Taylor & Francis
Taylor & Francis Group

Nasal and oral calls in mother and young trunk-nosed saiga antelopes, *Saiga tatarica*

Ilya A. Volodin[a,b]*, Olga V. Sibiryakova[a], Lyudmila E. Kokshunova[c], Roland Frey[d] and Elena V. Volodina[b]

[a]Department of Vertebrate Zoology, Faculty of Biology, Lomonosov Moscow State University, Vorobievy Gory, 1/12, Moscow 119991, Russia; [b]Scientific Research Department, Moscow Zoo, B. Gruzinskaya, 1, Moscow 123242, Russia; [c]Department of Pedagogical Education and Biology, Kalmyk State University, Pushkin Street, 11, Elista 358000, Russia; [d]Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Str., 17, 10315 Berlin, Germany

The trunk-like nose of the saiga antelope *Saiga tatarica* is a striking example of an exaggerated trait, assumed to having evolved as a dust filter for inhaled air. In addition, it functions to elongate the vocal tract in harem saiga males for producing low-formant calls that serve as a cue to body size for conspecifics. This study applies the source–filter theory to the acoustics of nasal, oral and nasal-and-oral calls that were recorded from a captive herd of 24 mother and 32 neonate saigas within the first 10 days postpartum. Anatomical measurements of the nasal and oral vocal tracts of two specimens (one per age class) helped to establish the settings for the analysis of formants. In both mother and young, the lower formants of nasal calls/call parts were in agreement with the source–filter theory, which suggests lower formants for the longer nasal vocal tract than for the shorter oral vocal tract. Similar fundamental frequencies of the nasal and oral parts of nasal-and-oral calls were also in agreement with the source–filter theory, which postulates the independence of source and filter. However, the fundamental frequency was higher in oral than in nasal calls, probably due to the higher emotional arousal during the production of oral calls. We discuss production mechanisms and the ontogeny of formant patterns of oral and nasal calls among bovid and cervid species with and without a trunk-like nose.

Keywords: source–filter theory; nasal and oral calls; acoustic analysis; vocal anatomy; mammals; mother–offspring communication

Introduction

Proboscis noses evolved several times independently among mammals and, accordingly, differ considerably in terms of anatomical structure and functions: e.g. in elephants (Gaeth et al. 1999; Todd 2010), tapirs (Witmer et al. 1999), coatis (Gompper 1995; Gompper and Decker 1998), elephant shrews (Kratzing and Woodall 1988), elephant seals (Huntley et al. 1984; Sanvito et al. 2007), proboscis monkeys, dikdik antelopes (Kamau et al. 1984; Frey and Hofmann 1996, 1997), pigs and some insectivores. Although vocal function of proboscis noses is secondary relative to their main functions, vocal emission entirely or partially through the nose is well documented for African elephants *Loxodonta africana* (McComb et al. 2003; Stoeger-Horwath et al. 2007; Stoeger, Heilmann, et al. 2012), Asian elephants *Elephas maximus* (Stoeger, Mietchen, et al. 2012) and saigas *Saiga tatarica* (Frey et al. 2007; Volodin et al. 2009). Vocalizations through the nose were also reported

---

*Corresponding author. Email: volodinsvoc@gmail.com

for proboscis monkeys *Nasalis larvatus* (Kawabe and Mano 1972). Rutting male Southern elephant seals *Mirounga leonina* emit their roars through the widely opened mouth; however, the large proboscis-like nose becomes inflated during each roar and might function as a resonator for increasing the amplitude of the roars (Sanvito and Galimberti 2000; Sanvito et al. 2007).

The source–filter theory (Fant 1960; Taylor and Reby 2010) provides a framework for explaining how nasal and oral vocal production modes affect the acoustic structure of a call. According to this theory, the fundamental frequency (f_0) of mammalian vocalizations is generated by vibrations of the vocal folds in the larynx (the source). Subsequently, while passing through the supralaryngeal vocal tract towards the mouth, it is subjected to an acoustic filtering process revealing vocal tract resonances (formants). Formant frequencies are inversely related to vocal tract length (Fitch and Reby 2001; Taylor and Reby 2010). Nasal vocal tracts are always longer than oral vocal tracts in the same individual and, therefore, formants of nasal calls should always be lower than formants of oral calls. This has been confirmed by studies of goitred gazelles *Gazella subgutturosa*, a species with a typical, i.e. not trunk-like, ruminant nose (Efremova et al. 2011; Volodin et al. 2011), African elephants (Stoeger, Heilmann, et al. 2012) and a single juvenile saiga (Volodin et al. 2009).

The source–filter theory suggests the independence of source and filter; i.e. vocal tract filtering should not affect the f_0 of the sound created in the larynx (Fant 1960; Titze 1994; Taylor and Reby 2010). Recent research showed that the source–filter theory is applicable to many sophisticated cases of mammalian vocal production modes, e.g. to vocalizing with a mobile larynx (e.g. Reby and McComb 2003; McElligott et al. 2006; Frey et al. 2011, 2012). Occasional departures from this theory due to source–filter coupling were suggested for certain rarely produced calls of Iberian red deer *Cervus elaphus hispanicus* (Volodin et al. 2013) and for Rocky Mountain elk *Cervus elaphus nelsoni* (Titze and Riede 2010). A non-independent (coupled) source and filter can be proposed when the vocal folds start oscillating at one of the formant frequencies (Titze 2008). Further potential departures concern the higher f_0 in oral than in nasal calls in domestic ewes *Ovis aries* (Sebe et al. 2010), goitred gazelles (Volodin et al. 2011), African elephants (Stoeger, Heilmann, et al. 2012) and in a single juvenile saiga (Volodin et al. 2009).

The saiga has a prominent trunk-like nose that, apart from other functions, is used for vocalizing (Frey et al. 2007; Volodin et al. 2009). A specific feature of the saiga biology is the seasonal covering of large distances while constantly moving in large herds (Bannikov et al. 1961; Heptner et al. 1989; Danilkin 2005). The most frequently used locomotory mode of saigas is the amble during which their lowered heads are almost permanently immersed in clouds of dust raised by the hooves of the innumerous animals of a herd (Frey et al. 2007). The trunk-like nose of saigas has probably evolved for filtering dust from the inhaled air (Bannikov et al. 1961; Heptner et al. 1989) and is present in all sex- and age-classes (Frey et al. 2007; Volodin et al. 2009). The young follow their mothers in herds from the first days of their life (Danilkin 2005) and, therefore, natural selection for large noses is acting on both sexes and all age classes (Frey et al. 2007).

At maturation, however, the saiga nose becomes disproportionally larger in males, a size difference likely due to sexual selection (Frey et al. 2007). Harem saiga males use their trunk-like noses for the emission of rutting roars while adopting a specific vocal posture. This posture, in combination with strong tension and rostral extension (by about 20%) of the trunk-like nose, causes a decrease of the formant frequencies of the roars (Frey et al. 2007). This represents a particular case of nasal vocal tract elongation. In mammals, two main ways of vocal tract elongation are possible (Fitch and Hauser 2002). The

elongation of the oral vocal tract is achieved *via* a descent and/or retraction of the larynx associated with an oral call production mode as, e.g. in male red deer, fallow deer *Dama dama*, goitred gazelles *G. subgutturosa* and Mongolian gazelles *Procapra gutturosa* (Reby and McComb 2003; McElligott et al. 2006; Frey et al. 2008, 2011, 2012). The elongation of the nasal vocal tract is achieved by evolving an extensible, trunk-like nose, eventually combined with nose extension, and a nasal call production mode as in male saiga (Frey et al. 2007). Elongated vocal tracts promote low-formant calls, which serve to advertise male dominance and attract receptive females (Reby et al. 2005; Charlton et al. 2007). Lip protrusion can provide additional vocal tract elongation at the rostral end, as in the *coo* calls of rhesus monkeys *Macaca mulatta* (Hauser and Schön Ybarra 1994).

So far, only nasal calls have been described for adult female saigas, whereas both nasal and oral calls have been reported for a single 1.5-month-old saiga (Volodin et al. 2009). These calls function as contact calls during interactions between mothers and young. The acoustic structures of these calls suggested that they were of the same "pulsed" call type as the rutting roars of adult saiga males (Volodin et al. 2009).

In the last 15 years, the number of saigas in the Russian Federation has undergone a dramatic tenfold decrease. Currently the saiga population does not exceed 10,000 individuals in Russia (Kokshunova 2012) and has been included in the International Union for Conservation of Nature (IUCN) Red List of threatened species (Mallon 2008). Along with protection in the wild, the creation of saiga breeding centres is an important issue for conservation of this species. In addition, such breeding centres provide an opportunity to study the vocal behaviour of saiga mothers and young. Taking acoustic recordings of saigas in their natural habitat is extremely difficult, as this species lives in open steppe and semi-desert areas and their flight distance is large.

Our purpose was to analyse the effects of nasalization on vocalization within the perspective of the source–filter theory for a mammal with a trunk-like nose. We compare fundamental frequency, duration and formant frequency measures between nasal and oral call types emitted by adult female saigas and their newborn young. In addition, we compare acoustic variables of nasal and oral call parts of the same calls.

Materials and methods

Study site and housing

Calls of adult mothers and young were recorded in May 2002 in the "Centre for Wild Animals of Kalmykia" (Russia, Republic of Kalmykia, settlement Har-Buluk, 46°19′01″N, 44°01′26″E). This State breeding centre was created in 2000 for the captive breeding of saigas as a conservation measure (Kokshunova 2012). Females and their respective offspring were kept together in an open-door 48 m × 34 m enclosure, equipped with permanent water bowls and feeders. The enclosure contained canvas covers providing shade and was surrounded by a solid fence, providing additional shade for the animals and protection from harsh winds. The subject adult females were fed with grain, mineral supplements and pre-dried grass. Hay and water were available *ad libitum* (Kokshunova 2012). The young were nursed by their mothers but also tasted pre-dried grass from their fifth day of life.

Subjects

Our subjects were 24 adult females and 32 young. Fifteen of the 24 females were wild-born in 2000 and then human-raised from the age of 1 to 2 days. At the start of this study, they were 2 years old and already had breeding experience. The other nine females were

captive-born in 2001 and were nursed by their mothers. During this study, they were 1 year old and bred for the first time. The neonate saigas were sired by two adult males, normally kept in separate enclosures. The young were born from 20 of the 24 females between 30 April and 9 May 2002. The young saigas were 0–10 days of age during the recordings. Twenty-four of the 32 young were twin-siblings and the remaining eight young were singletons. Nineteen of the 32 young were born before the start of data collection on 2 May 2002, and the other 13 young were born during the period of data collection.

Acoustic recordings

Data were collected 10 years before this study from untagged animals as a by-product of another research project not focusing on vocal behaviour. Accordingly, all the acoustic recordings were made without identification of individuals and sex of the young. Audio and video recordings were made using a Sony CCD-TR640 Hi8 camcorder (Sony Corp., Tokyo, Japan). This camcorder uses an analogue audio format. The frequency response of the microphone was from 20 Hz to 20 kHz.

A total of 144 min of simultaneous audio and video recordings were collected on 2, 3, 4 and 9 May 2002. The recording sessions were scheduled between 6:30 and 17:00 during periods when the animals were active. All recordings were made by one researcher (LK) from an observation tower 3.5 m high, located inside the enclosure. The camcorder served as a device for both video and audio recordings. As a consequence, some calling animals were not in the focus of the camcorder during the recording sessions. The distance from the camcorder to the calling animals varied between 2 and 20 m. The researcher commented on the camcorder by voice as to whether the calls were produced by mothers or by young.

The animals emitted calls through the nose and through the mouth when interacting with each other, e.g. when anticipating food and during routine procedures of feeding and cleaning of the enclosure. The young called when being separated from their mothers by a distance of more than 5 m or by other animals and when searching for their mother among other adult females. Females vocalized in response to calls of either their own young or in response to calls of alien young or when disturbed by other females.

Call samples

Sound tracks of video recordings were digitized at 48 kHz and 16-bit precision using Avisoft-SASLab Pro (Avisoft Bioacoustics, Berlin, Germany). All recorded calls were of the "pulsed" type, most commonly produced by all sex- and age-classes of saigas (Frey et al. 2007; Volodin et al. 2009). For acoustic analyses, we selected only calls of good quality, i.e. not overlapped by calls of other animals, recorded with good signal-to-noise ratio and not impaired by wind noise. For all calls satisfying these criteria, start and end points were labelled in a sound file.

Calls were subdivided into categories: originating from mothers or young, based on the video clips and according to substantial differences between the f_0 ranges of adult females and young that had been established in a previous study (Volodin et al. 2009). Mother and young calls were further subdivided by their production mode, i.e. either through the nose or through the mouth (mainly following Efremova et al. 2011; Volodin et al. 2011), into nasal, oral and nasal-and-oral calls (Figure 1). Nasal calls were produced through the nose, while the mouth was kept closed. Oral calls often had a very short nasal onset, but then were produced by emission of the sound entirely through the opened mouth. Nasal-and-oral calls started nasally and ended orally when an animal opened its mouth during a call (Figure 1).
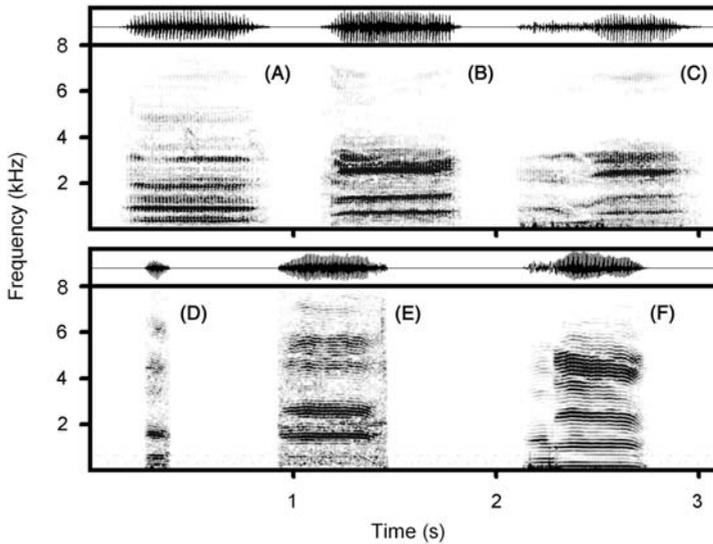
Figure 1.   Spectrogram and waveform showing examples of saiga antelope adult female (upper panel) and young (bottom panel) calls: nasal (A, D), oral (B, E) and nasal-and-oral (C, F). Each nasal-and-oral call comprises a nasal part at the start and an oral part at the end. The spectrogram was created with Hamming window, 24 kHz sampling rate, FFT 1024 points, frame 50% and overlap 93.75%. Both female and young calls are tonal, but at the same spectrogram settings the calls show either a pulsed (mother) or a harmonic (young) representation due to the lower fundamental frequency (f_0) of the female. See Supplementary Data for the sound files, which were used for creating these spectrograms.

We classified calls to nasal, oral and nasal-and-oral based on 82 calls from video clips, in which we could see whether the mouth was open or closed during calling, based on corresponding spectrograms of these calls, in which we could visually estimate differences in energy distribution (Figure 1), and based on specifically "nasal" sound quality in calls produced through the nose. Using this classification, two researchers (IV and OS) classified all calls of the acoustic files by spectrograms and by hearing. For analysis, we took only calls, for which both researchers coincided in their estimates of call types.

We analysed 158 nasal, 95 oral and 62 nasal-and-oral calls of mothers and 11 nasal, 168 oral and 19 nasal-and-oral calls of young, 513 calls in total. These calls were taken from about 20 mothers and 30 neonates, because many different animals vocalized during the recording period. Measurements for the nasal and oral parts of the nasal-and-oral calls were taken separately. If taking measurements from both nasal and oral parts of a nasal-and-oral call was not possible, we excluded this call from the analysis. As acoustic variables of nasal and oral parts of the same nasal-and-oral call could not be included into statistical analyses as independent variants, we randomly took either the nasal or the oral part of each nasal-and-oral call for the statistic analyses.

Measuring nasal and oral vocal tract lengths

We investigated two specimens anatomically. The first specimen was a deep frozen-thawed adult female (of 1 year and 3 months of age). It had already bred and died of disease in Moscow Zoo in 2007. This specimen was obtained as a carcass without internal organs and weighed 14.5 kg. The second specimen was a neonate male that had lived only for a few days and died of disease in Cologne Zoo in 2002. As the post-mortem position of

the larynx approximately corresponds to its resting position in the live animal, the resting lengths of the oral and nasal vocal tract could be estimated. The oral tract was measured from the vocal folds to the edges of the lips along the pharyngeal and oral cavities, and the nasal vocal tract was measured from the vocal folds to the nostrils along the pharyngeal and nasal cavities (Figures 2 and 3).

Measurements in the adult female specimen were taken during macroscopic anatomical dissection of the head-and-neck region. After removal of the left half of the mandible including the attaching musculature (masseter, pterygoid, temporalis and buccal muscles), the oral and nasal vocal tract lengths were ascertained with a string measure and subsequently scaled with a tape measure (1 mm precision). Each measurement was
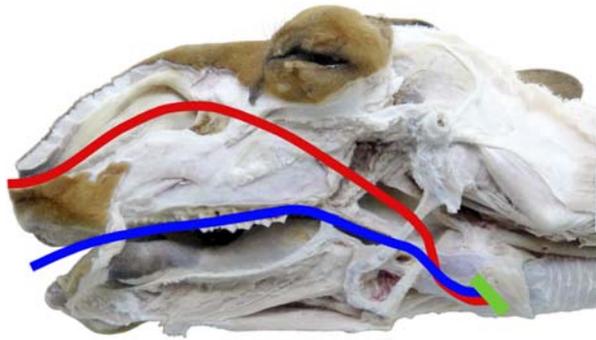
Figure 2.   (Colour online) Dissection image representing lengths of oral and nasal vocal tract in the dissected adult female saiga specimen. Vocal folds are marked in green, the nasal vocal tract length in red and the oral vocal tract length in blue. Scale bar is 10 mm.
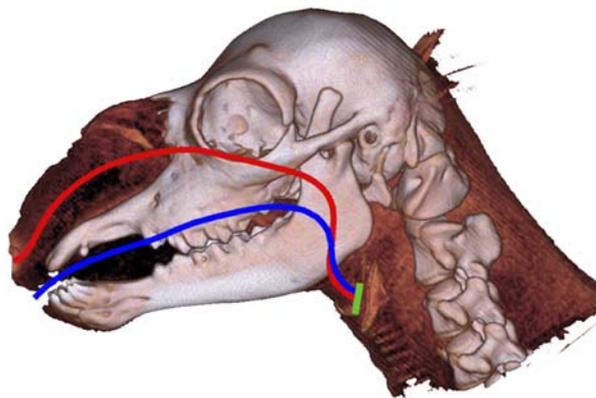
Figure 3.   (Colour online) CT image representing lengths of the oral and nasal vocal tract in a neonate male saiga specimen. Vocal folds are marked in green, the nasal vocal tract length in red and the oral vocal tract length in blue. Scale bar is 10 mm.

repeated three times, and the mean value was calculated. The photo documentation of the dissections was done with a Panasonic Lumix DMC-FZ150 (Panasonic Corp., Osaka, Japan) digital camera, a Panasonic DMC-LA 5E adapter tube and a Panasonic DMC-LC55E close-up lens. Images were stored on secure digital high capacity memory card, transferred to a PC and processed with Adobe Photoshop CS4 (Adobe Systems, Inc., San Jose, CA, USA).

The neonate head-and-neck specimen was subjected to computer tomographic (CT) investigation in a General Electric Lightspeed 4-slice helical CT (General Electric, Milwaukee, WI, USA). Scans were set at 120.0 kV, 120.0 mA, 0.6 mm slice thickness. A volume rendering software (GE VolumeViewer) was applied for processing of the slice data. High-resolution 3D reconstructions and virtual sagittal sections allowed for a detailed analysis of the vocal organs and associated structures. Colour markings of vocal fold position and nasal and oral vocal tract lengths were set in a virtual sagittal section and an unsectioned virtual 3D reconstruction. Measurements in the neonate specimen were then taken directly from the screen by a string measure. Scaling was done by means of the scale measure from the original scan data contained in the images (1 mm precision).

Call analysis

The generally agreed basic model for the analysis of formants is that of a uniform tube closed at one end, considering the sound source (larynx and vocal folds) as the closed end, whilst the mouth or nostrils represent the open end (Fitch 1997; Fitch and Reby 2001). According to this model, expected formant frequencies can be calculated as

$$\mathrm{Fn} = \frac{(2n - 1) \times c}{4L},$$

where n are formant numbers (1, 2, 3, etc.), L is vocal tract length and c is the speed of sound in air, approximated as $350 \, \mathrm{m \, s}^{-1}$. The age of our subject animals over the study period corresponded to that of the dissected specimens. Therefore, the anatomically ascertained mean values of vocal tract lengths served to establish the settings (maximum number of formants and the upper limit of frequency range) for linear prediction coding (LPC) and further analysis of the formant frequencies of the nasal and oral calls with the Praat DSP package v. 4.3.21 (Boersma and Weenink 2009).

From each nasal and oral call and also from the nasal and oral parts of each nasal-and-oral call, we measured the same six acoustic variables: duration, fundamental frequency period and four formant frequencies (Figure 4). Prior to analysis, calls were downsampled to 24 kHz. We measured call duration with Avisoft SASLab Pro from the screen with the standard marker cursor in the main window of Avisoft. The mean f_0 period (i.e. the mean distance from the previous pulse to the following pulse) was measured from the screen with the standard marker cursor in the main window of Avisoft, displaying the spectrogram and the waveform, with the following settings: Hamming window, Fast Fourier Transform (FFT) length 512, frame 50%. Frequency resolution of the spectrographic analysis was 43 Hz, time resolution varied between 0.3 and 0.5 ms, depending on call duration. All measurements were exported automatically to Microsoft Excel (Microsoft Corp., Redmond, WA, USA). Then we calculated the mean f_0 of each call as the inversed value of the mean f_0 period of the call (Figure 4).

The four first formants (F1, F2, F3 and F4) were measured using LPC with Praat. The LPC settings for creating the formant tracks were Burg analysis, window length 0.04 s,
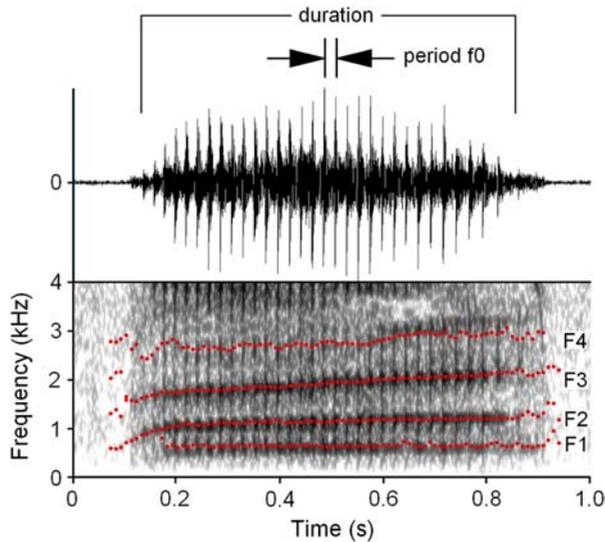
Figure 4.   (Colour online) Measured acoustic variables from waveform (above) and spectrogram (below) of the oral call of an adult female saiga: duration, fundamental frequency period (period f_0) and tracks of the first four formants (F1–F4). The LPC settings were Burg analysis, window length 0.04 s, time step 0.01 s, maximum number of formants 4, maximum formant frequency 3200 Hz.

time step 0.01 s; maximum number of formants 4–5. For mothers, the maximum formant frequency (the upper limit of frequency range) was 2000–2500 Hz for nasal calls and the nasal parts of nasal-and-oral calls, and 3000–3700 Hz for oral calls and the oral parts of nasal-and-oral calls. For the young, the maximum formant frequency was 4900–5300 Hz for nasal calls and the nasal parts of nasal-and-oral calls, and 5800–6700 Hz for oral calls and the oral parts of nasal-and-oral calls (Figure 4). Point values of formant tracks were extracted, exported to Excel and the value of each formant for the given call or call part was calculated as the average value from the point values. Applying the model of a uniform tube closed at one end, we calculated the formant dispersion (dF) for nasal and oral calls of mother and young saigas by using linear regression according to Reby and McComb (2003).

Statistics

All statistical analyses were carried out with STATISTICA, v. 6.0 (StatSoft, Inc., Tulsa, OK, USA). Means are given as mean ± SD. All statistical tests were two tailed and differences were considered significant where $P < 0.05$. Only 7 of 48 distributions of values of acoustic variables weakly departed from normality ($0.01 < P < 0.05$, Kolmogorov–Smirnov test). As parametrical ANOVA are relatively robust to departures from normality (Dillon and Goldstein 1984), this was not an obstacle to the application of these tests. Acoustic parameters of the nasal and oral parts of nasal-and-oral calls were included into the statistical analyses as independent variants by randomly selecting either the nasal or the oral part of each nasal-and-oral call. We used one-way ANOVA with Tukey post hoc test to compare the values of acoustic variables among nasal calls, oral calls, nasal parts of nasal-and-oral calls and oral parts of nasal-and-oral calls for mothers and young, respectively. We used repeated measures ANOVA

with Tukey post hoc test to compare the values of formant frequencies within nasal and oral calls.

Ethics

As the video recordings did not form part of an experiment and as routine management of the saigas followed welfare guidelines, no ethical permission was required by the Centre for Wild Animals of Kalmykia. During our work, we adhered to the "Guidelines for the treatment of animals in behavioural research and teaching" (Animal Behaviour, 2006, 71, 245–253) and to the laws on animal welfare for scientific research of the Russian Federation and Germany, where the study was conducted. The study was approved by the Committee of Bio-ethics of Lomonosov Moscow State University (protocol # 2010-36).

Results

The nasal vocal tract length was 320 mm in the female and 142 mm in the neonate specimen. The oral vocal tract length was 240 mm in the female and 116 mm in the neonate specimen.

Mothers and young do not raise the head during calling and only slightly tense the nose when producing nasal calls. Females were silent during parturition, but started calling within 1 h postpartum. Mothers and young called before nursing or when they were separated from each other by distance or by other animals. During and after nursing the females were silent. Although we could not count the total number of nasal and oral calls produced by mothers and young, the used call samples appeared to be a good representation of the percentages of nasal and oral calls. They suggest that the mothers produced more nasal than oral or nasal-and-oral calls, whereas the young produced predominantly oral calls and very rarely nasal or nasal-and-oral calls.

Mother calls

Values of the first four formants did not differ significantly between nasal calls and the nasal parts of nasal-and-oral calls, and between oral calls and the oral parts of nasal-and-oral calls (Table 1). However, all the four formants of nasal calls/call parts were significantly lower than respective formants of oral calls/call parts. At the same time, the formant distances F2–F1 and F3–F2 were significantly shorter in nasal calls/call parts than in oral calls/call parts, whereas the distance F4–F3 was significantly shorter only in nasal than in oral calls, but not between nasal and oral call parts (Table 1).

Distances between neighbouring formants differed also within nasal ($F_{2,314} = 47.05$, $P < 0.001$) and within oral calls ($F_{2,188} = 37.29$, $P < 0.001$). Within nasal calls, F3–F2 and F4–F3 did not differ significantly, but both were larger than F2–F1 ($P < 0.001$ in both cases, Tukey post hoc test). In contrast, within oral calls, F3–F2 was significantly larger than either F2–F1 or F4–F3 ($P < 0.001$ in both cases), and F2–F1 was significantly larger than F3–F4 ($P < 0.001$, Tukey post hoc test). Thus, distances between neighbouring formants were uneven in both nasal and oral calls, and the formant patterns of nasal calls differed strongly from those of oral calls (Table 1).

The estimated formant dispersion of 576 Hz corresponds to an estimated nasal vocal tract length of 304 mm during nasal calls (Figure 5). The estimated formant dispersion of 879 Hz corresponds to an estimated oral vocal tract length of 199 mm during oral calls. Thus, the vocal tract lengths, both nasal and oral, estimated by formant dispersion using

Table 1. Comparison between oral and nasal calls/call parts of adult female saigas: mean ± SD values of acoustic variables and ANOVA results.

Call/call part variable	Nasal calls, $n = 158$	Nasal parts of nasal-and-oral calls, $n = 31$	Oral parts of nasal-and-oral calls, $n = 31$	Oral calls, $n = 95$	ANOVA results
F1 (Hz)	435 ± 62[a]	406 ± 45[a]	646 ± 81[b]	666 ± 73[b]	$F_{3,311} = 305.64$, $P < 0.001$
F2 (Hz)	891 ± 52[a]	914 ± 86[a]	1386 ± 94[b]	1413 ± 111[b]	$F_{3,311} = 992.97$, $P < 0.001$
F3 (Hz)	1431 ± 103[a]	1423 ± 106[a]	2309 ± 149[b]	2297 ± 199[b]	$F_{3,311} = 932.43$, $P < 0.001$
F4 (Hz)	1992 ± 105[a]	2049 ± 85[a]	2886 ± 166[b]	2935 ± 119[b]	$F_{3,311} = 1622.02$, $P < 0.001$
F2–F1 (Hz)	456 ± 83[a]	508 ± 98[a]	740 ± 105[b]	748 ± 133[b]	$F_{3,311} = 186.03$, $P < 0.001$
F3–F2 (Hz)	540 ± 107[a]	509 ± 94[a]	923 ± 124[b]	884 ± 141[b]	$F_{3,311} = 231.10$, $P < 0.001$
F4–F3 (Hz)	561 ± 83[a]	627 ± 93[a,b]	577 ± 162[a,b]	638 ± 212[b]	$F_{3,311} = 6.51$, $P < 0.001$
Duration (s)	0.45 ± 0.16[a]	0.37 ± 0.12[a]	0.46 ± 0.21[a]	0.55 ± 0.17[b]	$F_{3,311} = 12.27$, $P < 0.001$
f_0 (Hz)	54.7 ± 6.4[a]	64.5 ± 4.8[b]	64.5 ± 7.7[b]	65.8 ± 10.3[b]	$F_{3,311} = 49.09$, $P < 0.001$

Notes: f_0, fundamental frequency; duration, call duration; F1, F2, F3, F4, frequencies of the first four formants; F2–F1, F3–F2, F4–F3, distances between neighbouring formants. The same superscripts indicate which calls/call parts did not differ significantly ($P > 0.05$, Tukey post hoc test).

the linear regression method (Reby and McComb 2003) were shorter than those measured directly during the anatomical dissection. For the nasal vocal tract, the difference was 5%, and for the oral vocal tract it was 17%.

Oral calls were significantly longer than nasal calls, whereas nasal and oral call parts did not differ in duration (Table 1). The entire duration of nasal-and-oral calls (0.77 ± 0.15 s) was significantly longer than the duration of either nasal or oral calls ($F_{2,312} = 85.58$, $P < 0.001$, Tukey post hoc test, $P < 0.001$ in both cases).

The very low fundamental frequency is visible as a pulsation in narrowband spectrograms of both nasal and oral calls (Figure 1; Table 1). The f_0 was significantly lower in nasal calls (54.7 Hz) than in oral calls (65.8 Hz) and lower in nasal calls than in nasal or oral call parts (64.5 Hz in both cases; Figure 6; Table 1). The f_0 did not differ between oral calls and nasal and oral call parts.

Young calls

As in the mothers, the values of the first four formants did not differ significantly between nasal calls and nasal call parts in the young (Table 2). Correspondingly, the values of the first four formants did not differ significantly between oral calls and oral call parts. The values of all the four first formants of nasal calls/call parts were significantly lower than the values of the corresponding formants of oral calls/call parts. Unlike in mothers, only the F3–F2 distance was significantly shorter in nasal calls/call parts than in oral calls/call parts. The F2–F1 distances did not differ between nasal and oral calls/call parts, whereas
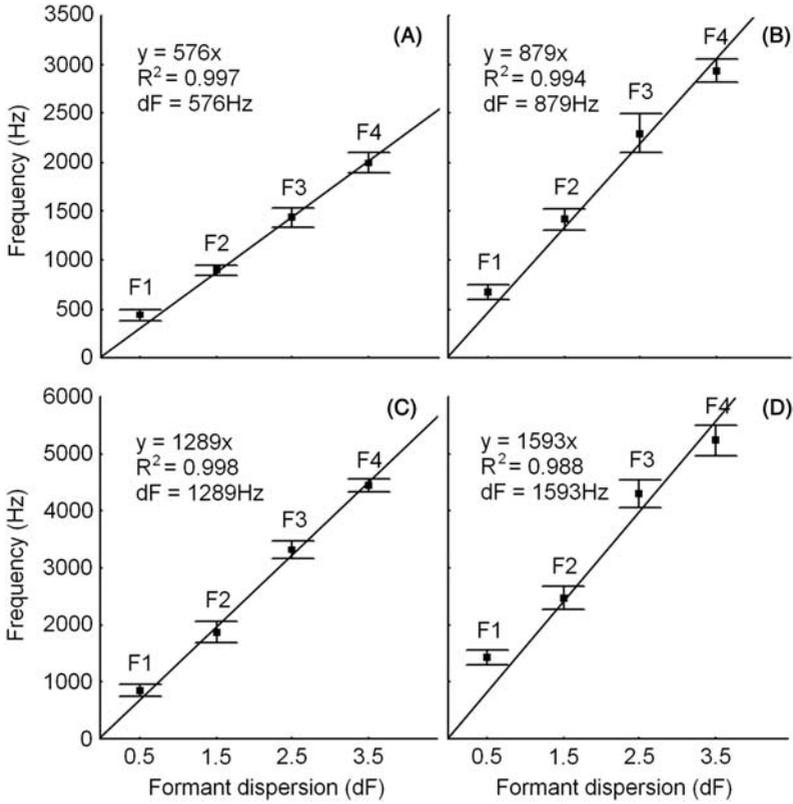
Figure 5.   Estimation of formant dispersion (dF) in the mother nasal calls (A), mother oral calls (B), young nasal calls (C) and young oral calls (D) of saiga by using linear regression. Central points show the means of the first four formants (F1–F4), whiskers show the SD.
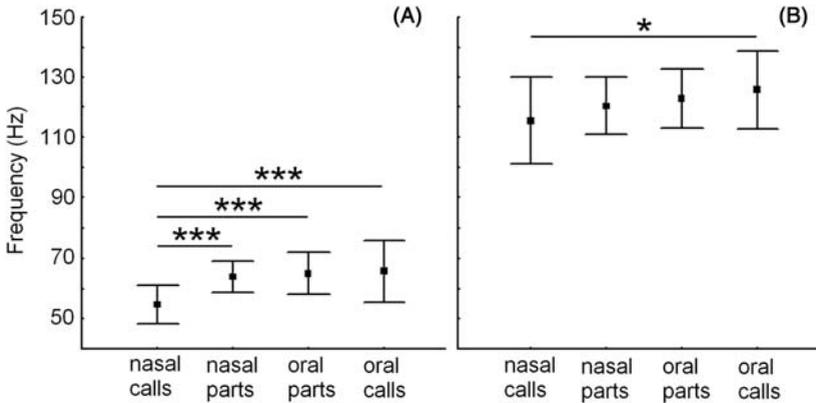

Figure 6.   Fundamental frequency (f_0) of saiga mother (A) and young (B) nasal calls, nasal parts of nasal-and-oral calls, oral parts of nasal-and-oral calls and oral calls. Central points show the means, whiskers show the SD. $*P < 0.05$; $***P < 0.001$, Tukey post hoc test.

Table 2. Comparison between nasal and oral calls/call parts of neonate saigas: mean ± SD values of acoustic variables and ANOVA results.

Call/call part variable	Nasal calls, $n = 11$	Nasal parts of nasal-and-oral calls, $n = 10$	Oral parts of nasal-and-oral calls, $n = 9$	Oral calls, $n = 168$	ANOVA results
F1 (Hz)	849 ± 108[a]	834 ± 108[a]	1383 ± 167[b]	1427 ± 127[b]	$F_{3,194} = 131.38$, $P < 0.001$
F2 (Hz)	1870 ± 184[a]	1922 ± 195[a]	2421 ± 154[b]	2465 ± 209[b]	$F_{3,194} = 48.08$, $P < 0.001$
F3 (Hz)	3316 ± 157[a]	3159 ± 271[a]	4190 ± 291[b]	4300 ± 243[b]	$F_{3,194} = 118.81$, $P < 0.001$
F4 (Hz)	4443 ± 115[a]	4393 ± 221[a]	5231 ± 281[b]	5227 ± 266[b]	$F_{3,194} = 60.94$, $P < 0.001$
F2–F1 (Hz)	1021 ± 189[a]	1088 ± 135[a]	1038 ± 190[a]	1037 ± 184[a]	$F_{3,194} = 0.28$, $P = 0.84$
F3–F2 (Hz)	1447 ± 218[a]	1237 ± 251[a]	1769 ± 260[b]	1835 ± 245[b]	$F_{3,194} = 26.04$, $P < 0.001$
F4–F3 (Hz)	1126 ± 201[a,b]	1234 ± 144[a]	1041 ± 391[a,b]	928 ± 263[b]	$F_{3,194} = 6.31$, $P < 0.001$
Duration (s)	0.37 ± 0.14[a]	0.27 ± 0.12[a]	0.34 ± 0.10[a]	0.50 ± 0.17[b]	$F_{3,194} = 10.16$, $P < 0.001$
f_0 (Hz)	115.6 ± 14.5[a]	120.5 ± 10.4[a,b]	120.9 ± 9.9[a,b]	125.8 ± 12.8[b]	$F_{3,194} = 2.91$, $P < 0.05$

Notes: f_0, fundamental frequency; duration, call duration; F1, F2, F3, F4, values of the first four formants; F2–F1, F3–F2, F4–F3, distances between neighbouring formants. The same superscripts indicate which calls/call parts did not differ significantly ($P > 0.05$, Tukey post hoc test).

the F4–F3 distances differed only between oral calls and the nasal parts of nasal-and-oral calls (Table 2).

As in the mothers, the distances between neighbouring formants differed within nasal ($F_{2,20} = 8.86$, $P < 0.01$) and within oral calls ($F_{2,334} = 588.50$, $P < 0.001$) of the young. Within nasal calls, F2–F1 and F4–F3 did not differ significantly, and both were smaller than F3–F2 ($P < 0.05$ in both cases, Tukey post hoc test). Within oral calls, F3–F2 was significantly larger than either F2–F1 or F4–F3 ($P < 0.001$ in both cases), and F2–F1 was significantly larger than F3–F4 ($P < 0.001$, Tukey post hoc test). Thus, unlike the mothers, the young showed matching formant patterns for oral and nasal calls (Table 2).

The estimated formant dispersion of 1289 Hz corresponds to an estimated nasal vocal tract length of 136 mm during young nasal calls (Figure 5). The estimated formant dispersion of 1593 Hz corresponds to an estimated oral vocal tract length of 110 mm during young oral calls. As for the mothers, both nasal and oral vocal tract lengths, estimated by the linear regression method (Reby and McComb 2003), were shorter than those measured directly in the dissected specimen. However, the differences did not exceed 5%.

As in the mothers, oral calls were significantly longer than nasal calls, whereas nasal and oral call parts did not differ in duration (Table 2). However, the entire duration of nasal-and-oral calls (0.56 ± 0.11 s) was significantly longer than those of nasal calls ($F_{2,195} = 4.79$, $P < 0.01$, Tukey post hoc test, $P < 0.01$), but did not differ significantly from that of oral calls (Tukey post hoc test, $P = 0.32$).

The f_0 of the young was twice higher than that of the mothers, both in nasal and oral calls (Figures 1 and 6). The maximum f_0 values of the mothers were lower than the

minimum f_0 values of the young, both in nasal (minimum–maximum: 44.6–73.3 Hz vs. 91.6–142.9 Hz, respectively) and in oral calls (minimum–maximum: 42.5–80.3 Hz vs. 106.5–159.2 Hz, respectively). The f_0 values of the mother and the young calls did not overlap, so applying tests for significance was redundant. In the young, f_0 was significantly lower in nasal calls (115.6 Hz) than in oral calls (125.8 Hz; Figure 6; Table 2). The f_0 values of nasal and oral call parts (120.5 and 120.9 Hz, respectively) were intermediate between the f_0 of nasal and oral calls and did not differ from them significantly.

Discussion

Source–filter theory and formants of nasal and oral calls

Lower formants in nasal than in oral calls/call parts of mother and young saigas, found in this study, were in agreement with the source–filter theory, suggesting lower formants for the longer nasal vocal tract than for the shorter oral vocal tract (Fant 1960; Titze 1994; Taylor and Reby 2010). In the nasal-and-oral calls of both mother and young saigas, formants of nasal and oral parts were undistinguishable from respective formants of pure nasal and oral calls. Lower formants in nasal than in oral calls were already found in contact calls of juvenile goitred gazelles, a species without an enlarged nose (Volodin et al. 2011), and in rumbles of subadult African elephants (Stoeger, Heilmann, et al. 2012).

Formant-dispersion-based estimations of saiga vocal tract lengths agreed rather well with those ascertained in the dissected specimens. Thus, by using morphological data for the LPC settings prior to the use of the uniform-tube-closed-at-one-end model (Fitch 1997; Fitch and Reby 2001; Reby and McComb 2003), the positions of the formants in nasal and oral calls of mother and young saigas could be correctly predicted (see also Frey et al. 2007 for data in adult males). In a previous study, the lack of basic morphological data resulted in incorrect calculation of the position of the first formant, although formant dispersion was measured correctly in both mother and young saiga calls (Volodin et al. 2009).

Although the uniform tube vocal tract model is widely applied, real vocal tract shapes of mammals strongly depart from uniformity (e.g. Riede et al. 2005; Frey et al. 2007; Gamba et al. 2012; Frey and Riede 2013). In the saiga calls of this study, differences between neighbouring formants were unequal in both oral and nasal calls. This indicates that both nasal and oral vocal tracts were not uniform. Unequal distances between neighbouring formants were found also in adult male and young goitred gazelles (Frey et al. 2011; Volodin et al. 2011), red deer stags (Reby and McComb 2003; Kidjo et al. 2008; Frey et al. 2012), fallow deer bucks (McElligott et al. 2006) and adult male American bison *Bison bison* (Wyman et al. 2012).

Among the species of Bovidae, formant patterns may differ between species, sexes, age-classes and according to nasalization of calls (Figure 7). In oral calls of domestic goat kids *Capra hircus*, the distance between F3 and F4 formants is short, whereas the F1–F2 and F2–F3 distances are equal and wide (Briefer and McElligott 2011). In oral calls of saiga young, both F1–F2 and F3–F4 distances are short (this study, Figure 1), whereas in oral calls of goitred gazelle young, only the F2–F3 distance is short (Volodin et al. 2011). In nasal calls of both saigas and goitred gazelles, the F1–F2 and F3–F4 distances are short and the F2–F3 distance is wide (Volodin et al. 2011). Ontogenetically, this pattern is retained up to 6 months in goitred gazelles (Lapshina et al. 2012). In adult saigas the pattern has switched to closely spaced F1–F2 and evenly spaced other formants (females: this study, Figure 1; males: Frey et al. 2007; our unpublished data). This difference in nasal formant patterns between neonate and adult saigas most probably is the result of the
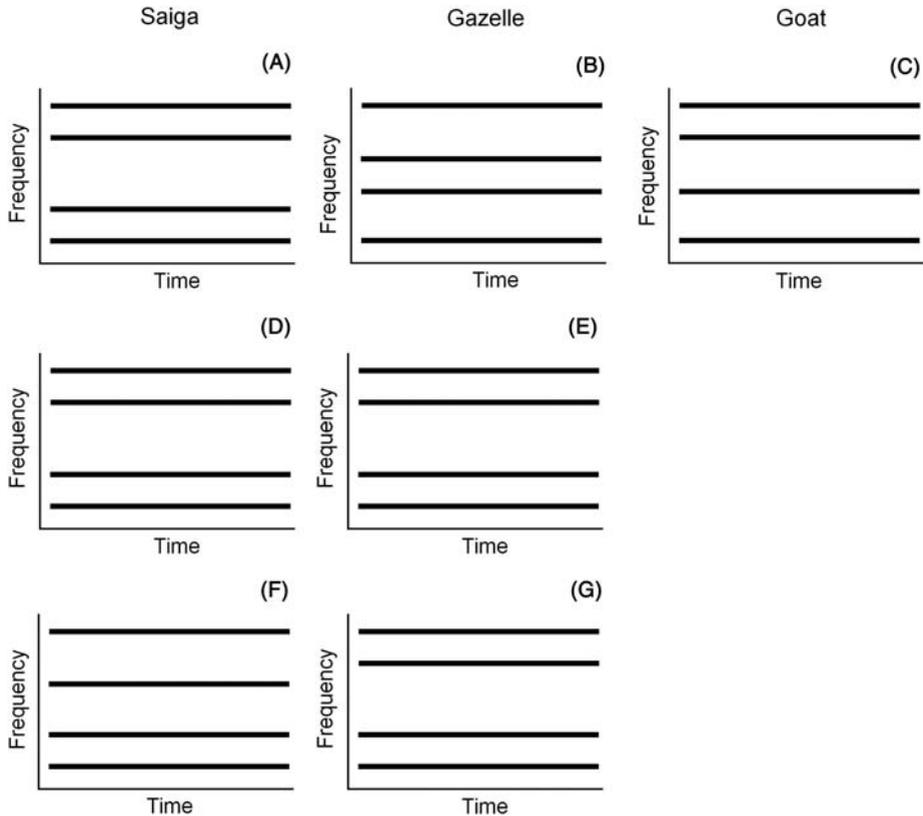
Saiga Gazelle Goat

Figure 7.   Schematic representation of the formant patterns (relative differences between the formants) in three species of Bovidae: saiga young oral calls (A); goitred gazelle young oral calls (B); domestic goat kid oral calls (C); saiga young nasal calls (D); goitred gazelle young nasal calls (E); saiga adult male and female nasal calls (F); goitred gazelle subadult male and female nasal calls (G). The bands represent the four first formants.

prominent differences in the size of the proboscis, which is much larger in adults than in neonates with a relatively small nose.

At the same time, the measured elevated values of the first formant compared with values predicted from the uniform tube model probably represent a common pattern of all species of Bovidae and Cervidae irrespective of age, sex or nasalization. The F1 values of saiga mother and young nasal and oral calls were higher than the values predicted by the uniform tube model (Figure 5). Consistently, nasal rutting roars of adult male saigas also had higher F1 values than expected from the model (Frey et al. 2007). An elevated F1 is also found in the oral rutting roars of red deer stags (Reby and McComb 2003; Frey et al. 2012), in the oral rutting roars and growls of adult male goitred gazelles (Frey et al. 2011) and in the oral rutting bellows of adult male American bison (Wyman et al. 2012). Therefore, this might turn out to be a common rule for the formant patterns of ruminants. Also, this phenomenon is probably not related to emotional arousal, as high-arousal rutting calls of adult male saigas are produced through the nose. The elevated F1 might also be influenced by extra air volumes connected to the vocal tract (e.g. small ventricle-like air sacs as in male Mongolian gazelle or the nasal cavities during oral calls). In experiments, such volumes significantly shifted the position of the first formant, whereas the upper formants remained weakly affected (Riede et al. 2008).

In addition to the length of the vocal tract, emotional arousal could be the factor, increasing the call formant frequencies (Briefer 2012). However, all comparisons of formant frequencies were made within call type; e.g. in the nasal calls of African elephants, the F1 increased with negative arousal (Soltis et al. 2011). So, the effects of the vocal tract length cannot be separated from the effects of emotional arousal on formant of saiga calls.

Source–filter theory and fundamental frequencies of nasal and oral calls

In mother and young saigas, the f_0 was lower in the nasal than in the oral calls, conforming to data on nasal and oral contact calls of goitred gazelle young (Volodin et al. 2011), nasal and oral bleats of adult ewes (Sebe et al. 2010) and nasal and oral rumbles of subadult African elephants (Stoeger, Heilmann, et al. 2012). This study additionally compared the f_0 within saiga nasal-and-oral calls that switch from a nasal to an oral production mode. We found that f_0 did not differ between the nasal and oral call parts neither in mothers nor in young. Therefore, the nasal and oral vocal tracts did not affect the f_0 of saiga calls. This is in perfect agreement with the source–filter theory, postulating the independence of source and filter (Fant 1960; Titze 1994; Taylor and Reby 2010).

The higher f_0 of oral calls than that of nasal calls could result from the higher emotional arousal of animals during the emission of oral calls. The increase of f_0 in correspondence with an elevated degree of emotional arousal represents a common phenomenon for all mammals (Briefer 2012) and particularly applies to ruminants (Charlton and Reby 2011; Lingle et al. 2012). Potentially, the f_0 increases, when an enhanced arousal triggers certain changes in the vocal folds, e.g. elongation of the vocal folds, increased vocal fold tension by an increase in tissue rigidity or a functional reduction of the oscillating tissue mass (Titze 1994; Riede 2010). Mechanically, the raising of f_0 in oral calls compared with nasal calls could result from a rotation of the thyroid cartilage towards the cricoid cartilage (Volodin et al. 2011), resulting in increased vocal fold tension as a consequence of intrinsic laryngeal muscle contractions supported by contractions of the sternothyroid and sternohyoid muscles (Sapir et al. 1981; Hong et al. 1997).

In addition, a high-emotional arousal could entail a more powerful airflow from the lungs and thus, a higher speed of the expiratory air. The oral vocal tract is expected to have a lower impedance than the nasal vocal tract with its multiple, closely spaced surfaces. Accordingly, a phonatory air stream of increased volume and speed can more easily be released through the oral than through the nasal vocal tract. This would explain why states of higher emotional arousal correlate with oral emission. With other parameters of vocal production kept constant, the enhanced subglottal pressure should contribute to an increase of f_0 and sound amplitude (Hsiao et al. 1994; Titze 1994). Consistently, in domestic ewes and African elephants, the amplitude was higher in oral than in nasal calls (Sebe et al. 2010; Stoeger, Heilmann, et al. 2012). Probably, the emission of nasal-and-oral calls in saigas occurred when the aroused animal started calling through the nose at a fundamental frequency that was higher than that of the nasal calls (Figure 6), but was incapable of maintaining this f_0 throughout the call without opening the mouth.

Ratios and proposed functions of nasal and oral calls in mothers and young

Neonate saiga young in this study produced preferentially oral calls and only a small number of nasal calls. Although the precise proportion of nasal and oral calls could not be ascertained owing to data collection from individually unidentified animals, the nasal and

nasal-and-oral call samples were much smaller than the oral call samples. Previously, we found that at the age of 1.5 months, a young saiga produced both nasal and oral calls, whereas at the age of 3 months only nasal calls were produced (Volodin et al. 2009). This suggests a ceasing of the initially predominating oral calls along ontogeny.

However, saiga mothers were capable of producing oral calls in addition to nasal calls postpartum and during the first days of nursing. During the emission of the oral calls the gape of the mouth was well visible on the video clips. In a previous study we found that all 285 mother calls, recorded from at least 20 mothers with young aged 2–3 months, were nasal calls (Volodin et al. 2009). This suggests that the time postpartum may be an important factor decreasing the proportion of oral calls of female saigas. Subadult males at 8–9 months of age and adult males during the rut or out of the rut vocalize exclusively through the nose (Frey et al. 2007; Volodin et al. 2009).

We propose that the predominant use of oral calls by mother and young saigas in the first 10 days postpartum reflects their high level of arousal in this period. Similar data have been obtained for the goitred gazelle, in which neonates vocalize through the nose and through the mouth, but the proportion of nasal calls increases with time. At the age of 6 months, subadult goitred gazelles vocalize exclusively through the nose (Efremova et al. 2011). The goitred gazelle young produce their oral calls in situations of higher emotional arousal (Efremova et al. 2011), as well as adult ewes (Sebe et al. 2010). Concordantly, African elephant young produce both nasal and oral rumbles up to the age of 18 months, whereas older elephants produce calls primarily through the nose (Stoeger-Horwath et al. 2007). Occasionally, adult females vocalize through the mouth, whereas adult male elephants, as adult male saigas, call exclusively through the nose (Stoeger, Heilmann, et al. 2012).

Apart from expressing enhanced emotional arousal, nasalization can strongly affect vocal individuality. In juvenile goitred gazelles and in adult ewes, the oral contact calls were significantly more individualized than the nasal contact calls (Sebe et al. 2010; Volodin et al. 2011). However, neonate lambs approach any female emitting oral calls, but discriminate between their own and alien mothers only when they emit nasal calls (Sebe et al. 2010).

This study has been conducted on saigas kept in an enclosure, and females gave birth in a dense aggregation of mother and young. Are such dense aggregation and mother–offspring vocal communication comparable with those in nature? Indeed, this is similar to the natural breeding conditions of saigas, because prior to parturition, free-ranging females gather in aggregations of over 1000 individuals on restricted areas. Within these large aggregations, the females stay together in groups of 15–20 individuals with a distance of around 20 m between the individuals and of 200–300 m between groups (Bannikov et al. 1961; Sokolov and Zhirnov 1998; Danilkin 2005). In the last few years, all free-ranging saigas in Russia gave birth on a restricted area in the central parts of the Black Earths State Reserve in the Republic of Kalmykia (Danilkin 2005; Kokshunova 2012). Aggregations of pregnant saiga females are considered as an adaptation for decreasing neonate mortality by predation. Wolves *Canis lupus*, as the main predators of saiga, are strictly territorial in the spring season. Thus, only few wolf packs will hunt in the aggregation area (Bannikov et al. 1961; Sokolov and Zhirnov 1998). Saiga young are precocial followers, and within 30 min after birth are already capable of standing on their feet, suckling, walking and even try to run (Danilkin 2005; Kokshunova 2012). A few hours after birth they transfer to another place together with their mothers; after 4–5 days they follow their mothers permanently and after 10 days they are capable of following the herd and running as quickly as adults in case of danger (Sokolov and Zhirnov 1998; Danilkin 2005). Thus, dense aggregations of

pregnant females and neonates represent the typical situation for this species around the short birth period.

In conclusion, our results suggest that mother and young saigas can switch between nasal and oral vocal pathways and, thereby, can vary vocal tract lengths correspondingly. This vocal behaviour is associated with considerable variations in formants and moderate to small variations of f_0. Differences in the structure of nasal and oral contact calls are important features of mother–offspring communication.

Supplementary material

Supplementary data (sound file). Nasal, oral and nasal-and-oral calls of a mother saiga and nasal, oral and nasal-and-oral calls of a young saiga. Supplementary material for this article is available via the supplementary tab on the article's online page at http://dx.doi.org/10.1080/09524622.2013.826598.

Acknowledgements

References

Bannikov AG, Jirnov LV, Lebedeva LS, Fandeev AA. 1961. Saiga biology. Moscow: Agricultural Literature Press [in Russian].
Boersma P, Weenink D. 2009. Praat: doing phonetics by computer. Available from: http://www.praat.org/
Briefer EF. 2012. Vocal expression of emotions in mammals: mechanisms of production and evidence. Journal of Zoology 288:1–20.
Briefer EF, McElligott AG. 2011. Mutual mother–offspring recognition in an ungulate hider species (*Capra hircus*). Animal Cognition 14:585–598.
Charlton BD, Reby D. 2011. Context-related acoustic variation in male fallow deer (*Dama dama*) groans. PLoS ONE 6(6):e21066. doi:10.1371/journal.pone.0021066.
Charlton BD, Reby D, McComb K. 2007. Female red deer prefer the roars of larger males. Biology Letters 3:382–385.
Danilkin AA. 2005. Hollow-horned ruminants (Bovidae). Moscow: KMK Scientific Press [in Russian].
Dillon WR, Goldstein M. 1984. Multivariate analysis: methods and applications. New York: Wiley.
Efremova KO, Volodin IA, Volodina EV, Frey R, Lapshina EN, Soldatova NV. 2011. Developmental changes of nasal and oral calls in the goitred gazelle *Gazella subgutturosa*, a nonhuman mammal with a sexually dimorphic and descended larynx. Naturwissenschaften 98:919–931.
Fant G. 1960. Acoustic theory of speech production. Hague: Mouton & Co.
Fitch WT. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. Journal of the Acoustical Society of America 102:1213–1222.
Fitch WT, Hauser MD. 2002. Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals. In: Simmons A, Fay RR, Popper AN, editors. Acoustic communication, Springer handbook of auditory research. New York: Springer. p. 65–137.
Fitch WT, Reby D. 2001. The descended larynx is not uniquely human. Proceedings of the Royal Society B 268:1669–1675.

Frey R, Gebler A, Olson KA, Odonkhuu D, Fritsch G, Batsaikhan N, Stuermer IW. 2008. Mobile larynx in Mongolian gazelle: retraction of the larynx during rutting barks in male Mongolian gazelle (*Procapra gutturosa* Pallas, 1777). Journal of Morphology 269:1223–1237.

Frey R, Hofmann RR. 1996. Evolutionary morphology of the proboscideal nose of Guenther's dikdik (*Rhynchotragus guentheri* – Thomas, 1894) (Mammalia, Bovidae). Zoologischer Anzeiger 235:31–51.

Frey R, Hofmann RR. 1997. Skull, proboscis musculature and preorbital gland in the saiga antelope and Guenther's dikdik (Mammalia, Artiodactyla, Bovidae). Zoologischer Anzeiger 235:183–199.

Frey R, Riede T. 2013. The anatomy of vocal divergence in North American Elk and European red deer. Journal of Morphology 274:307–319.

Frey R, Volodin I, Volodina E, Carranza J, Torres-Porras J. 2012. Vocal anatomy, tongue protrusion behaviour and the acoustics of rutting roars in free-ranging Iberian red deer stags (*Cervus elaphus hispanicus*). Journal of Anatomy 220:271–292.

Frey R, Volodin IA, Volodina EV. 2007. A nose that roars: anatomical specializations and behavioural features of rutting male saiga. Journal of Anatomy 211:717–736.

Frey R, Volodin IA, Volodina EV, Soldatova NV, Juldaschev ET. 2011. Descended and mobile larynx, vocal tract elongation and rutting roars in male goitred gazelles (*Gazella subgutturosa* Güldenstaedt, 1780). Journal of Anatomy 218:566–585.

Gaeth AP, Short RV, Renfree MB. 1999. The developing renal, reproductive, and respiratory systems of the African elephant suggest an aquatic ancestry. Proceedings of the National Academy of Sciences 96:5555–5558.

Gamba M, Friard O, Giacoma K. 2012. Vocal tract morphology determines species-specific features in vocal signals of lemurs (*Eulemur*). International Journal of Primatology 33:1453–1466.

Gompper ME. 1995. *Nasua narica*. Mammalian Species 487:1–10.

Gompper ME, Decker DM. 1998. *Nasua nasua*. Mammalian Species 580:1–9.

Hauser MD, Schön Ybarra M. 1994. The role of lip configuration in monkey vocalizations: experiments using xylocaine as a nerve block. Brain and Language 46:232–244.

Heptner VG, Nasimovich AA, Bannikov AG. 1989. Mammals of the Soviet Union. Vol. 1, Ungulates Leiden (NY): E.J. Brill.

Hong KH, Ye M, Kim TM, Kevorkian KF, Berke GS. 1997. The role of strap muscles in phonation – *in vivo* canine laryngeal model. Journal of Voice 11:23–32.

Hsiao T, Solomon NP, Luschei ES, Titze IR, Kang L, Fu T, Hsu M. 1994. Effect of subglottic pressure on fundamental frequency of the canine larynx with active muscle tensions. Annals of Otology, Rhinology, and Laryngology 103:817–821.

Huntley AC, Costa DP, Rubin RD. 1984. The contribution of nasal countercurrent heat exchange to water balance in the northern elephant seal, *Mirounga angustirostris*. Journal of Experimental Biology 113:447–454.

Kamau JM, Maina JN, Maloiy GM. 1984. The design and the role of the nasal passages in temperature regulation in the dik–dik antelope (*Rhynchotragus kirkii*) with observations on the carotid rete. Respiration Physiology 56:183–194.

Kawabe M, Mano T. 1972. Ecology and behavior of the wild proboscis monkey, *Nasalis larvatus* (Wurmb), in Sabah, Malaysia. Primates 13:213–228.

Kidjo N, Cargnelutti B, Charlton BD, Wilson C, Reby D. 2008. Vocal behaviour in the endangered Corsican deer: description and phylogenetic implications. Bioacoustics 18:159–181.

Kokshunova L. 2012. Ecology and behaviour of the saiga antelope (*Saiga tatarica tatarica* L., 1766) in conditions of extreme anthropogenic press. Saarbrücken: Lambert Academic Publishing [in Russian].

Kratzing JE, Woodall PF. 1988. The rostral nasal anatomy of two elephant shrews. Journal of Anatomy 157:135–143.

Lapshina EN, Volodin IA, Volodina EV, Frey R, Efremova KO, Soldatova NV. 2012. The ontogeny of acoustic individuality in the nasal calls of captive goitred gazelles, *Gazella subgutturosa*. Behavioural Processes 90:323–330.

Lingle S, Wyman MT, Kotrba R, Teichroeb LJ, Romanow CA. 2012. What makes a cry a cry? A review of infant distress vocalizations. Current Zoology 58:698–726.

Mallon DP. 2008. *Saiga tatarica*. In: IUCN 2012, IUCN Red List of Threatened Species. Version 2012.2. Available from: http://www.iucnredlist.org

McComb K, Reby D, Baker L, Moss C, Sayialel S. 2003. Long-distance communication of acoustic cues to social identity in African elephants. Animal Behaviour 65:317–329.

McElligott AG, Birrer M, Vannoni E. 2006. Retraction of the mobile descended larynx during groaning enables fallow bucks (*Dama dama*) to lower their formant frequencies. Journal of Zoology 270:340–345.

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, McComb K, Cargnelutti B, Darwin CJ, Fitch WT, Clutton-Brock TH. 2005. Red deer stags use formants as assessment cues during intra-sexual agonistic interactions. Proceedings of the Royal Society B 272:941–947.

Riede T. 2010. Elasticity and stress relaxation of rhesus monkey *Macaca mulatta* vocal folds. Journal of Experimental Biology 213:2924–2932.

Riede T, Bronson E, Hatzikirou H, Zuberbühler K. 2005. Vocal production mechanisms in a non-human primate: morphological data and a model. Journal of Human Evolution 48:85–96.

Riede T, Tokuda IT, Munger JB, Thomson SL. 2008. Mammalian laryngeal air sacs add variability to the vocal tract impedance: physical and computational modeling. Journal of the Acoustical Society of America 124:634–647.

Sanvito S, Galimberti F. 2000. Bioacoustics of southern elephant seals. I. Acoustic structure of male aggressive vocalisations. Bioacoustics 10:259–285.

Sanvito S, Galimberti F, Miller EH. 2007. Vocal signalling of male southern elephant seals is honest but imprecise. Animal Behaviour 73:287–299.

Sapir S, Larson C, Campbell C. 1981. Effect of geniohyoid, cricothyroid and sternothyroid muscle stimulation on voice fundamental frequency of electrically elicited phonation in rhesus macaque. Laryngoscope 91:457–468.

Sebe F, Duboscq J, Aubin T, Ligout S, Poindron P. 2010. Early vocal recognition of mother by lambs: contribution of low- and high-frequency vocalizations. Animal Behaviour 79:1055–1066.

Sokolov VE, Zhirnov LV. 1998. The saiga antelope. Phylogeny, systematic, ecology, conservation and use. Moscow: Russian Academy of Sciences [in Russian].

Soltis J, Blowers TE, Savage A. 2011. Measuring positive and negative affect in the voiced sounds of African elephants (*Loxodonta africana*). Journal of the Acoustical Society of America 129:1059–1066.

Stoeger AS, Heilmann G, Zeppelzauer M, Ganswindt A, Hensman S, Charlton BD. 2012. Visualizing sound emission of elephant vocalizations: evidence for two rumble production types. PLoS ONE 7(11):e48907. doi:10.1371/journal.pone.0048907.

Stoeger AS, Mietchen D, Oh S, de Silva S, Herbst CT, Kwon S, Fitch WT. 2012. An Asian elephant imitates human speech. Current Biology 22:2144–2148.

Stoeger-Horwath AS, Stoeger S, Schwammer HM, Kratochvil H. 2007. Call repertoire of infant African elephants: first insights into the early vocal ontogeny. Journal of the Acoustical Society of America 121:3922–3931.

Taylor AM, Reby D. 2010. The contribution of source–filter theory to mammal vocal communication research. Journal of Zoology 280:221–236.

Titze IR. 1994. Principles of voice production. Englewood Cliffs, NJ: Prentice-Hall.

Titze IR. 2008. Nonlinear source–filter coupling in phonation: theory. Journal of the Acoustical Society of America 123:2733–2749.

Titze IR, Riede T. 2010. A cervid vocal fold model suggests greater glottal efficiency in calling at high frequencies. PLoS Computational Biology 6(8):e1000897. doi:10.1371/journal.pcbi.1000897.

Todd NE. 2010. Qualitative comparison of the cranio-dental osteology of the extant elephants, *Elephas maximus* (Asian elephant) and *Loxodonta africana* (African elephant). Anatomical Records 293(1):62–73.

Volodin I, Volodina E, Frey R, Carranza J, Torres-Porras J. 2013. Spectrographic analysis points to source-filter coupling in rutting roars of Iberian red deer. Acta Ethologica 16:57–63.

Volodin IA, Lapshina EN, Volodina EV, Frey R, Soldatova NV. 2011. Nasal and oral calls in juvenile goitred gazelles (*Gazella subgutturosa*) and their potential to encode sex and identity. Ethology 117:294–308.

Volodin IA, Volodina EV, Efremova KO. 2009. Antelope, calling through the nose: structure of sounds and effect of sexual selection on the vocal behavior of the saiga (*Saiga tatarica*). Zoologicheskii Zhurnal 88:113–124 [in Russian].

Witmer LM, Sampson SD, Solounias N. 1999. The proboscis of tapirs (Mammalia: Perissodactyla): a case study in novel narial anatomy. Journal of Zoology 249:249–267.

Wyman MT, Mooring MS, McCowan B, Penedo MCT, Reby D, Hart LA. 2012. Acoustic cues to size and quality in the vocalizations of male North American bison, *Bison bison*. Animal Behaviour 84:1381–1391.