. They can record 24-bit, 96 kHz audio and, because of the memory stick system, they should be reliable even in forest conditions. These recorders are a promising novelty on the market, but, again, only extended usage will test their reliability in field conditions.

Equipment that removes information from the recorded sound by applying sound compression, such as MiniDisc, Digital Compact Cassette (DCC) and Mpeg Layer 3 (MP3) should be avoided. Whether these alter recordings of wildlife sounds is still controversial and is much debated. Therefore, we strongly discourage the use of these compressed formats to record and store sounds to be used for scientific analysis. Additional information on this topic can be found in several specialist

papers (Viellard, 1993; Kroodsma *et al.*, 1996; Hopp *et al.*, 1998; Geissmann, 2003).

With some sensible precautions, even digital camcorders can provide useful recordings. The main tip is to check the quality of the built-in microphone, possibly add a professional one, and to adjust the recording volume manually, discarding the function whereby audio is amplified to match what the picture is zooming in on (“audio zoom”). A high quality recorder is necessary, but this alone is not enough to ensure a high quality result.

Microphones are also very important. Apart from unwanted noises (vehicles, waterfalls, wind, etc.) the main obstacle to obtaining a high-quality recording is the distance the microphone is from the emitter of the desired sound. Most of the cheap commercial microphones will not be able to capture signals that are strong enough, unless they are placed very close to the subject. Once the microphone has been placed as close as possible to the emitter of the sound, researchers should rely on

the capabilities of ultra-directional microphones (e.g. Sennheiser ME88 with K3U module). The use of parabolic reflectors, windshields, pistol grips and shock mounts should also be considered (see the references cited above). In the field (e.g. when moving in the forest following a subject), it is very important to have secure connections between microphones and recorders. Sometimes it is helpful to fasten the plug connectors with adhesive tape. Headphones can be very useful to detect any problems caused by rapid movements and to monitor the signal being recorded.

Additional information about tape identification, equipment employed, date, place, weather/temperature and peculiar behavioural contexts are usually important and should be entered on the tape itself, preferably at the beginning of the tape (or file) and at the beginning of each of the recording sessions. They can then be replicated on the media case or cover.

Computers, via sound cards, can acquire recordings from analogue sources or DAT. There is a multitude of suitable soundcards that may be used to transfer natural sounds to a computer (some information can be found on the Internet, see Info on the Web below).

In recent years, cheap computer systems have reduced the cost and improved the accessibility of acoustic analysis methods. Sound analysis software is still something for specialists but can be easily found via the internet, sometimes at no cost. Excellent information about sound analysis software and freeware can be found on the Internet (again see websites cited in the Info on the Web section).

Sound analyses presented in this paper were done using PRAAT and CANARY.

PRAAT is a program for phonetic analysis that was developed at the Institute of Phonetic Sciences of the University of Amsterdam (Boersma & Weenink, 2005). PRAAT use can be combined with AKUSTYK, which is a comprehensive vowel analysis software package by B. Plichta (2005) at the Michigan State University. Both these programs are distributed under General Public License and are, therefore, free of charge.

CANARY is a software tool for the digital acquisition, manipulation, analysis and measurement of sound on Macintosh computers (it works

properly on computers booted under Mac OS 9.2 and older). It was developed by the Cornell Bioacoustics Research Program and provides a powerful user-friendly research tool in bioacoustics (Cornell Laboratory of Ornithology, 1993). It has, however, now been discontinued. Additional analyses were conducted using COLEA (Loizou, 2003), a MATLAB software tool for speech analysis (The MathWorks Inc., 1992).

Primate phonation

Studies of primate vocal communication are quite common. In contrast, primate phonation has rarely been investigated, either in terms of its mechanical properties or the implication for primates' linguistic ability (Lieberman, 1968; Brown *et al.*, 2003).

Vocal registers offer an interesting frame for the classification of vocal sounds in broad categories. Registers are series of consecutively phonated frequencies produced with nearly identical vocal quality and with minimum overlap in fundamental frequencies between adjacent registers (Hollien, 1974). A register is defined by Titze (1994) as a perceptual category that divides the regions of vocal quality into distinct regions that are maintained over some ranges of pitch and loudness. The register is a categorical feature which is quantally perceived and which is related to categorical perception.

Human vocal registers have been widely investigated and are usually defined in relation to the phonological functions of dynamic control of voice quality, including changes in the spectral structure or in the temporal perturbation (Laver, 1980). In the past, the vocal registers were often considered to be the consequence of the perceived fundamental frequency of voice, but recent works refer to vocal registers as a category of vocal fold vibrations (Zemlin, 1997).

Hollien (1974) introduced different registers for speaking and singing. He distinguished between pulse, modal and falsetto registers for speech, but he used chest, head and falsetto registers for singing. These labels are often regarded as the counterparts of the phonatory settings, although the register categories are more complex and involve supralaryngeal settings.

Normal speech is usually produced in the so

called modal register, in which fundamental frequency ranges between 100 and 300 Hz. Humans are also capable of producing sounds with fundamental frequencies above 300 Hz and this is called the falsetto register, or below 100 Hz, in the fry register. Fundamental frequency is not the only feature distinguishing between vocal registers, and additional information on this topic may be found in several papers (Hollien, 1974; Titze, 1994; Blomgren *et al.*, 1998).

In a recent paper by Riede and Zuberbühler (2003a), a fourth register, called the pulse register, was applied to the alarm calls of Diana monkeys (*Cercopithecus diana*). This separate register, covering the frequency range below 100 Hz but involving vibration of the ventricular folds, has been previously described in human phonation (Fuks *et al.*, 1998; Lindestad *et al.*, 2001). The ventricular folds are superior lateral pleats of laryngeal mucosa. Their role in phonation was judged irrelevant for a long time and is still poorly known.

However, empirical studies on non-human primate phonation are very complex to set up, requiring a wide range of technical appliances. Brown and colleagues (2003) studied the laryngeal biomechanics of the squirrel monkey, describing four different regimes of vocal fold activation. They recognised a first regime corresponding to human modal phonation. A second regime of vocal fold vibration showed different frequency stacks and was identified as biphonation. A third register exhibited a rhythmical series of glottal pulses. Finally, a fourth one lacked an apparent fundamental frequency and showed a quasi-irregular phonation mode.

Even though empirical experiments are really hard to set up due to the equipment required and to the shortage of dissected organs available, the study of vocal registers could really provide new criteria to define broad categories within the vocal production of particular species.

Although there is unquestionably an element of subjectivity in labelling primate vocalisations, what will emerge from a basic qualitative analysis of the vocal repertoire of non-human primates is the presence of tonal calls, low-pitched sounds and loud emissions. These general categories of vocalizations are shown in figures 2a-c. Pictured below are sounds emitted by a male black lemur

(*Eulemur macaco macaco*), consisting of a tonal call, long grunt and alarm call.

Tonal calls show a distinct fundamental frequency and harmonic overtones and are exchanged among conspecifics during short range communication.

In these highly harmonic utterances it is hard to distinguish spectral information (multiples of the fundamental frequency) from vocal tract resonances. However, with a good knowledge of the anatomical structure involved in phonation and information about articulatory manoeuvres, it is possible to detect them properly (Fig. 2a).

Low-pitched sounds are present in the vocal repertoire of most lemurs; they are frequently emitted and exhibit a noisy structure and wider energy bands (Fig. 2b). In this category, formants are much more recognizable than source-related parameters, but fundamental frequency and harmonics are masked and are very difficult to measure.

It has to be made clear, however, that both tonal calls and low-pitched sounds contain acoustic cues related to both laryngeal vibration (e.g. fundamental frequency) and vocal tract resonance (e.g. formants), even if one set of acoustic features is usually more evident than the other.

Loud call is a general term often used to indicate those vocal signals employed in long distance communication. Depending on the role of these calls and the species (e.g. advertisement calls, intra-group spacing, pair duets), these loud signals show a wide range of characteristics from tonal to harsh and noisy emissions. In the black lemur, the alarm call is often emitted when aerial predators are perceived (Fig. 2c).

Investigators who have studied the vocal behaviour of lemurs found have another interesting phenomenon: tonal calls and grunts were discovered to be associated in one utterance. This combination of harmonic calls and low-pitched emissions is found in all *Eulemur* species and is more often emitted when there is an increased level of arousal (Fig. 2d). These calls show an abrupt change in the spectral structure, involving, among other parameters, an impressive change of fundamental frequency values.

Considering the similarities across primate phonation, we can argue that tonal calls show a type of phonation similar to human modal register.

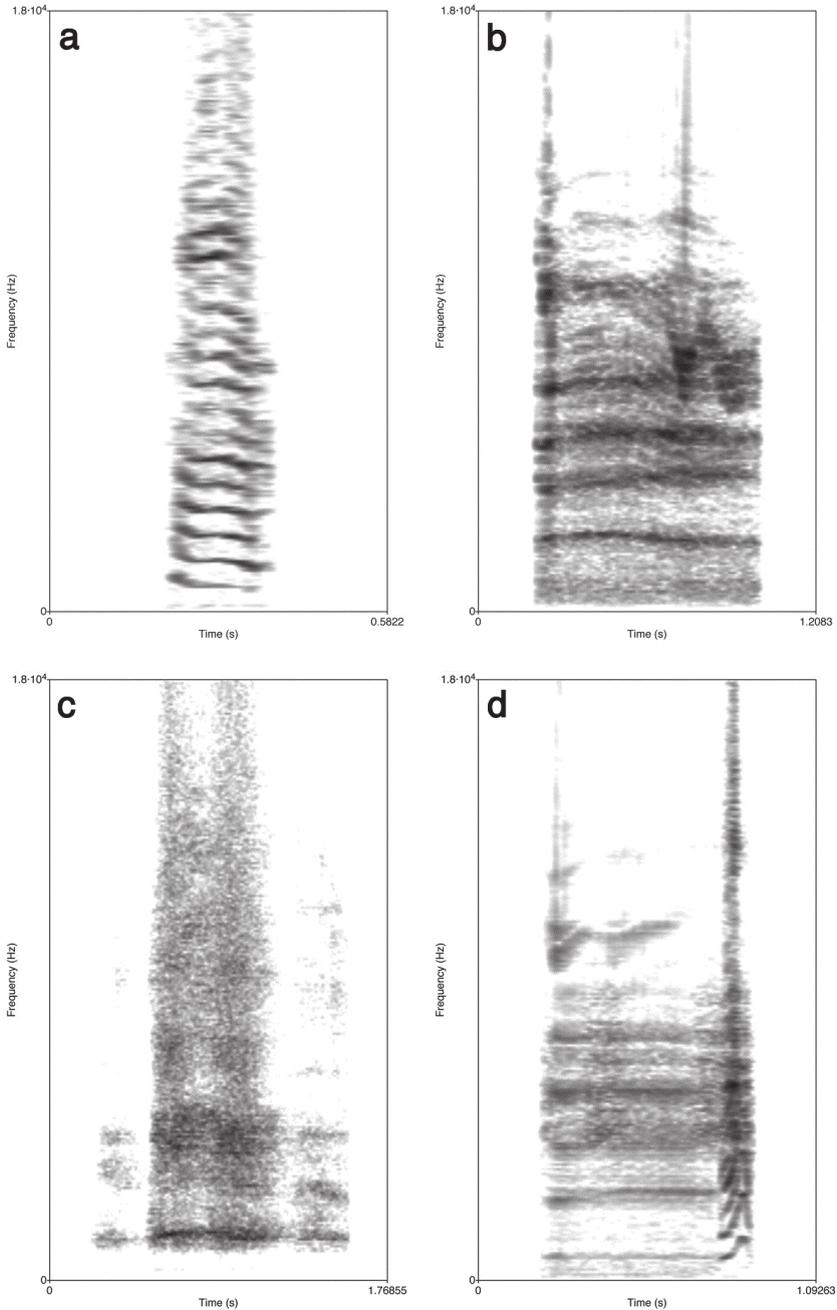


Fig. 2 – Spectrograms of 4 vocalizations emitted by a male black lemur (*Eulemur macaco macaco*): a) "long grunt" showing a pulse register phonation; b) "tonal call" produced with modal register; c) "alarm call" exhibiting a modal register with chaotic noise pattern; d) "long grunt clear call" presenting mixed pulse register and modal register phonation. Spectrograms display the variation in frequency (y-axis) over time (x-axis). Parameters used to generate all the representations in this paper are: dynamic range 35-40 dB; window length 0,05-1,5 s; pre-emphasis 7 dB/oct.

Long grunts exhibits pulse register phonation and the “long grunt - clear call” is a transition between pulse register and modal register.

The study of fundamental frequency

As said before, source-related acoustic cues refer to those aspects of vocalization associated with vocal fold vibration. The parameter most often used in the investigation of vocal communication is probably the fundamental frequency (F0), which corresponds to the rate of vocal fold vibration and is perceived as pitch in human speech.

There are a number of standard methods that researchers use to extract F0, based on various mathematical principles.

Theoretically F0 can be calculated from the waveform. The period is usually defined as the duration between maximum positive peaks, the inverse of which is the frequency of oscillation (Fig. 3a). The problem comes when the waveform consists of more than a simple sinusoid. In these cases (virtually all field recordings and natural sounds), the pitch value can be roughly measured from the spectrogram (Fig. 3b) and more precisely extracted with other methods. A power spectrum calculated with the Fast Fourier transform (FFT) algorithm for a limited number of frames can properly outline the fundamental frequency and its harmonics that appear as a number of prominent equally spaced components (Fig. 3c).

A second reliable way of obtaining an estimate of the fundamental frequency for long stationary vocal signals is to use the *cepstrum*. This is a Fourier analysis of the logarithmic amplitude spectrum of the signal. If the log amplitude spectrum contains many regularly spaced harmonics, then the Fourier analysis of the spectrum will show a peak corresponding to the spacing between the harmonics: the fundamental frequency.

Another powerful method is based on the autocorrelation algorithm, which exploits the fact that a periodic signal, even if it is not a pure sine wave, will be similar from one period to the next. The pitch extraction is provided taking a section of the signal, with a length at least twice as long as the longest period that we might detect. The autocorrelation function is defined as the sum of the pointwise absolute difference between the two signals over some intervals. The fundamental

period is identified as the first minimum of the autocorrelation function. Using the autocorrelation or cross-correlation algorithms it possible to determine the pitch contour of a signal, representing variation of the fundamental frequency in function of time (Fig. 3d).

Previous studies on the variation of fundamental frequency in the human voice showed that declination of the fundamental frequency spans coherent units of an utterance (Fourcin, 2002). This effect of vocal tract configuration was investigated in non-human primates. It was demonstrated that vocal production of vervet monkeys (*Cercopithecus aethiops*) and rhesus macaques (*Macaca mulatta*) showed fundamental frequency declination (Hauser & Fowler, 1992), suggesting that this phenomenon is not only widespread across languages but present in non-human primates too, possibly serving a communicative function.

In addition, the fundamental frequency of human phonation is involved in conveying information about the emotional state of the speaker (Williams & Stevens, 1972). In the human voice, it has been demonstrated that F0-related parameters (e.g. short-term perturbations, long-term variability, mean value, range, maximum) are among the measures often reported to correlate with elevated levels of emotional stress in the speaker and sometimes represent the primary indicators of stress among measured vocal cues (Protopapas & Lieberman, 1997).

In non-human primates, Leinonen and colleagues (2003) observed that vocalisations given by *Macaca arctoides*, previously categorised according to their social contexts and by human listeners, showed similar F0 levels to those of female human vocalisations attributed to the same context. In the same study, macaque vocalisations associated with “anger” and “fear” and human vowel samples, extracted from emotional-motivational simulations, both showed the highest fundamental frequency values.

In lemurs, for instance in the crowned lemur (*Eulemur coronatus*), various tonal calls are emitted in different contexts and their harmonic structure is frequently detectable. Low amplitude, tonal close calls are emitted when these lemurs are slowly progressing in close contact in the forest (Macedonia & Stanger, 1994). Quavered structured

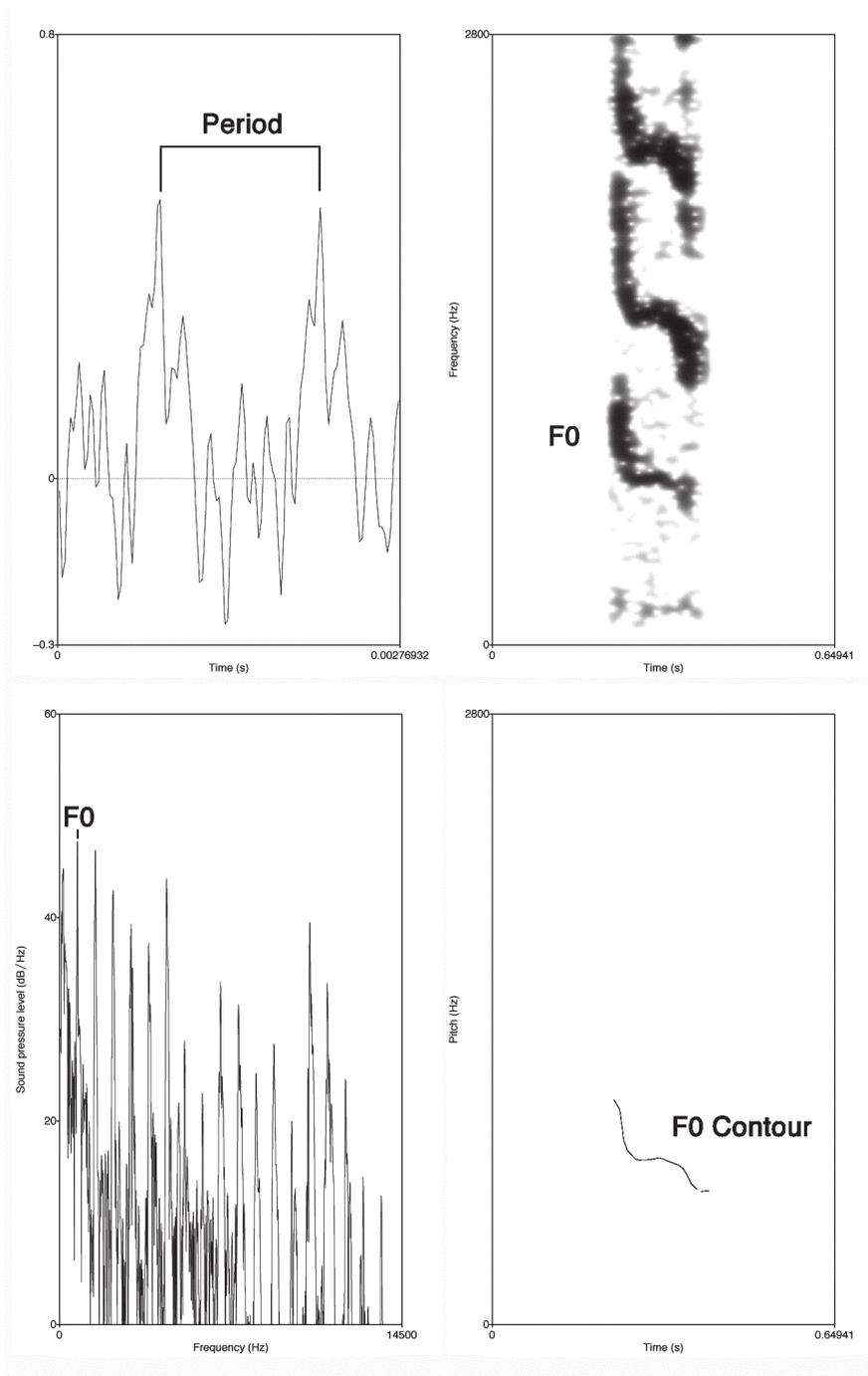


Fig. 3 – Standard methods that researchers use to extract F0 from a sound signal: a) waveform b) sonogram c) Fast Fourier Transform spectrum and d) pitch contour.

tonal calls or screams are given when they are stressed or manipulated (Petter & Charles Dominique, 1979).

The question we addressed in this example is whether or not tonal vocalizations emitted by crowned lemurs, in different behavioural contexts, show significant differences in their acoustic structure (e.g. the maximum F0). As shown in figure 4, the maximum F0 values of close contact calls and screams show striking differences. Scream vocalizations uttered in distress situations show higher maximum F0 values in all the recorded individuals.

These results are in agreement with several previous studies that demonstrated that non-human primates in severe behavioural conditions (e.g. separation from conspecifics, manipulation, high level of arousal) emit vocalization with strong frequency modulation and a very high maximum F0.

In crowned lemurs, however, such an large increase in the fundamental frequency value is rare or even absent in the rest of the vocal repertoire. On the other hand, vocalizations aimed at maintaining social closeness or social contact show a stable, clear harmonic structure with moderate frequency modulation.

The study of formants

The term “formants” was originally used to describe vocal tract resonances in speech signals (Fant, 1960). The study of formants in animal communication was started by Lieberman in the 1960s (Lieberman, 1968). Lieberman’s work showed that it is possible to detect formant-like phenomena in primate vocalizations and they were soon termed formants (Lieberman *et al.*, 1969; Owren & Bernacki, 1998; Fitch, 1997; Riede & Fitch, 1999). Even if the use of the term formants is widely accepted in many research fields, some scientists insist in avoiding the use of this term for newborn and non-human uttering. However, recognizable formant-like spectral features revealed to be formants when acoustic and anatomical measurements were combined (Owren & Bernacki, 1998; Fitch, 1997; Riede & Fitch, 1999). We suggest that the use of the term “formants” for non-human primates vocal tract resonances should be kept independently from either the F0/F1 ratio or

the vocal tract morphology.

Formants in non-human primates were studied using Linear Predictive Coding (LPC). LPC is one of the most powerful speech analysis techniques and a robust method for encoding good quality speech. It provides extremely accurate estimates of speech parameters and is relatively efficient for computation.

Linear predictive coding is an especially effective technique for estimating formants (Rabiner & Schafer, 1978). LPC is based on the assumption that a speech signal is produced by a source at the end of a tube (see the Source-Filter Theory above). The glottis produces a buzz, which is characterized by its intensity and frequency. LPC analyzes the speech signal by estimating the formants, removing their effects from the speech signal, and estimating the intensity and frequency of the remaining buzz. The process of removing the formants is called *inverse filtering*, and the remaining signal is called the *residue*.

Because speech signals vary with time, this process is done on short chunks of the speech signal, which are called *frames*. Usually 30 to 50 frames per second give intelligible speech with good compression.

The LPC system estimates formants from the speech signal computing a difference equation, which expresses each sample of the signal as a linear combination of previous samples. Such an equation is called a *linear predictor*, which is why this is called Linear Predictive Coding. The coefficients of the difference equation (the *prediction coefficients*) characterize the formants, so the LPC system needs to estimate these coefficients. The LPC system does the estimates minimizing the mean-square error between the predicted signal and the actual signal. This process involves the computation of a matrix of coefficient values and the solution of a set of linear equations. The convergence of this series of equations into a unique solution can be assured using different methods (e.g. autocorrelation, covariance, etc.)

Of course, there is a certain degree of approximation in this method, due to the assumption of the simple tube and no interaction between the oral and the nasal tracts. In fact, the real vocal tract shows side branches that may produce, as in human nasal sounds, what are usually called *zeros* or *anti-formants*, which require

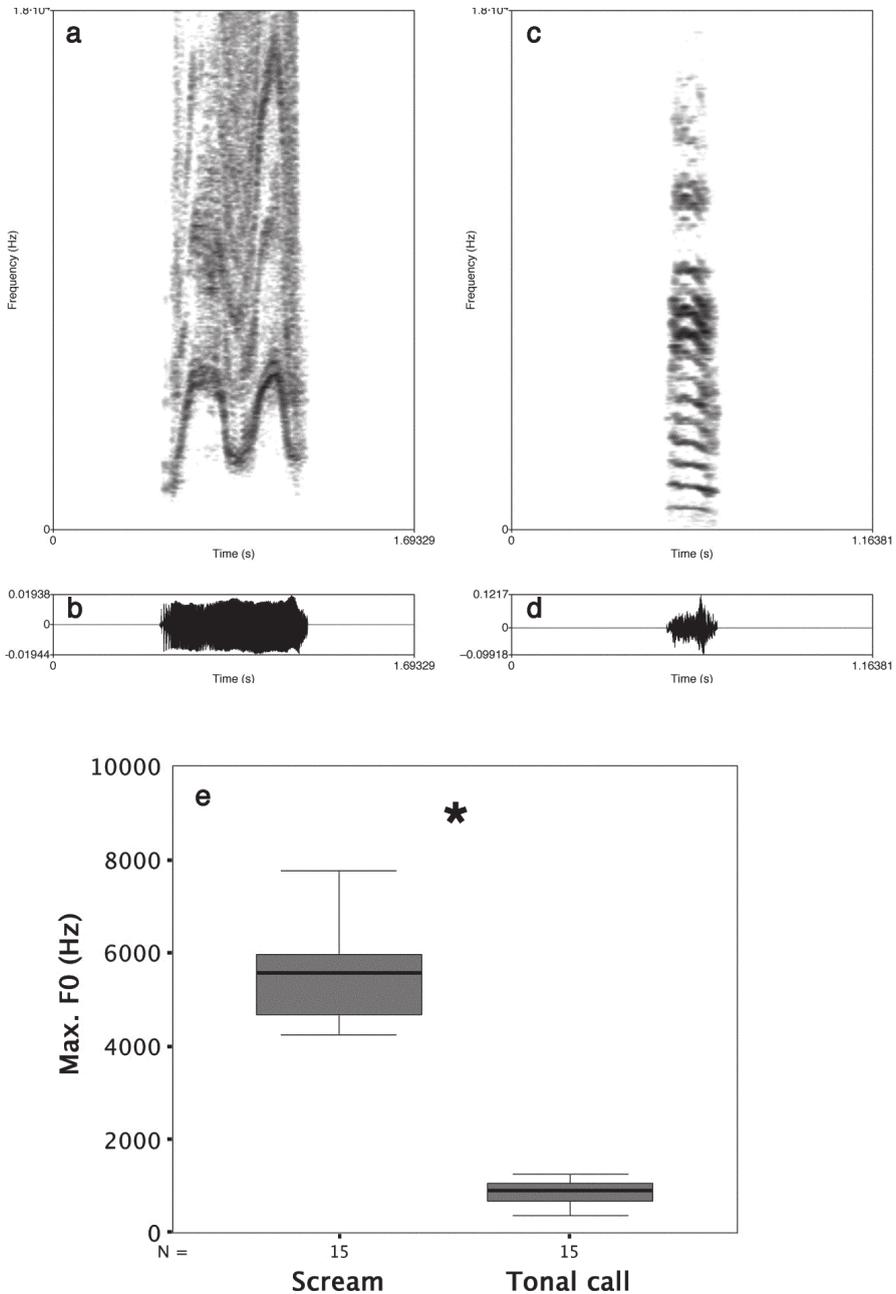


Fig. 4 - Spectrograms and oscillograms of screams (a, b) and tonal close contact calls (c, d) emitted by crowned lemurs (*Eulemur coronatus*). These two vocal types showed highly significant differences (e) in the maximum fundamental frequency (ANOVA on average individual means, $df=28$, $F=286,608$; $p<0,001$). Oscillograms display the variation in amplitude (dB, y-axis) over time (x-axis).

much more complex equations. Anti-formants are frequency regions in which the amplitude of the source signal are attenuated because the nasal cavities absorb energy from the sound wave. The effects of these absorptions are more marked in human nasal consonants because they are articulated with a complete occlusion of the oral cavity.

Detailed information about LPC analysis settings in mammal vocalization and working procedures can be found in the literature (Owren & Bernacki, 1998; Reby & McComb, 2003; Rendall *et al.*, 2005).

Further studies were then conducted showing that vocal tract resonances are meaningful for the vocal communication of non-human primate species and that different calls may show different formant patterns (Seyfarth *et al.*, 1994; Rendall *et al.*, 1998; Fitch, 1997; Riede & Zuberbuhler 2003a; 2003b).

In the Diana monkey, it has been shown that referential calls, such as leopard and eagle alarm vocalizations, differed mainly according to the modulation of the first formant (Riede & Zuberbuhler, 2003b).

Several technical tips should be considered when applying formant detection methods to primate vocalizations. It is very important to know any available information about the vocal tract anatomy because the length, the shape and the volume of the vocal tract determine the resonance phenomena. Knowledge of the basic anatomy of the subject species is very important to define analysis parameters in the software, in order to extract sound estimate of formant frequency and for the interpretation of the results. There could be important variation between species and specimens, therefore, analysis parameters should be adjusted and verified per species or per individual (Gamba & Giacoma, unpublished data).

It is common practice to verify LPC-detected poles (formants) by interpreting mammal phonation on the basis of the uniform tube model (Lieberman, 1968; Lieberman *et al.*, 1969; Lieberman and Blumstein, 1988; Shipley *et al.*, 1991; Reby & McComb, 2003). A few studies have suggested that the uniform vocal tract model might not properly explain some primate vocalization formant patterns (Owren *et*

al., 1997; Riede *et al.*, 2005a). First of all, it has to be appreciated that different calls in the vocal repertoire of a particular species can be the result of a different phonation mechanism and vocal tract configurations (Gamba & Giacoma, 2005).

A common phenomenon when analysing non-human primate calls is that they often show very low formants when compared to vocal tract length. For instance, Stevens suggested that lip narrowing may cause lowering of formants (see Riede *et al.*, 2005a).

The long debated capacity to lower the larynx is now very well described in several mammal species (Fitch & Reby, 2001; Fitch, 2000; Fitch, 2002), at least as a temporary phenomenon occurring during vocalisation. Of course, lowering of the larynx can represent another important source of variation in primate vocalisation, producing vocal tract elongation.

It has been recently proposed that non-uniform tube models should be applied to primate phonation. In fact, it has been demonstrated that formant frequencies measured in *Cercopithecus diana* alarm calls can be better simulated when 4-tube model is considered (Riede *et al.*, 2005a).

Studies on snake alarm calls and eagle alarm calls emitted by vervet monkeys (Owren & Bernacki, 1988) show that a non-human primate species possesses the ability to alter the vocal tract configuration to modify resonances in its vocalisations. Evidence of this ability had been observed earlier, in *Papio hamadryas* and *Theropithecus gelada*, but had not been described in detail (Andrew, 1976, Owren & Bernacki, 1988).

Recent studies by Riede and colleagues (2005a) made clear that non-human primates do produce changes in the shape of their vocal tracts during phonation. They also showed that changes in the configuration of the vocal tract may be investigated by describing variation in the formant pattern of the vocalisations (Riede *et al.*, 2005a).

It is well known that, in humans, changes in the vocal tract configuration are at the basis of the human phonatory ability and these changes affect the output vocal signal. As stated above, this has hardly been investigated in non-human primates and even less so in prosimians.

For the purpose of this paper, we therefore analyzed the correlation between vocal tract configuration and formant pattern in the red-

bellied lemur.

This species is conspicuously vocal and both captive and wild individuals have been observed emitting long grunts and tonal calls. Long grunts are always emitted with the mouth closed and some expulsion of air from the nose is often detectable. Tonal calls are emitted with the mouth open and show a clear harmonic structure as described above. Spectrograms of long grunts and tonal calls produced by an adult male of *Eulemur rubriventer* are shown in figure 5. Both spectrograms were produced using 50 ms as FFT window and Hanning algorithm. Note the different formant patterns between the long grunt and tonal call (see also LPC spectra positioned on left and right side of the spectrogram in Fig. 5).

In order to perform a quantitative analysis to test whether the two configurations produce different formant patterns or not, we measured first (F1) and second (F2) formants in a set of spontaneously emitted vocalizations. Formant values were extracted by LPC analysis in PRAAT software. The graph in figure 6 shows the F1/F2 plot for long grunts and tonal calls in *E. rubriventer*.

Mean values of the long grunts and tonal calls, emitted with closed and open mouth respectively, display significant differences in the formant values (Long grunt: F1 = 851 Hz, F2 = 1949; Tonal call: F1 = 1389 Hz, F2 = 2763; ANOVA F1: F1, 39 = 87,767, $p < 0,001$; ANOVA F2: F1, 39 = 30,337, $p < 0,001$).

Although primatologists must be ready to face extreme variation in the signal structure and in the phonation mechanism, it is evident that formant analysis has a great value in the interpretation of variability in primate acoustic signals and can provide valuable data in the daunting task of decoding vocal communication mechanisms in primates.

Non-linear phenomena in primate vocalisations

In the literature on vocal communication, the use of the term non-linear is sometime misleading. Several events in the phonatory process can be labelled as non-linear. We will, in this section, introduce the reader to some of the most common non-linear phenomena, recently described in non-human mammals, but we emphasizes the fact that

authors often disagree in what the term “non-linear” stands for.

In humans, non-linear phenomena occur more frequently in pathological voices (see Tokuda *et al.*, 2002 for extensive references). It is possible to observe, under certain conditions, an irregular oscillation of the vocal folds that may lead to non-linear phenomena, such as subharmonics, biphonation and deterministic chaos (Fitch *et al.*, 2002).

Subharmonics are important acoustic properties for objective evaluation of rough voices among voice technicians. They are known to occur in the cries of human infants (Mende *et al.*, 1990; Gobermann & Robb, 1999) and also in newborn lemurs (Gabutti *et al.*, 2004), because of the irregular vibration of the two vocal folds.

Subharmonics appear as frequency components located between two consecutive harmonics in the spectrum and can be produced because the natural vibratory frequencies of the two vocal folds are different, but synchronous. This phenomenon may occur, for instance, because there is more tension on one vocal fold than the other, because there are pathological modifications (polyps, polypoid degeneration, plicae ventricularis) or because vibration of the ventricular folds is regular enough to create a distinct independent stack of components.

Biphonation is characterized by a sequence of glottal cycles of different shapes and lengths. But, in this case, two glottal cycles are never identical and therefore biphonation is characterized by discrete spectra with irrational ratios between the frequencies stacks. A biphonation event can be induced either by left-right asymmetry or by desynchronized anterior-posterior vibratory modes.

Deterministic chaos in vocal signal can be the result of non-periodic, irregular vibrations produced by desynchronized coupled oscillators (Fitch *et al.* 2002). These chaotic events are characterized by a broadband spectrum with diffused energy (see Fig. 2c for an example) and are perceived as harsh and noisy sounds.

These non-linear phenomena have been described in animals with voice disorders (Riede, 2000) and in the vocal communication of several mammal species (see Wilden *et al.*, 1998). Species they have been detected in include: *Lycaon pictus* (Wilden *et al.*, 1998), *Canis lupus familiaris*

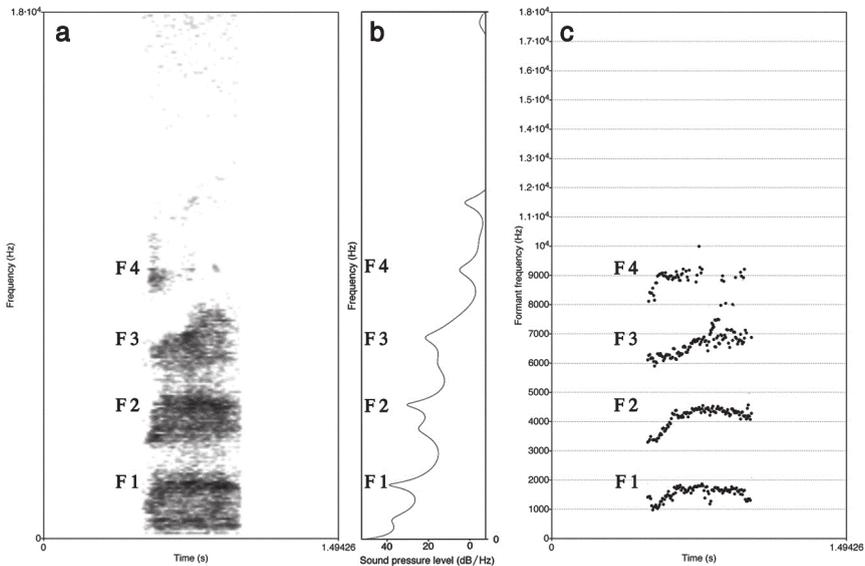


Fig. 5 – Formants in lemur vocalizations can be detected (a) from the spectrogram and measured (b) using an LPC spectrum. A formant-tracking diagram (c) can also be generated using software like PRAAT or COLEA for MATLAB.

(Riede *et al.*, 2000), *Canis latrans*, *Canis familiaris* (Riede *et al.*, 2005b) and *Cuon alpinus* (Volodin & Volodina, 2002).

Interest in non-human primate chaotic signals seems to have increased in the last few years. They have been described in *Macaca fuscata* (Riede *et al.*, 1997), *Macaca mulatta* (Fitch *et al.*, 2002), *Papio cynocephalus ursinus* (Fischer *et al.*, 2002), *Pongo* sp. (Davila Ross & Geissmann, 2004) and *Pan troglodytes* (Riede *et al.*, 2004).

Studies of the human voice have provided several measurements that could be effectively related to pathology (reviewed in Riede, 2000; Tokuda *et al.*, 2002). Among them, harmonics-to-noise ratio (HNR) and jitter and shimmer. The HNR is a measure that compares the acoustic energy of the harmonic components with that of the noise in time series (Awan & Frenkel, 1994; Qi & Hillmann, 1997). HNR has been found useful as a parameter for quantifying the noise within a signal. Furthermore, it has recently been used in order to describe different groups of atonal calls in canids (Riede *et al.*, 2000).

Even if they are rarely measured in the vocal signals of non-human primates, HNR, jitter and

shimmer can potentially provide information about the emotional status of non-human primate individuals. In fact, these parameters have often been used to investigate non-verbal cues in human vocal communication (e.g. Bachorowski, 1999; Pittam & Scherer, 1993).

In his study on baboon grunts, Rendall (2003) found that jitter did not vary consistently between high- and low-arousal conditions in all contexts. He commented that these results were consistent with previous mixed evidence on the relation between jitter and arousal or anxiety. Further investigations are needed, but Riede (2000) provided a possible methodological explanation: increase in the modulation noise causes broadened harmonics resulting in an overestimation of noise energy, consequently the minimum peaks of the spectrum are less deep and noise energy is overestimated.

Repertoire size and the classification of vocalizations

Ethograms are lists describing the inventory of behaviours observed in a certain species. In the same manner, the vocal repertoire can be defined as

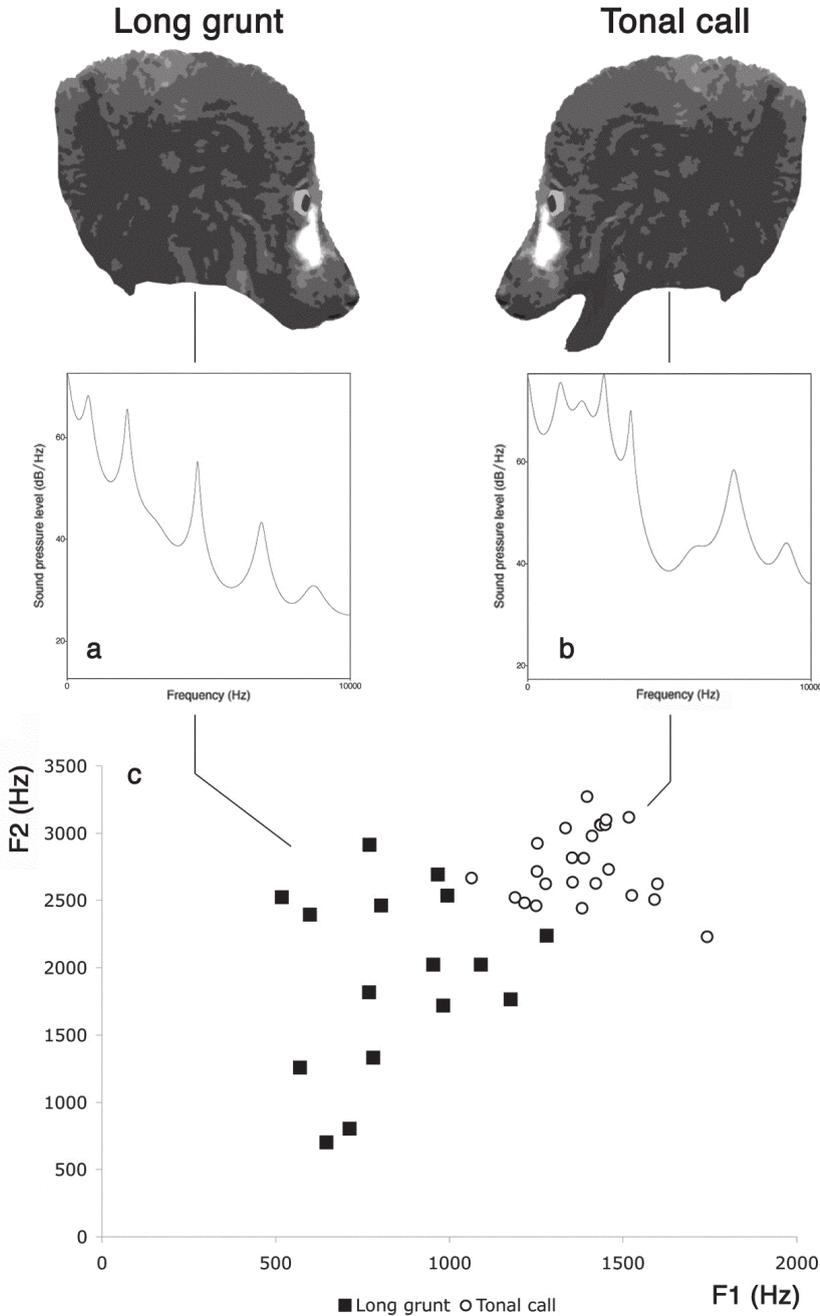


Fig. 6 – The long grunt is emitted with mouth closed and in absence of articulation (up, left), while tonal call exhibits open-mouth phonation mechanism (up, right). F1/F2 vowel space plot (c) shows first formants in *E. rubriventer* as derived from the LPC spectra. Typical LPC spectra of Long Grunt (a) and Tonal Call (b) are shown above the plot.

the list of calls that a species gives during social, affiliative, agonistic or aggressive interactions. The number of calls emitted by a certain species has sometimes been investigated in order to quantify the acoustic ability of that species, and these data have been related to its biological, physiological and social features. However, a basic, preliminary question still remains: How can we recognize categories within the vocal repertoire of a particular species? A common element, shared by most studies of vocal repertoires, is some kind of spectrographic representation of calls. However, it is not rare to find repertoires described on the basis of a qualitative analysis, relying on only observation and visual classification. Although we acknowledge the worth of these papers, especially when there is a context-functional account to the description of the different calls, the cheap computers and reliable, free software now available really encourage deepening the analytic level into variability and classification issues.

Primate vocal repertoires have been studied using many approaches, as different authors studying non-human animal communication systems have often chosen different methods to deal with the intricate task of dividing the vocal repertoire into different vocal types. Roughly, we can recognise 3 basic approaches: a contextual approach that may be very useful in identifying vocalizations emitted in different situations (Baldwin & Baldwin, 1976); a behavioural approach that identifies vocalizations and contiguous behaviours to infer functional properties of the calls (Pereira *et al.*, 1988); and an acoustic approach that will mainly look at the identification of the different vocal types on the basis of their physical properties. Of course, most studies combine some of the above to investigate stereotypy of behaviour in non-human primates and to avoid dangerous errors that may easily occur when only one point of view is taken into account. These approaches, when assumed in combination with the various applicable criteria and the different aspects that might be investigated, generate a number of methodological settings.

As a result, studies of vocal communication in non-human primates show many different approaches including looking at ontogenetic traits (Hammerschmidt *et al.*, 1994; 2001), gender characteristics (Fischer *et al.*, 2001; 2002), inter-

specific variability (de Waal, 1988), phylogenetic implications (Macedonia & Stanger, 1994), neuroethology (Jurgens, 1982), breathing rates (Riede & Zuberbühler; 2003a), perceptual units (Miller *et al.*, 2001, Ghazanfar *et al.*, 2001), anti-predator behaviour (Macedonia, 1990; 1993; Oda, 1996, Riede & Zuberbühler; 2003b), intra- and inter-individual variability (Zimmermann, 1995a; Hafen *et al.*, 1998; Oda & Masataka, 1996) and subspecific differences (Macedonia & Taylor, 1985), to name a few. Such a range of approaches is difficult to summarise in a comparative view. Although these studies provide important insights in the field of primate communication, most of them cannot be used as a source of information regarding the repertoire size and structure of these species.

The pioneering work of Marler (1965) and Gautier and Gautier-Hion (1982), offers a number of insights into the answers to several questions about vocal repertoire variability and the discreteness vs. continuum/gradedness of calls.

Qualitative comparisons of sounds can be made by visual evaluation of graphic representations, such as spectrograms or power spectrums. However, qualitative comparisons can be rarely utilised to discriminate between complex repertoires, where there is a need for investigation of variability of call types and variability across species or individuals. A quantitative assessment of a vocal repertoire is very often performed through using multivariate statistics. This approach is particularly suitable for classification tasks; especially when it involves the use of Discriminant Function Analysis (DFA). Discriminant analysis itself is a valuable tool to define which variable(s) discriminate best between groups (e.g. vocal types, species or individuals). Among the measurements taken from the various representations of a vocal signal, it is possible to show which ones are useful to discriminate between groups, and then to combine them into a number of classification scores ($n-1$). Stepwise Discriminant Function Analysis (SDFA) is a special variant of Discriminant analysis. It is a stepwise method of analysis that will enter variables into the model until there is no increase in the classification accuracy. This particular technique is helpful in providing the best set of variables to be combined in the classification functions. An additional property of DFA is that following one classification

run it is possible to utilise the resulting score to classify cases belonging to other samples or groups. More sophisticated discriminant procedures can also be employed, ranging from cross-validation to leave-one out tests. DFA was used by Fischer *et al* (2002) to identify differences between baboon loud calls emitted in different contexts and by various individuals. This study showed that contest *wahoos* and alarm *wahoos* emitted by male *Papio cynocephalus ursinus* could be discriminated by mean of a significant discriminant model based on their acoustic features. Community identity in wild chimpanzees (*Pan troglodytes verus*) was investigated by means of a DFA analysis of their pant hoots. Calls were correctly classified by means of a DFA and then results were tested via sets of permuted DFA (Crockford *et al.*, 2004). These additional tests are particularly useful when there is a risk of replication in the data submitted for analysis (e.g. individuals represented more than once in the groups).

We employed S DFA to classify vocalisation in the vocal repertoire of the black lemur. A preliminary visual analysis of the spectrograms made from a huge dataset of recordings from captive and wild individuals allowed the recognition of 13 spectrographically different calls. Of these, 11 were quantitatively characterized taking into account both source- and filter-related parameters and then submitted to statistical analysis. Multiple acoustic features were measured. Vocalizations were preliminarily divided into different categories according to their vocal type. This is a necessary step because the procedure requires specifying for each call the frequency range within which the fundamental frequency is expected. Source-related measures were collected using a PRAAT script procedure monitored step by step by the operator. This is important to ensure no distortion due to the semi-automatic procedure affected the analysis. Measurements of source-related parameters were mean, standard deviation, maximum and minimum F0. The average fundamental frequency variation per unit time and the average number of inflexions in F0 contour per unit time were also calculated (after Reby & McComb, 2003). We also collected average estimates of jitter, shimmer and HNR for each call and the percentage of voice and unvoiced frames.

Using another PRAAT procedure, we extracted filter-related parameters as first six formants (F1-F6) and bandwidths (B1-B6), plus the respective standard deviations.

Even if S DFA classification is relatively robust, it is preferable to perform some collinearity diagnostics before entering all the measurements into the discriminant model.

S DFA applied to discriminate between vocalisations in *E. macaco* showed it is possible to assign 94% percent of the vocalizations (N = 263; Wilk's Lambda = 0,005, F10, 227 = 304,893; $p < 0,001$) to the correct vocal type (Grunt = 91,1%; Grunted Hoot = 90,4%; Long Grunt = 81,8%; Click = 100%; Hoot = 100%; Long Grunt Clear Call = 94,7%; Snort = 100%; Alarm call = 93,8%; Chatter = 100%; Scream = 100%; Tonal call = 95,5%). In the leave-one out classification, 88,9% of cross-validated grouped cases was correctly assigned. Figure 7 shows the plot of the first two canonical roots derived S DFA.

The one described above is a relatively simple application of the S DFA that allows a statistically significant validation of *a priori* defined vocal categories.

The use of techniques used for speech recognition in humans can be considered for the analysis of non-human primate vocalisations. Various applications of Artificial Neural Network (ANN) techniques can be applied to the study of acoustic signals (Rabiner & Schafer, 1978). One of the easiest applications involves the processing of data tables that could otherwise be submitted to DFA or other statistical analyses. However this kind of application of neural properties is weak in many senses and hybrid in its nature. More sophisticated networks (e.g. Back Propagation networks) can be designed for a more precise classification of vocalisations, especially by analyzing a limited set of parameters at each time slice of a vocal signal.

Nowadays, researchers' attention is focused on establishing new techniques that may provide an independent, objective call classification, based on the acoustic structure of the call. Studies using these new methods have been applied to several mammal and bird species, including killer whales (Deeke *et al.*, 1999), bats (Burnett & Masters, 1999), deer (Reby *et al.*, 1997) and zebra finches (Werfel, 2001). Unsupervised neural networks are rarely applied to calls of non-human primates, mainly

because of the complex pattern of primate vocal communication. However, some of these neural techniques, such as self organising maps and digital time warping, will be useful to solve these matters (Thorn, 2004).

In recent years, new evidence has suggested the need to investigate primate vocal repertoires more broadly, defining acoustically different categories of calls by objective analysis and then investigating animals' responses to these different categories.

None of the methods mentioned above can be taken as the final solution to the study of the primate vocal repertoire. Call classification will provide statistically homogeneous groups of calls but will not investigate the non-human primates' point of view on this matter. Playback experiments can be very misleading if, for instance, they are not based on preliminary work that identifies the

correct calls to be played back to the subject animals.

However, we shall note that it is still unclear, and unlikely, that all the call types could evoke a detectable univocal response in the subjects. With exception of few strongly referential vocalizations, call types are often emitted in multiple contexts.

Species-specific potential in non-human primate vocalizations

The species-specific vocalizations of non-human primates play a key role in reproduction, survival and social interaction. Among primate species, specificity of vocalizations has been investigated in a limited number of taxa, with most of the recent studies being on perceptual and neural mechanisms (Wang, 1998; 2000; Ghazhanfar *et al.*, 2001). The impressive long calls of gibbons are

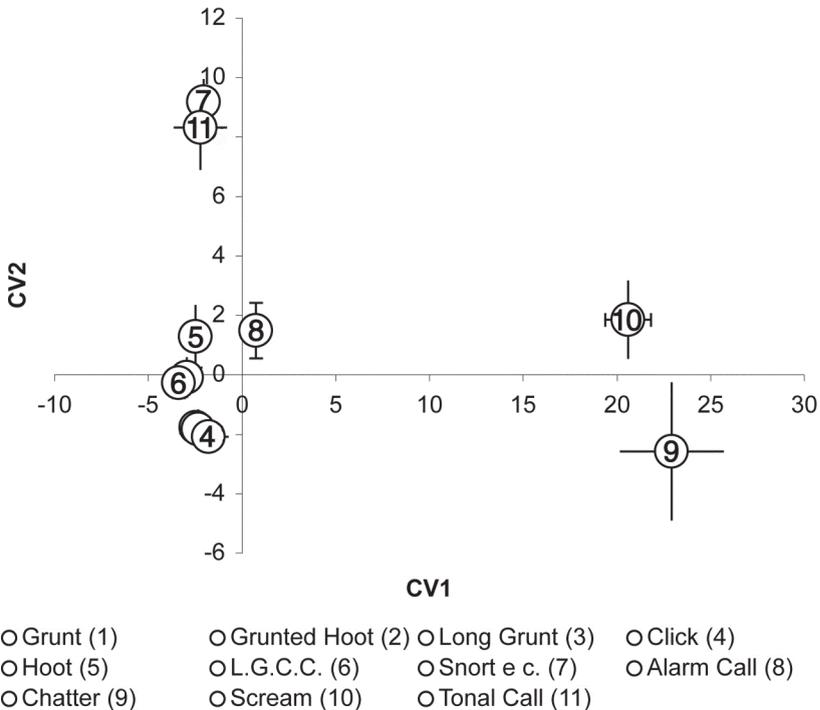


Fig. 7 – Plot of the mean discriminant scores for each vocal type in the vocal repertoire of the black lemur. Correct classification was 94% (Grunt = 91,1%; Grunted Hoot = 90,4%; Long Grunt = 81,8%; Click = 100%; Hoot = 100%; Long Grunt Clear Call = 94,7%; Snort = 100%; Alarm call = 93,8%; Chatter = 100%; Scream = 100%; Tonal call = 95,5%). In the leave-one out classification, 88,9% of cross-validated grouped cases were correctly assigned.

strongly influenced by species and context (Marshall & Marshall 1976; Marler & Tenaza 1977; Geissmann 1993, 1995). There is also some evidence supporting the idea that genetic inheritance is the most important determinant of the acoustic feature of gibbon calls (Marshall & Marshall 1976; Marler & Tenaza 1977; Geissmann 1993, 1995, 2002).

As Zimmermann (1995b) emphasises, loud calls, usually employed in long distance communication among conspecifics, are an ideal subject for the investigation of species-specificity among non-human primates (Petter & Charles Dominique, 1979; Snowdon, 1989; Gautier & Gautier-Hion, 1977; Marshall & Marshall, 1976, Mitani, 1987, Hohmann and Fruth, 1994).

However, more complete approaches to the study of behaviour, including vocal behaviour, and its species-specific potential should be encouraged (Giacoma & Balletto, 1988). We know little about intraspecific and interspecific variation across the entire vocal repertoire and such data would probably open new perspectives in the study of the genetic basis of primate vocal communication and species-specific recognition.

Acoustic discrimination of vocalizations in vertebrate species is based on several cues, as tested by studies ranging from human (Orlikoff & Baken, 1999) and non-human primate sounds (Crockford *et al.*, 2004; Fischer *et al.*, 2002) to amphibians (Sullivan *et al.*, 1995; Rosso, 2003). Thus, a multivariate analysis is suitable for such a complex task, and DFA is particularly effective (see above). Although several discriminant techniques can be applied, the use of individual average values is preferred. This method will avoid replication problems that may affect the resulting percentage of correct assignment. Discriminant analysis can, then, be used to determine which variables are the best predictors of which group subjects should be assigned to (e.g. Thumser *et al.*, 1996).

Thus, we adopted SDFA to analyze the species-specific potential in the acoustic structure of calls emitted by *Eulemur* spp. To provide an example, we chose a complex vocalization called long grunt clear call, which occurs naturally in the repertoire of wild and captive individuals. These vocalizations could be elicited by different stimuli but they occur only when the level of arousal is high and animals are moving very rapidly across the trees or into the

enclosure.

For the purpose of this paper, we analyzed long grunt clear call (LGCC) vocalizations emitted by 25 captive lemurs, 5 individuals for each species belonging to genus *Eulemur* (*E. macaco*, *E. fulvus*, *E. mongoz*, *E. coronatus*, *E. rubriventer*).

Calls were analyzed with PRAAT and source- and filter-related measures were saved into a text file and then exported into MICROSOFT EXCEL. We derived some parameters from previous measurements as the index of formant dispersion (Fitch, 1997), estimated vocal tract length (Reby & McComb, 2003) and the average number of periods per time unit (Gamba, 2005). All calculated parameters and measured variables were then imported into statistical software, SPSS 11 FOR MACOS and tested for multicollinearity. Suitable parameters were then used to calculate average individual means, then we submitted these mean values to the SDFA.

Among the measures we took, related to the fundamental frequency, formants and duration, three were determined to be the ones that discriminated best between groups. They were *sumvar* (the cumulative F0 variation, see Reby & McComb, 2003 for details; $F = 24,63$), the frequency of unvoiced frames ($F = 75,96$) and the average number of period per time unit ($F = 23,67$).

The SDFA used to discriminate among species showed Factor 1 primarily based on the percentage of unvoiced frames, while Factor 2 was primarily based on the average number of periods per time unit. The DFA model correctly categorized 80% of *E. macaco*, 100% of *E. fulvus*, 66,7% of *E. mongoz*, 57,1% of *E. coronatus* and 60% of *E. rubriventer*. The discriminant model was highly significant and correctly classified 76% of cases according to species (Wilk's Lambda = 0,022; $F_{12,47} = 25,068$, $p < 0,001$). Mean canonical scores of the first two discriminant functions are shown in figure 8.

Vocalizations as a taxonomic tool

Ornithologists have, on a vast scale, undertaken the investigation of taxonomic relationships among species using vocalizations, (Kroodsma, 1977; Catchpole, 1980; Payne, 1986; Thumser *et al.*, 1996; McCracken & Sheldon, 1997; Baptista & Martinez, 2002). In a number of cases, taxonomists

have described several subspecies on the basis of acoustical differences. Bioacoustic studies have also suggested that some subspecies should be considered to be different species (reviewed in Alstrom, 2001). Moreover, studies of sounds emitted by birds have also stimulated further work that lead to the discovery of new species (Goodman *et al.*, 1996).

In primate studies, vocal characteristics have been used to assess systematic relationships among several species (Gautier, 1988; Geissmann, 1993; Oates & Trocco, 1983; Macedonia & Taylor, 1985; Snowdon, 1993; Struhsaker, 1981; Wilson & Wilson, 1975; Zimmermann *et al.*, 1988; Zimmermann, 1990). Among the non-human primate studies, the most intriguing and complex application of communication characters for a phylogenetic reconstruction is the one proposed by Macedonia and Stanger (1994).

The use of acoustic analysis data for phylogenetic

reconstruction presents a very difficult step in the definition of characters to be submitted to the phylogenetic programs. The most common problem in this phase is to arbitrarily create distinct characters where there is, in fact, a continuum of variation within a certain parameters. Macedonia and Stanger work showed a robust determination of the characters, mostly based on the vocal types and in few cases on the acoustic structure of calls.

Data presented in that paper revisited the phylogeny of Lemuridae using vocal, visual and olfactory characters. We present here a slightly revised version of this analysis that includes more recent information on the vocal repertoires of some Eulemur species (Gamba & Giacoma, 2001). Following Macedonia and Stanger (1994), we used PAUP 4.0 for the phylogenetic analysis and used their methodology. Our results (Fig. 9) are consistent with the ones presented by Macedonia and Stanger and with those from several DNA

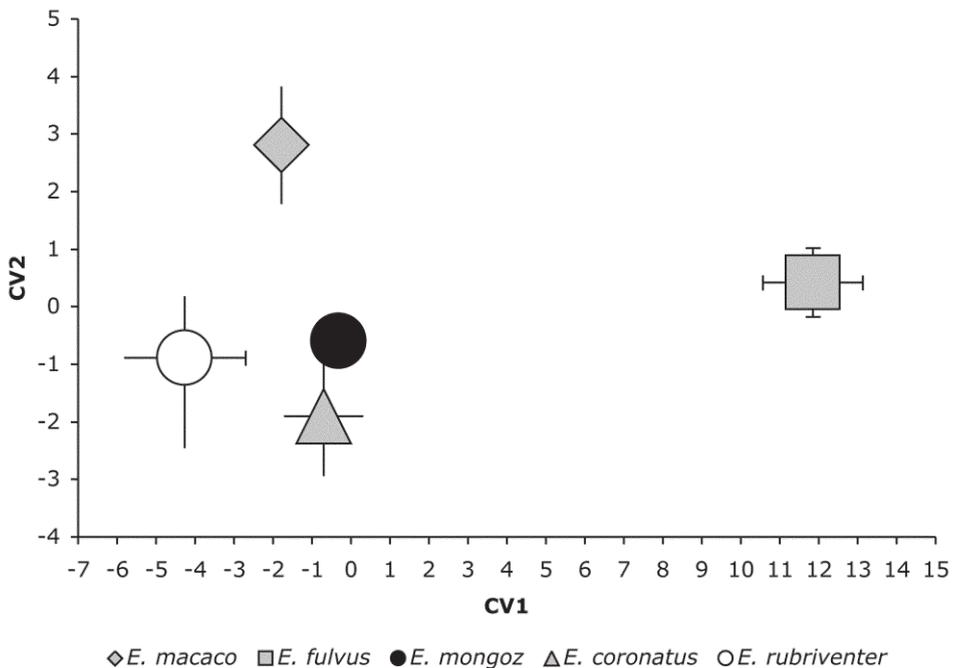


Fig. 8 - A scatterplot of the similarity between the long grunt clear calls of the five Eulemur species (*E. macaco*, *E. fulvus*, *E. mongoz*, *E. coronatus*, *E. rubriventer*). Mean canonical scores of the first two functions for each species is shown in figure. Score were derived from the stepwise Discriminant analysis.

studies (for other references see Macedonia and Stanger, 1994).

Interdisciplinary implications

The evolution of language is one of the most debated topics in biology nowadays, and it is difficult to imagine a more interdisciplinary topic. Language and communication studies should be, and commonly are, intimately bound to other subjects, for instance, cognitive abilities, social factors and gestures.

Research teams in vocal communication are usually made up of motivated individuals well adapted to an interdisciplinary research effort. They need to develop skills and acquire knowledge in several different fields. For instance, data collection involves not only the technical ability to track behavioural displays but also to record high-quality audio and video documents that can then be analysed. An ideal team should integrate interdisciplinary competences, possibly being made up of physicists, biologists, psychologists, computer scientists and engineers. All these scientists, from a wide variety of disciplines, need to work closely together in order to study non-human animal and human vocalizations. Behavioural data and vocalization can be analyzed and combined to determine the behavioural context of vocalizations, factors influencing vocalization structure and the overall relationship between vocalizations and social behaviour.

The world of primatology, in most cases, owes a debt to the study of the human voice, often making use of new techniques that come from various fields, such as linguistics, voice pathology and forensic science.

Language acquisition and evolution as well as child development and genetic determinants are topics that have been of interest to anthropologist for a long time, while sociolinguistic and ontogenetic studies are mainstream approaches in the study of modern human language (see Duranti, 2004). However, if it is true that the investigation of communication in non-human primates will be a big step in the knowledge of the evolution of language, there is still a lot of work to undertake.

What we know about primate vocal communication is still very limited, even more so in terms of the evolutionary perspective. Traditional

methods used in tracing the evolutionary history of humans are not helpful when considering the evolution of language. Though it may be true that the most distinctive anatomical feature of humans is the central nervous system and that there is no close analogue of human language in other primate species, we should also realise that the data from non-human primate studies may throw light on the evolutionary processes beyond language and speech (see Fitch, 2000) and that few authors have investigated communicative ancestral patterns and capacities (Lieberman & Crelin 1971).

Interestingly, one recent examination of vocal repertoires across species found associations between group size and repertoire size in non-human primates. Although the data used in this study came from methodologically different approaches, the study represents one of the first statistically significant and solid attempts to connect social factors and vocal complexity in non-human primates (McComb & Semple, 2005).

To what extent human and animal communication is genetically determined is still an open question. The study of vocal communication in primates may contribute crucial data to anthropology about the role that genetic factors play in vocal behaviour. Even if the ontogenetical causes of certain vocal behaviours are poorly known, it has been demonstrated that the acquisition of the vocal repertoire in non-human primates is, to a large extent, genetically driven. This means that non-human primates know innately how to emit proper vocal sounds. Genetically programmed calls are also found in human vocalisations. For instance, even across different languages, specific emotions are reflected in the variation of certain acoustical parameters (see Jurgens, 1990).

The study of primate vocalizations has suffered for a long time from widespread scepticism. In fact, the use of vocal signals as a phenotypical trait is a powerful, underestimated tool for the investigation of behavioural variation in multi-levelled systems. This may result from a large-scale investigation of variability and stereotypy in vocalizations, across vocal types and across taxa. A more detailed consideration of intra- and inter- individual variability and intra- and interspecific variation would really help in clarifying sources of variation in the acoustic signals of primates.

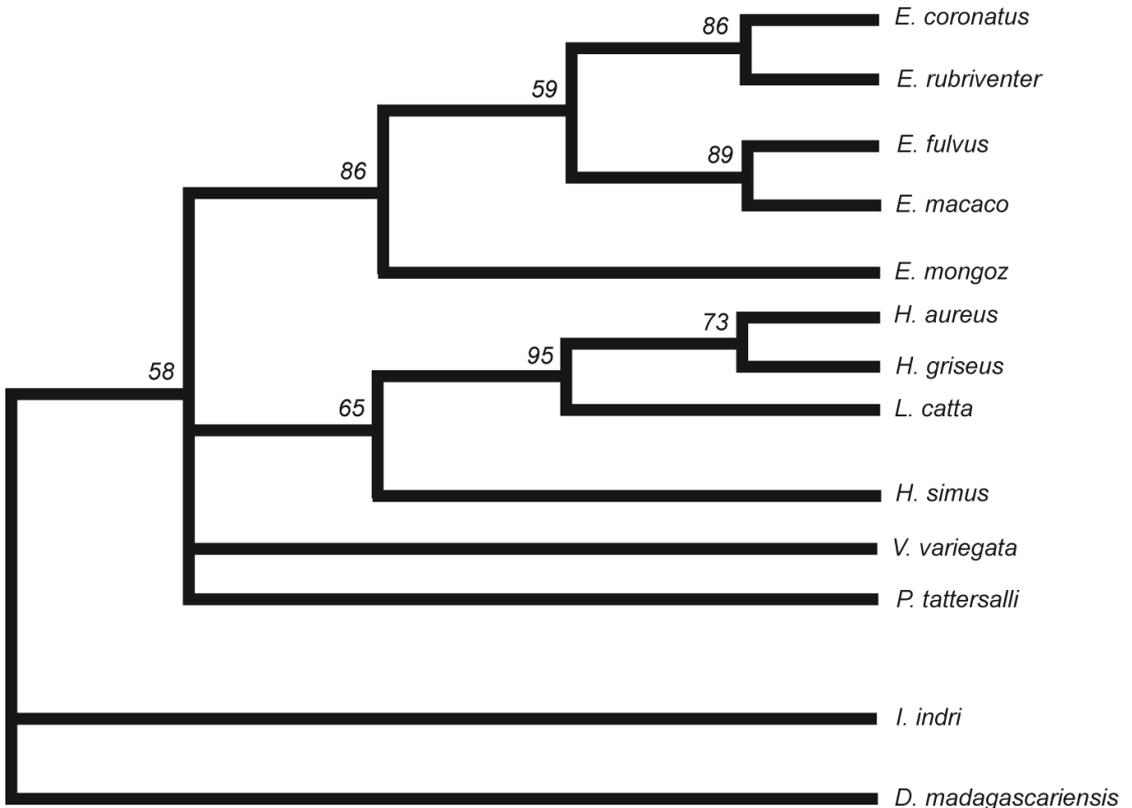


Fig. 9 – Single most parsimonious tree produced with bootstrap method following instructions reported in the original paper (Macedonia & Stanger, 1994). Numbers indicate the percentage of time that the bootstrap sampling procedure produced the clade shown to the immediate right of each number in a total of 1000 replicates.

Acknowledgments

We thank Luca Pozzi for his help in the phylogenetic analysis and Sergio Castellano and Alessandra Rosso for advice and support throughout the preparation of this paper. Patrizia Messeri, Franca Ligabue Stricker, Caroline Harcourt, Melchiorre Masali, Emiliano Bruner and two anonymous referees provided helpful comments on earlier versions of this manuscript. The collection of recordings analysed for the purpose of this paper was supported by grants to M.G. by Parco Natura Viva – Garda Zoological Park and the University of Torino. We thank Cesare Avesani

Zaborra for his continued support.

Info on the Web

<http://zeeman.ehc.edu/envs/Hopp/>

Extremely useful compilation of web resources, by well-known scientist Steve Hopp, which provides software and information for bioacousticians.

<http://www.fon.hum.uva.nl/praat/>

The official PRAAT website provides manuals and instructions on this free, powerful tool for phoneticians. It also provides tutorials for speech analysis and synthesis and regular updates about

PRAAT new features.

<http://www.msu.edu/~plichtab/>

As introduced in the paper, AKUSTYK is a comprehensive vowel analysis software package. This website also provides information about signal acquisition, educational tools, field recording and linguistic studies.

<http://www.haskins.yale.edu/>

Haskins Laboratories is a private non-profit research institute with a primary focus on speech, language and reading, and their biological basis. This website provides an extensive introduction to speech analysis and shows various approaches to these studies.

<http://www.ibac.info/>

The International Bioacoustics Council was founded in the 1960s and its objective is to promote international participation throughout the entire field of bioacoustics activity. It regularly organizes international symposia all around the world.

http://www.epanorama.net/documents/pc/soundcard_tips.html

Provides detailed information about sound cards, audio-acquisition systems and hardware configuration.

References

- Alstrom P. 2001. *The Use Of Sounds In Bird Systematics*. Uppsala University, Uppsala.
- Andrew R.J. 1976. Use of formants in the grunts baboons and other nonhuman primates. *Ann (N.Y.) Acad. Sci.*, 280: 673-693.
- Awan S.N. & Frenkel M.L. 1994. Improvements in estimating the harmonics-to-noise ratio of the voice. *J. Voice*, 8: 255-262.
- Bachorowski J.A. 1999. Vocal expression and perception of emotion. *Current Directions in Psychological Science*, 8: 53-57.
- Baldwin J.D. & Baldwin J.I. 1976. Vocalizations of howler monkeys (*Alouatta palliata*) In southwestern Panama. *Folia Primatol.*, 26: 81-108.
- Baptista L.F. & Martínez Gómez J.E. 2002. La investigación bioacústica de las aves del archipiélago de Revillagigedo: un reporte de avance. *Huitzil* 3: 33-41.
- Blomgren M., Chen Y., Ng M.L. & Gilbert H.R. 1998. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J. Acoust. Soc. Am.*, 103: 2649-2658.
- Boersma P. & Weenink D. 2005. PRAAT: doing phonetics by computer; <http://www.praat.org>
- Brown C.H., Alipour F., Berry D.A. & Montequin D. 2003. Laryngeal biomechanics and vocal communication in the squirrel monkey (*Saimiri boliviensis*). *J. Acoust. Soc. Am.*, 113: 2114-2126.
- Burnett S.C. & Masters W.M. 1999. The use of neural networks to classify echolocation calls of bats. *J. Acoust. Soc. Am.*, 106: 2189.
- Catchpole C.K. 1980. Sexual selection and the evolution of complex songs among European warblers of the genus *Acrocephalus*. *Behaviour* 74: 149-166.
- Cornell Laboratory of Ornithology 1993. CANARY: The Cornell Bioacoustics Workstation (Cornell Laboratory of Ornithology, Ithaca, NY) Version 1.4.2
- Crockford C., Herbinger I., Vigilant L. & Boesch C. 2004. Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, 110: 221-243.
- Davila Ross M. & Geissmann T. 2004. Biphonation in orangutan long calls. *Folia Primatol.*, 75: 368-369.
- Deecke V.B., Ford J.K. & Spong P. 1999. Quantifying complex patterns of bioacoustic variation: use of a neural network to compare killer whale (*Orcinus orca*) dialects. *J. Acoust. Soc. Am.*, 105: 2499-2507.
- Dejonckere P.H., Bradley P., Clemente P., Cornut G., Crevier-Buchman L., Friedrich G., Van De Heyning P., Remacle M. & Woisard V. 2001. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of phonosurgical treatments and evaluating new assessment techniques: Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society ELS. *Eur. Arch. Otorhinolaryngol.*, 258: 77-82.
- Duranti A. (ed.) 2004. *Companion to Linguistic Anthropology*, Blackwell Publishing, Malden.
- Fant G. 1960. *Acoustic Theory of Speech Production*.

- Mouton, The Hague.
- Fischer J., Hammerschmidt K., Cheney D.L. & Seyfarth R.M. 2001. Acoustic features of female chacma baboon barks. *Ethology*, 107: 33–54.
- Fischer J., Hammerschmidt K., Cheney D.L. & Seyfarth R.M. 2002. Acoustic features of male baboon loud calls: Influences of context, age, and individuality. *J. Acoust. Soc. Am.*, 111: 1465–1474.
- Fitch W.T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.*, 102: 1213–1222.
- Fitch, W.T. 2000. The evolution of speech: a comparative review. *Trends Cogn. Sci.*, 4: 258–267.
- Fitch W.T. 2002. Comparative Vocal Production and the Evolution of Speech: Reinterpreting the Descent of the Larynx. In Alison Wray, editor, *The Transition to Language*, pp. 21–45. Oxford University Press, Oxford.
- Fitch W.T. & Reby D. 2001. The descended larynx is not uniquely human. *Proc. Roy. Soc. London*. 268: 1669–1675.
- Fitch W.T., Neubauer J. & Herzel H. 2002. Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63: 407–418.
- Flanagan J.L. 1965. *Speech Analysis Synthesis and Perception*. Springer-Verlag, New York.
- Fourcin A. 2002. Patterns, Structure and Measurement in Voice Pathology. *First International Conference on Speech Sciences ICSS*, Seoul.
- Fuks L., Hammarberg B. & Sundberg J. 1998. A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glottographic evidences. *TMH-QPRS* 3, 49–59.
- Gabutti A., Gamba M. & Giacoma C. 2004. Subharmonics in *Varecia variegata variegata* vocal production. *Folia Primatol.*, 75: 377–378.
- Gamba M. 2005 – *Evoluzione della Comunicazione Vocale nei Lemuri del Madagascar*. PhD Thesis, Università degli Studi di Torino, Italy.
- Gamba M. & Giacoma C., 2001. Vocal communication in Lemuridae: new evidences in *Eulemur* and *Varecia*. *Advances in Ethology*, 36. *Supplement to Ethology*. Contributions to the XXVII International Ethological Conference, p. 160, Tuebingen, Germany.
- Gamba M. & Giacoma C. 2005. Vocal production mechanisms in ruffed lemurs: a prosimian model for the basis of primate phonation. *MAVEBA 2005*, Firenze University Press, Florence.
- Gautier J.P. 1988. Interspecific affinities among the guenons as deduced from vocalizations. In: Gautier-Hion A., Bourlier F., Gautier J.P. & Kingdon J. (eds), *A Primate Radiation: Evolutionary Biology of the African Guenons*, pp. 194–226, Cambridge University Press, Cambridge.
- Gautier J.P. & Gautier-Hion A. 1977. Communication in old world monkeys. In: Sebeok T. (ed) *How animals communicate*, pp. 890–964, Indiana University Press, Bloomington.
- Gautier J.P. & Gautier-Hion A. 1982. Vocal Communication within a Group of Monkeys: Analysis by Biotelemetry. In: *Primate Communication*, Snowdon C.T., Brown C.H. & Petersen M. (eds), pp. 5–29, Cambridge University Press, Cambridge.
- Geissmann, T. 1993. *Evolution of communication in gibbons (Hylobatidae)*. Ph.D. thesis, Anthropological Institute, Philosoph. Faculty II, Zurich University.
- Geissmann T. 1995. Gibbon systematics and species identification. *Int. Zoo News*, 42: 467–501.
- Geissmann T. 2002. Taxonomy and evolution of gibbons. In: Anthropology and primatology into the third millennium: The Centenary Congress of the Zürich Anthropological Institute. *Evol. Anthropol.*, 11(1). Soligo C., Anzenberger G. & Martin R.D. (eds.), pp. 28–31, New York: Wiley-Liss.
- Geissmann T. 2003. Tape-recording primate vocalizations. In: *Field and Laboratory Methods in Primatology: A Practical Guide*, Setchell J.M. & Curtis D.J. (eds), Cambridge University Press, Cambridge.
- Ghazanfar A.A., Flombaum J.T., 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. *J. Comp. Physiol.*, 187: 27–35.
- Giacoma C. & Balletto E. 1988. Phylogeny of the salamandrid genus *Triturus*. *Boll. Zool.* 55: 337–360.

- Goberman A.M. & Robb M.P. 1999. Acoustic examination of preterm and full-term infant cries: The long-time average spectrum. *J. Speech Lang. Hear. Res.*, 42: 850-861.
- Goodman, S. M., Andrianarimisa A., Olson L. E. and Soarimalala V. 1996. Patterns of elevational distribution of birds and small mammals in the humid forests of Montagne d'Ambre, Madagascar. *Ecotropica*, 2: 87-98.
- Hafen T., Neveu H., Rumpel Y., Wilden I. & Zimmermann E. 1998. Acoustically dimorphic advertisement calls separate morphologically and genetically homogenous populations of the Grey Mouse Lemur (*Microcebus murinus*). *Folia Primatol.*, 69: 342-345.
- Hammerschmidt K., Ansoorge V. & Fischer J. 1994. Age-related variations in the vocal repertoire of Barbary macaques. In: *Current Primatology*, Roeder J.J., Thierry B., Anderson J. R. & Herrenschildt N. (eds), pp. 287-295. Université Louis Pasteur, Strasbourg.
- Hammerschmidt K., Freudenstein T. & Jurgens U. 2001. Vocal development in squirrel monkeys. *Behaviour*, 138: 1179-1204.
- Hauser M.D. & Fowler C. 1992. Fundamental frequency declination is not unique to human speech: Evidence from nonhuman primates. *J. Acoust. Soc. Am.*, 91: 363-369.
- Hohmann G. & Fruth B. 1994. Structure and use of distance calls in wild bonobos (*Pan paniscus*). *Int. J. Primatol.*, 15: 767-782.
- Hollien H. 1974. On vocal registers. *J. Phonetics*, 2: 125-143.
- Hopp S.L., Owren M.J. & Evans C.S. (eds.) 1998. *Animal Acoustic Communication. Sound analysis and Research Methods*. Springer-Verlag, Germany.
- Jurgens U. 1982. A neuro-ethological approach to the study of vocalizations in the squirrel monkey. In: Snowdon C.T., Brown C.H. & Petersen M.R. (eds), *Primate Vocal Communication*, pp. 50-62, Cambridge Univ. Press, Cambridge.
- Jurgens U. 1990. Vocal communication in primates. In: Kesner R.P. & Olton D.S. (Eds.), *Neurobiology of Comparative Cognition*, pp. 31-42, Lawrence Erlbaum Associates, Hillsdale.
- Kempelen von W. 1791. *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*, J.V. Degen, Wien.
- Kroodsma D.E. 1977. Correlates of song organization among North American wrens. *Am. Nat.* 111: 995-1008.
- Kroodsma D.E., Budney G.F., Grotke R.W., Vielliard J.M.E., Gaunt S.L.L., Ranft R. & Veprintseva O.D. 1996. Natural sound archives: guidance for recordists and a request for cooperation. In: Kroodsma D.E. & Miller E.H. (eds), *Ecology and Evolution of Acoustic Communication in Birds*, pp. 474-486. Cornell University Press, Ithaca, NY.
- Laver J. 1980. *The Phonetic Description of Voice Quality*. Cambridge University Press, Cambridge.
- Leinonen L., Laakso M.L., Carlson S. & Linnankoski I. 2003. Shared features in emotional vocalization of *Macaca arctoides* and man. *Logoped. Phoniatr. Vocol.* 28: 12-21.
- Lieberman P. 1968. Primate vocalization and human linguistic ability. *J. Acoust. Soc. Am.*, 44: 1574-1584.
- Lieberman P. & Blumstein S.E. 1988. *Speech physiology, speech perception, and acoustic phonetics*. Cambridge Studies in Speech Science and Communication. Cambridge University Press, Cambridge.
- Lieberman P. & Crelin E.S. 1971. On the speech of Neanderthal man. *Linguistic Inquiry*, 2: 203-222.
- Lieberman P., Klatt D.H. & Wilson W.A. 1969. Vocal tract limitations on the vocal repertoires of rhesus monkey and other non-human primates. *Science*, 164: 1185-1187.
- Lindestad P., Soedersten M., Merker, B. & Granqvist S. 2001. Voice source characteristics in Mongolian 'throat singing' studied with highspeed imaging technique, acoustic spectra, and inverse filtering. *J. Voice*, 15: 78-85.
- Loizou P. 2003. COLEA; <http://www.utdallas.edu/~loizou/speech/colea.htm>
- Macedonia J.M. 1990. *Vocal communication and antipredator behavior in the ringtailed lemur (Lemur catta)*. PhD Thesis, Duke University.
- Macedonia J.M. 1993. Adaptation and phylogenetic constraints in the antipredator behavior of ring-tailed and ruffed lemurs. In: Kappeler P.M. & Ganzhorn J.U. (eds). *Lemur social systems and their ecological basis*, pp 67-83. Plenum Press, New York.

- Macedonia J.M. & Stanger K.F. 1994. Phylogeny of the Lemuridae revisited: Evidence from communication signals. *Folia Primatol.*, 63: 1-43.
- Macedonia J.M. & Taylor L.L. 1985. Subspecific divergence in a loud call of ruffed lemur (*Varecia variegata*). *Am. J. Primatol.*, 9: 295-304.
- Marler P. 1965. Communication in monkeys and apes. In: DeVore I. (ed), *Primate Behavior: Field Studies of Monkeys and Apes*, pp. 544-584. Holt, Rinehart and Winston, New York.
- Marler P. & Tenaza R., 1977. Signaling behavior of apes with special reference to vocalization. In: Sebeok T.A. (ed), *How animals communicate*, pp. 965-1033. Indiana University Press, Bloomington and London.
- Marshall J.T. & Marshall E.R. 1976. Gibbons and their territorial songs. *Science*, 193: 235-237.
- McComb K. & Semple S. 2005. Coevolution of vocal communication and sociality in primates. *Biol. Lett.* doi:10.1098/rsbl.2005.0366
- McCracken K.G. & Sheldon F.H. 1997. Avian vocalizations and phylogenetic signal. *Proc. Natl. Acad. Sci. USA* 94: 3833-3836.
- Mende W., Herzog H. & Wermke K. 1990. Bifurcations and chaos in newborn infant cries. *Phys. Lett.* 145: 418-424.
- Miller C.T., Miller J., Gil-da-Costa R. & Hauser M.D. 2001. Selective phonotaxis by cotton-top tamarins (*Saguinus oedipus*). *Behaviour*, 138: 811-826.
- Mitani J. C. 1987. Species discrimination of male song in gibbons. *Am. J. Primatol.*, 13: 413-423.
- Müller J. 1848. *The physiology of the senses, voice and muscular motion with the mental faculties*. W. Baly, Walton & Maberly, London.
- Oates J.F. & Trocco T.F. 1983. Taxonomy and phylogeny of black and white colubus monkeys: Inferences from analysis of loud call variation. *Folia Primatol.*, 40: 83-113
- Oda R. 1996. Effects of contextual and social variables on contact call production in free-ranging Ringtailed Lemur (*Lemur catta*). *Int. J. Primatol.*, 17: 191-205.
- Oda R. & Masataka N. 1996. Interspecific responses of ringtailed lemurs to playback of antipredator alarm calls given by Verreaux's sifakas. *Ethology* 102: 441-453.
- Orlikoff R.F. & Baken R.J. 1999. *Clinical Measurement of Speech and Voice*, Thomson Delmar Learning.
- Owren M.J. & Bernacki R.H. 1988. The acoustic features of vervet monkey alarm calls. *J. Acoust. Soc. Am.*, 83: 1927-1935.
- Owren M.J. & Bernacki R.H. 1998. Applying Linear Predictive Coding (LPC) to frequency-spectrum analysis of animal acoustic signals. In: Hopp, S.L., Owren, M.J. & Evans, C.S. (eds.), *Animals Acoustic Communication*, pp. 129-162. Springer, Berlin.
- Owren M.J., Seyfarth R.M. & Cheney D.L., 1997. The acoustic features of vowel-like grunt vocalization in chacma baboons (*Papio cyncephalus ursinus*). Implications for production processes and function. *J. Acoust. Soc. Am.*, 101: 2951-2963.
- Payne R. B. 1986. Bird songs and avian systematics. *Current Ornithology* 3: 87-126.
- Pereira M.E., Seeligson M.L. & Macedonia, J.M. 1988. The behavioral repertoire of the black-and-white ruffed lemur, *Varecia variegata variegata* (Primates: Lemuridae). *Folia Primatol.*, 51:1-32.
- Petter J.J. & Charles-Dominique P. 1979. Vocal communication in prosimians. In: *The Study of Prosimian Behavior*, Doyle G.A. & Martin R.D. (eds). Academic Press, New York.
- Pittam J. & Scherer K. R. 1993. Vocal expression and communication of emotion. In Lewis M. & Haviland J. (eds), *The handbook of emotions*, p. 185-197. Guilford, New York.
- Plichta B. 2005. AKUSTYK; <http://www.msu.edu/~plichta>
- Protopapas A. & Lieberman P. 1997. Fundamental frequency of phonation and perceived emotional stress. *J. Acoust. Soc. Am.*, 101: 2264-2277.
- Qi Y. & Hillman R.E. 1997. Temporal and spectral estimations of harmonics-to-noise ratio in human voice signals. *J. Acoust. Soc. Am.* 102: 537-543.
- Rabiner L. & Schafer R. 1978. *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs.
- Reby D. & McComb K. 2003. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, 65: 519-530.
- Reby D., Lek S., Dimopoulos I., Joachim J., Lauga J. & Aulagnier S. 1997. Artificial neural networks as a classification method in the behavioural sciences. *Behavioural Processes*, 40: 35-43.

- Rendall D. 2003. Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J. Acoust. Soc. Am.*, 113: 3390-3402
- Rendall D., Owren M.J. & Rodman P.S. 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalization. *J. Acoust. Soc. Am.*, 103: 602-614.
- Rendall D., Kollias S, Ney C. & Lloyd P. 2005. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *J. Acoust. Soc. Am.*, 117: 944-955.
- Riede T. 2000. *Vocal changes in animals during disorders*. PhD thesis. Humboldt-Universität, Berlin.
- Riede T. & Fitch T. 1999. Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). *J. Exp. Biol.* 202: 2859-2867.
- Riede T. & Zuberbuhler K. 2003a. Pulse register phonation in Diana monkey alarm calls. *J. Acoust. Soc. Am.*, 113: 2919-2926.
- Riede T. & Zuberbuhler K. 2003b. The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization. *J. Acoust. Soc. Am.*, 114: 1132-1142.
- Riede T., Wilden I. & Tembrock G. 1997. Subharmonics, biphonations, and frequency jumps. common components of mammalian vocalization or indicators for disorders. *Z. Säugetierkunde* 62: 198-203.
- Riede T., Owren M.J. & Arcadi A.C. 2004. Nonlinear acoustics in pant hoots of common chimpanzees (*Pan troglodytes*): frequency jumps, subharmonics, biphonation, and deterministic chaos. *Am. J. Primatol.*, 64: 277-290.
- Riede T., Bronson E., Hatzikirou H. & Zuberbuhler K. 2005a. Vocal production mechanisms in a non-human primate: morphological data and a model. *J. Hum. Evol.*, 48: 85-96.
- Riede T., Herzel H., Mehwald D., Seidner W., Trumler E., Bohme G. & Tempbrock G. 2000. Nonlinear phenomena in the natural howling of a dog-wolf mix. *J. Acoust. Soc. Am.*, 108: 1435-1442.
- Riede T., Mitchell B.R., Tokuda I. & Owren M.J. 2005b. Characterizing noise in nonhuman vocalizations: Acoustic analysis and human perception of barks by coyotes and dogs. *J. Acoust. Soc. Am.*, 118: 514-22.
- Rosso A., 2003. *Ciclo vitale, morfologia e canto nelle ragnelle italiane: uno studio di variazione geografica*. PhD thesis, Università degli Studi di Torino, Italia.
- Seyfarth R.M. & Cheney D.L., Harcourt A.H. & Stewart, K.J. 1994. The acoustic features of gorilla double grunts and their relation to behavior. *Am. J. Primatol.*, 33: 31-50.
- Shiple C., Carterette E.C. & Buchwald J.S. 1991. The effect of articulation on the acoustical structure of feline vocalization. *J. Acoust. Soc. Am.*, 89: 902-909.
- Snowdon C.T. 1989. Vocal communication in New World monkeys. *J. Hum. Evol.* 18: 611-633.
- Snowdon C.T. 1993. A vocal taxonomy of the callitrichids. In: Rylands A. (ed), *Marmosets & Tamarins: Systematics, Ecology & Behavior*, pp 78-94, Oxford Univ. Press, Oxford.
- Struhsaker T.T. 1981. Vocalizations, phylogeny & paleogeography of the red colobus monkey (*Colobus badius*). *Afr. J. Ecol.*, 19: 265-283.
- Sullivan B.K., Ryan M.J. and Verrell P.A., 1995. Female Choice and Mating System Structure. In: Heatwole H. & Sullivan B.K. (eds), *Amphibian Biology, Social Behaviour*, pp. 469-517, Surrey Beatty and Sons, NSW.
- The MathWorks Inc. 1992. MATLAB. The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Massachusetts 01760, U.S.A.
- Thorn A. 2004. Mapping Vocal Repertoires: Automated Classification, and Mapping of Relationships between Vocalisations. *ASAB Easter Meeting*, 2004, University of Sussex, Brighton.
- Thumser N.N., Karron J.D. & Ficken M.S. 1996. Interspecific variation in the calls of spheniscus penguins. *Wilson Bulletin.*, 108: 72-79.
- Titze I. R. 1994. *Principles of Voice Production*, Prentice Hall, Englewood Cliffs.
- Tokuda I., Riede T., Neubauer J., Owren M.J. & Herzel H. 2002. Nonlinear Analysis of Irregular Animal Vocalizations. *J. Acoust. Soc. Am.*, 111: 2908-2919.
- Vielliard J. 1993. Recording wildlife in tropical rainforest. *Bioacoustics* 4: 305- 311.
- Volodin I.A. & Volodina E.V. 2002. Biphonation as a prominent feature of the Dhole *Cuon*

- alpinus* sounds. *Bioacoustics*, 13: 105-120.
- Waal F.B.M. de 1988. The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour*, 106: 183-251.
- Wang X. 1998. What is the neural code of species-specific communication sounds in the auditory cortex of Primates? In: Palmer A.R., Rees A., Summersfield Q. & Meddis R., *Psychophysical and Physiological Advances in Hearing*, pp. 521-528, Whurr Publishers Ltd., London.
- Wang X. 2000. On cortical coding of vocal communication sounds in primates. *Proc. Natl. Acad. Sci. USA*, 97: 11843-11849.
- Werfel J. 2001. *Neural network models for zebra finch song production and reinforcement learning*. Master's thesis, MIT.
- Wilden I., Herzel H., Peters G. & Tembrock G. 1998. Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics* 9: 171-196.
- Williams C.E. & Stevens K.N. 1972. Emotions and speech: some acoustical correlates. *J. Acoust. Soc. Am.*, 52: 1238-1250.
- Wilson W.L. & Wilson C.C. 1975. Species specific vocalizations and the determination of phylogenetic affinities of the *Presbytis-aygula-melalophos* group in Sumatra. In: *Contemporary Primatology*. Proc. 5th Intl. Congr. Primatol., pp. 459-463, Nagoya.
- Zemlin W.R. 1997. *Speech and Hearing Science: Anatomy and Physiology*. 4th ed. Allyn and Bacon, Boston.
- Zimmermann E. 1990. Differentiation of vocalizations on bushbabies (Galaginae, Prosimiae, Primates) and significance for assessing phylogenetic relationships. *Z. Zool. Syst. Evol.*, 28: 217-239
- Zimmermann E. 1995a. Acoustic communication in nocturnal prosimians. In: Alterman L., Doyle G.A. & Izard K., *Creatures of the Dark: Biology of Nocturnal Prosimians*, pp 311-330, Plenum Press, New York.
- Zimmermann E. 1995b. Loud calls in nocturnal prosimians: Structure, evolution and ontogeny. In: Zimmermann E., Newman J.D. & Juergens, U., *Current Topics in Primate Vocal Communication*, pp. 47-72, Plenum Press, New York.
- Zimmermann E., Bearder S.K., Doyle G.A. & Andersson A. 1988. Variations in vocal patterns of Senegal and South African lesser bushbabies and their implications for taxonomic relationships. *Folia Primatol.*, 51: 87-105.

