

# Development of aggressive vocalizations in male southern elephant seals (*Mirounga leonina*): maturation or learning?

Simona Sanvito<sup>1,2,3</sup>, Filippo Galimberti<sup>2</sup> & Edward H. Miller<sup>1</sup>

(<sup>1</sup>Department of Biology, Memorial University of Newfoundland, St. John's NL A1B 3X9, Canada; <sup>2</sup>Elephant Seal Research Group, Sea Lion Island, Falkland Islands)

(Accepted: 28 September 2007)

---

## Summary

Vocalizations are an important component of male elephant seal agonistic behaviour. Acoustic and behavioural components of vocalizations emitted during agonistic contests show gross differences between young and old males, but the variation with age depends on the specific feature. Vocalizations become more frequent and effective at later ages. Acoustic features that are constrained by structural phenotype, which changes with age, also should change with age, while acoustic features that are independent from structural phenotype should show no relationship with age. We demonstrate that, in southern elephant seals, formant frequencies, which are constrained by the vocal tract length and, therefore, by body size, show a clear decrease with age, whereas temporal and structural features of sounds, which potentially are unconstrained, show no correlation with age. Formants ontogeny seems, therefore, to be mostly the result of body maturation, and hence formants may be reliable signals of age. In contrast, acoustic features such as temporal features and syllable structure, are free to change, and hence may serve as the raw material for vocal learning and individual recognition.

*Keywords:* vocalization, formants, vocal development, vocal learning, communication, Pinnipedia, southern elephant seal, *Mirounga leonina*.

---

<sup>3</sup>) Corresponding author's address: Simona Sanvito, c/o prof. Yolanda Schramm, Laboratorio de Ecología Molecular, Facultad de Ciencias Marinas, Universidad Autónoma de Baja California, km 103 carretera Tijuana-Ensenada, Ensenada, 22860, Baja California, Mexico, e-mail: simo\_esrg@eleseal.it

## Introduction

Vocal development has been investigated in many different mammalian taxa, such as primates (Gautier, 1998; Pistorio et al., 2006), marsupials (Aitkin et al., 1996), bats (Van Parijs & Corkeron, 2002; Vater et al., 2003), rodents (Mandelli & Sales, 2004), terrestrial Carnivora (Schassburger, 1993; Chaadaeva, 2002; Chaadaeva & Sokolova, 2005) and cetaceans (Snowdon & Hausberger, 1997; Tyack, 1997; Killebrew et al., 2001). Unfortunately, most of these studies dealt with a small part of the animals' lifespan, usually focusing on the very early stages of life and, therefore, do not offer a full picture of vocal development into adulthood. Moreover, most of these works were conducted in captive or semi captive settings, with few notable exceptions (Seyfarth & Cheney, 1986) and, therefore, may not be representative of the social conditions in which animals develop their natural vocal skills. In pinnipeds, vocal development has been quite extensively investigated (Shipley et al., 1986; Miller, 1991) but, as for other taxa, the vast majority of the studies regarded the development of calls in pups during the lactation period, i.e., in a very short part of their lifespan (see, for example, Job et al., 1995; Khan et al., 2006). In this paper we present the first detailed, longitudinal study on vocal development from puberty to adulthood, of any pinniped species.

Southern elephant seals (*Mirounga leonina*; SES hereafter) breed on land during the austral spring, when females gather in large groups called 'harems', and males compete to establish a more-or-less linear dominance hierarchy that regulates access to breeding females (Galimberti et al., 2003). Male-male competition and inequality in breeding success are among the highest recorded for any vertebrate species (Galimberti et al., 2002). The general structure of agonistic behaviour of elephant seals has already been described, in both the southern (Laws, 1956; McCann, 1981) and the northern elephant seal (Bartholomew & Collias, 1962; Sandegren, 1976; Cox, 1981). Agonistic behavioural sequences include optical and acoustic displays, chases, and direct aggression and most contests are settled by vocalizations (Sanvito et al., 2006a). A very important aspect of male competition in elephant seals is age. Fights usually involve males of similar age (Braschi, 2004), and the interactions between males of different age classes are often settled by vocalizations (Sanvito et al., 2006a). 'Aggressive' vocalizations (Sanvito & Galimberti, 2000) play an important role in agonistic encounters among males (Sanvito et al., 2006a), for settlement of contests and establishment of dominance relationships. Studies by other authors (Shipley et al., 1986) and previ-

ous preliminary results (unpublished data) showed that various behavioural and acoustic features of elephant seal males' vocalizations change with age. This relationship suggests that vocalizations may act as honest signals of the age component of male resource holding potential (Fitch & Hauser, 2002). For instance, the general macro-structure of male vocalizations (i.e., their temporal pattern and syllables composition) changes gradually from a plastic, no fixed structured vocal pattern in young animals (i.e., each male emits different vocal patterns in different vocalizations) towards a strongly structured and stereotyped one in adults (i.e., each male always emits the same vocal pattern; Sanvito et al., 2006a).

Vocal development can result from two main processes: (i) learning of acoustic features through auditory experience (Egnor & Hauser, 2004), for example, by imitation of older individuals (Sanvito et al., 2006a); and (ii) developmental changes in body morphology or body size (called 'structural phenotype' hereafter) that affect vocal tract characteristics (e.g., size of the vocal folds, or size or shape of resonating cavities). The following hypotheses can discriminate between these two processes:

1. Vocal learning is expected to shape acoustic features that are not constrained by structural phenotype, whereas physical maturation is expected to shape features that are directly linked to it
2. Vocal learning is expected to reduce the within-individual variation with age, because each male will increase his vocal competence, and will develop a fixed vocalization pattern from a flexible one (Sanvito et al., 2006a), whereas physical maturation is not expected to affect within-individual variation, because the acoustic features affected by maturation will be constrained by phenotype at all ages
3. Physical maturation is expected to produce directional trends with age in mean value of acoustic features that are structurally linked to physical growth; however, vocal learning is not expected to generate such a trend, because the features that can be learned are not constrained by structural phenotype
4. Physical maturation is expected to produce a trend in among-individual variability of acoustic features, because these features will reflect the among-individual variation in structural phenotype at each age, whereas vocal learning is not expected to produce such a trend because of the lack of constraints in the features that can be learned.

Previous evidences of vocal development in male elephant seals regards the northern species (*M. angustirostris*) only, and are limited by the cross-sectional sample collected during a brief period of time, the small number of acoustic features considered, and the lack of information on age or body size of the individuals involved (Shipley et al., 1986). In long-lived species, the study of vocal ontogeny requires a longitudinal data set on vocalizations of marked individuals for which age and body size over growth are known. In this paper we describe the development of vocalizations in a large sample of SES males at a small and localized breeding colony in the Falkland Islands. We analyze variation with age of acoustic and behavioural features of vocalizations, we assess within- and among-individual variation of acoustic features, and we compare the observed developmental trends of acoustic features with the expectations of the vocal learning and physical maturation processes.

## Material and methods

### *Study site and marking of animals*

Data were collected during eight breeding seasons (September–November, 1995 to 2002) at Sea Lion Island (Falkland Islands), which shelters a small and localized population of SES, comprising approx. 550 breeding females and approx. 60 breeding males (Galimberti & Sanvito, 2001). All males were individually recognizable because they had been marked with cattle tags during previous breeding seasons, soon after birth, or after arrival on land during their first breeding season. All breeding males were also marked with hair dye to permit fast and safe identification. Further details on study site, population demography and marking can be found in Galimberti & Boitani (1999).

### *Estimation of age*

Age was known for males tagged as pups and estimated to  $\pm 1$  year for other males using external morphology (Clinton, 1994; Galimberti & Boitani, 1999). Our estimate of age was based on scarring of the chest and development of the proboscis, and was independent of body size. We checked these estimated ages for (i) intra-observer reliability, (ii) inter-observer reliability, (iii) congruence of age attributed in following breeding seasons and

(iv) correspondence with actual age (for the males born after the beginning of our study). Reliability was calculated using the Kendall coefficient of concordance (Siegel & Castellan, 1988), on the age attributed to marked males during a random sample of 10 daily censuses. Mean intra-observer reliability was 0.95, and inter-observer reliability ranged from 0.93 to 0.99 for 2–4 observers. Congruence of the whole classification was checked using lifetime records of the males that were present for three or more breeding seasons and by comparing estimated age to actual age for males tagged at birth. In the analysis where age is involved, we used only males for which either (a) true age or (b) a reliable age estimate from morphological development over several breeding seasons were available (Clinton, 1994). Males were classified as ‘young’ ( $\leq 8$  years old) and ‘old’ ( $\geq 9$  years old).

#### *Behavioural data collection and analysis*

We obtained behavioural data during 7852 h of observation, during standard periods of 2-h length. A total of 25 671 social interactions between males were observed. Detailed behavioural sequences of interacting males were recorded for 5099 of these interactions. Behavioural data were available for 183 individually recognized males, present during one to six breeding seasons (mean of 1.83 seasons per males), spanning a range of ages from 6 to 14 years. During each observation period, we recorded, for each male present, its total number of interactions, number of interactions won, number of interactions settled by vocalization, number of interactions in which the male vocalized one or more times, total number of vocalizations emitted, number of transitions of modules in the behavioural sequence involving the vocalization, and number of bouts for each vocalization. For each of the preceding, we counted total occurrences and occurrences when the subject male initiated interactions or responded to initiation by another male. More details on the protocol are given in Galimberti et al. (2000).

Original behavioural data were counts, which were converted to percentages and means for presentation, but retained unconverted for the statistical analysis. To analyze relationships with age, we compared three models suitable for regression on count data (Long, 1997; Cameron & Trivedi, 1998): Poisson regression, negative binomial regression, and zero-inflated negative binomial regression. The Poisson regression is suitable only when the mean is approximately equal to the variance. Our data showed over-dispersion for

all variables, making the negative binomial regression a more appropriate model (Gardner et al., 1995). Moreover, due to influences of the presence of older males, younger males may have avoided interacting or vocalizing at all, which may have produced an excess of zero counts for the young class. Therefore, we also considered the zero-inflated variant of the negative binomial, with age as the generating variable for the zero inflation process (Cheung, 2002). Males that were observed for <20 h were excluded. To account for the lack of balance in the sampling of males, which was unavoidable due to the different patterns of presence on land of males of different age classes (Galimberti & Boitani, 1999), we incorporated the total number of observation periods in which a male was observed as exposure factor in regression models. We tested each model using a likelihood-ratio (LR) test of the model including age versus the null (intercept only) model. The data set was longitudinal, so autocorrelation within individuals may have produced an inflation of standard errors and probabilities (Diggle et al., 1994). Therefore, we calculated robust standard errors with the Huber–White sandwich estimator of variance (Williams, 2000), which accounts for within-male clustering, and we calculated Wald tests on the regression coefficients. All results of the Wald tests were in accordance with LR tests, so we present only the latter. We compared the models using the Akaike information criterion (AIC).

To simplify interpretation, we transformed regression coefficients to the expected percentage change in counts due to a unit change in the regressor (age in this case), and to the expected percentage change due to a standard deviation change in the regressor (Long & Freese, 2001).

### *Sound recording*

We recorded male agonistic vocalizations by standard stimulation of the animals (Sanvito & Galimberti, 2000) with Sony DAT recorders TCD-D7 and TCD-D100 and a Sennheiser MD 441 dynamic cardioid microphone (frequency range 30–20 000 Hz). Recordings were digitized at 48 kHz with 16-bit resolution.

Vocalizations were recorded over the 3-month breeding season each year, for a total of 7405 vocalizations (1–64 vocalizations per male per year). From the recordings, 2007 vocalizations from 196 different males were selected for analysis, with males recorded in 1–6 breeding seasons (mean  $\pm$  SD =  $1.7 \pm 1.1$ ). Fifty-five males were recorded for  $\geq 3$  seasons, and 29 were followed throughout their entire vocal development, from their initial variable

unstructured calls to their stereotyped structured adult vocalizations. To analyze balanced samples, five recordings per male per year were randomly chosen. The only exception was 2002, when two vocalizations per male per week were analyzed to investigate intra-seasonal variation.

### *Spectral analysis*

The following settings were used for spectral analysis (Charif et al., 1995): Hamming window function with frame length of 21.33 ms (1024 points) and corresponding filter bandwidth of 190.31 Hz; frame overlap of 50% with time-grid resolution of 10.67 ms and frequency-grid resolution of 11.72 Hz (fft = 4096 points). Spectral settings were chosen to resolve the pulse-train structure of the vocalizations and for good frequency resolution (Sanvito & Galimberti, 2000). Frequency modulation is almost absent from SES vocalizations (Sanvito & Galimberti, 2000), so for analysis of amplitude spectra we calculated average spectra for entire bouts (Phillips & Stirling, 2000).

### *Acoustic measurements*

Sound measurement and spectral analysis were carried out with Canary 1.2 (Macaulay Library, Cornell Lab of Ornithology, Ithaca, NY, USA), and with programs written in Igor Pro 4.0.9 (WaveMetrics) and in Revolution 2.0 (Runtime Revolution). Male elephant seal aggressive vocalizations are composed of a series of sound emissions, called 'bouts', which are divided into 'syllables' and 'syllable parts' (Sanvito & Galimberti, 2000). A syllable is a single acoustic event, with a continuous spectrographic track over time. A syllable part is a portion of a syllable with an approximately constant pulse rate.

We measured five classes of variables:

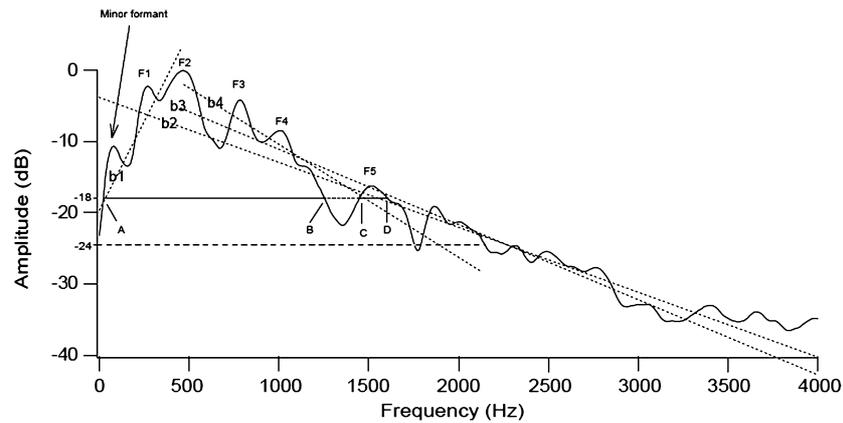
1. Temporal variables. Measured on waveforms (Table 1)
2. Frequency variables. Measured on average power spectra of bouts, spectrograms and waveforms (Figure 1 and Table 1). Particular attention is needed when considering formant like frequencies. Formants are parts of the frequency spectrum that are reinforced by resonant properties of the vocal tract (Fry, 1979; Miller & Murray, 1995; Reby & McComb, 2003). In spectra and spectrograms of SES vocalizations, some frequencies are enhanced over others, and appear as dark bands (Sanvito & Galimberti, 2000). We refer to these as

**Table 1a.** Description of temporal and frequency variables.

Variable name	Unit	Description
(A) Temporal		
Number of bouts per vocalization	bouts	Total number of bouts composing a vocalization
Bout duration	s	Time between beginning and end of bout
Interval between bouts	s	Time between end of bout and beginning of following bout
Vocalization duration	s	Time between beginning of first bout and the end of last bout of a vocalization
Duty cycle within bout		(sum of bout durations)/total vocalization duration (i.e., the proportion of vocalization with signal)
Relative peak time		(interval from beginning of bout to time of highest amplitude in the bout)/bout duration
Max syllable part	s	Duration of syllable part in which the maximal peak pressure occurs
(B) Frequency		
Dominant frequency	Hz	Frequency at which, on average, the highest amplitude in a bout occurs (= frequency of highest peak in a bout's amplitude spectrum)
Fundamental frequency (F0)	Hz	Pulse rate in longest or most intense portion of bout (calculated from waveform; Zuberbuhler et al., 1997)
Peak frequency	Hz	Frequency at which highest amplitude in a bout occurs (calculated on spectrogram of a bout as the frequency at which the highest amplitude peak occurs)
3, 6, 12, 18 dB bandwidth	Hz	Width of frequency band around dominant frequency where signal attenuates by 3, 6, 12, 18 dB (calculated on average spectrum of a bout by finding two frequencies $F_a$ and $F_b$ around the dominant frequency ( $F_a < \text{dominant frequency} < F_a$ ) where spectrum level is 3, 6, 12, 18 dB below peak value; bandwidth = $F_b - F_a$ )
Min frequency at -3, 6, 12, 18 dB	Hz	$F_a$ of previous definition (i.e., minimum frequency at which power spectrum goes 6 dB below the amplitude of the highest peak)
Max frequency at -3, 6, 12, 18 dB	Hz	$F_b$ of previous definition (i.e., maximum frequency at which power spectrum goes 6 dB below the amplitude of the highest peak)
3, 6, 12, 18 dB bandwidth proportion		Proportion of the frequency bandwidth in which spectrum is actually above the amplitude of -3, 6, 12, 18 dB
Formant like frequencies (F1-F5)	Hz	See text for explanation
Formant dispersion	Hz	$(F5 - F1)/4$ . This measure averages spacing between consecutive formants
Minor formant (Fm)	Hz	See text for explanation

**Table 1b.** Description of variables related to sound amplitude, energy distribution and internal bout structure.

Variable name	Unit	Description
<b>(C) Sound amplitude</b>		
Instantaneous relative peak intensity		Ratio between instantaneous peak intensity per Hertz (maximal intensity/Hz in bout, calculated from spectrogram) and peak pressure (maximal pressure in selected bout, calculated from waveform). It measures the instantaneous effect of the frequency band with maximal energy on all the energy emitted on all frequency bands
Peak intensity predominance		Ratio between relative average intensity of whole bout and peak pressure. The value ranges from 0 to 1 (1: maximal intensity is equal to average intensity of whole bout; <1: maximal intensity is increasingly greater than average intensity of whole bout)
<b>(D) Energy distribution</b>		
Power spectrum ascending slope (= spectrum ascending slope)		Slope of linear regression line fitted through the spectrum, from the beginning to the highest peak
Power spectrum 0/4000 Hz slope (= spectrum 0/4000 Hz)		Slope of linear regression line fitted through the spectrum, from 0 to 4000 Hz
Power spectrum maximal amplitude/4000 Hz slope (= spectrum max/4000 Hz slope)		Slope of linear regression line fitted through the spectrum, from the highest peak to 4000 Hz
Power spectrum maximal amplitude/-24dB slope (= spectrum max/-24 dB slope)		Slope of linear regression line fitted through the spectrum, from the highest peak to the point where the spectrum goes 24 dB below that peak value
<b>(E) Bout structure</b>		
Number of syllables per bout	syllables	Total number of syllables (equal or not) per bout
Syllable rate	syllable/s	Number of syllables/bout duration
Number of syllable types per bout	syllable types	Number of different kinds of syllables per bout
Shannon index		Index of evenness of syllable composition calculated as $H/H_{max}$ (Krebs, 1989); $H$ is the sum from 1 to $n$ of $\pi_i \ln(\pi_i)$ , $n$ is the number of syllables in the bout, $\pi_i$ is the relative frequency of syllables of type $i$ in the bout, $\ln$ is the base 2 logarithm, and $H_{max}$ is the expected maximum $H$ for the number of syllables of the bout, calculated as $\ln(n)$



**Figure 1.** Graphical illustration of some acoustic variables. Spectral peaks indicated F1–F5 are formant frequencies (dominant frequency = F2 in this example). Minor formant is frequency of first spectral peak, usually < 100 Hz (not always present). Variables relative to frequency bandwidth are indicated only for –18 dB; those for –3, –6 and –12 dB were computed similarly. Frequencies at A and D are minimal and maximal frequencies at –18 dB, respectively. AD represents the –18 dB bandwidth. 18 dB bandwidth proportion is calculated as  $(AB + CD)/(AD)$ . The proportion can be equal to 1 when there are no spectral troughs going below 18 dB from the highest spectral peak, within the –18 dB bandwidth. b1 = power spectrum ascending slope; b2 = power spectrum 0/4000 Hz slope; b3 = power spectrum max amplitude/4000 Hz slope; b4 = power spectrum max amplitude/–24 dB slope.

formant-like frequencies (formants hereafter). Presumably, they reflect the resonant frequencies of the vocal tract. Formants were measured from average amplitude spectra for each male, as the first 5 evident frequency peaks (F1 to F5; see below for further details on formant estimation). In some cases, we found a very first peak in the power spectrum, at lower frequency (often <100 Hz) and reduced intensity with respect to F1 and just preceding it. In many cases it was not present, or just slightly evident as a ‘shoulder’ on the lower side of the power spectrum. We called this minor formant (Fm), which seems to be the first formant produced by the resonant properties of the nasal tract (Sanvito et al., 2006b)

3. Sound amplitude variables. Measured from average power spectra, spectrograms and waveforms (Table 2)
4. Variables associated with energy distribution over frequency. Measured from power spectra (Figure 1 and Table 2). The energy distribution of sound is important in communication but is difficult to

**Table 2.** Relationships of behavioural measures of social interaction and vocalization to age of male southern elephant seals. int. = interaction, seq. = behavioural sequence; VO = vocalization.

	Age (years)													
	6	7	8	9	10	11	12	13	14	All				
Number of males	42	83	66	57	42	24	9	8	3	334				
% of males	12.6	24.9	19.8	17.1	12.6	7.2	2.7	2.4	0.9					
Mean number of interactions	52.9	80.7	102.4	129.6	209.5	238.5	290.2	338.9	392.0	132.0				
Mean% bi-directional int.	8.0	9.8	9.9	10.2	6.9	5.5	7.1	8.5	2.7	8.8				
Mean% int. in which male = actor	14.3	20.7	34.7	47.5	60.8	75.7	94.1	89.1	98.1	40.5				
Mean% int. won	11.7	19.2	33.8	47.9	60.8	75.0	93.8	88.6	97.4	39.7				
Mean% int. settled by VO	32.6	37.8	44.9	43.3	47.6	46.8	47.3	45.5	53.4	42.0				
Mean% int. settled by VO and won	2.1	4.4	14.0	19.6	28.1	34.1	44.2	42.9	51.3	16.1				
Mean% int. settled by VO with VO only	1.4	6.7	15.0	20.1	26.5	29.2	32.1	34.0	47.9	15.8				
Mean% int. with VO	5.7	14.2	25.2	35.8	50.1	62.0	62.1	74.0	82.4	30.3				
Mean% int. with bi-directional VO	1.4	3.9	6.2	5.7	10.3	4.6	4.3	3.6	0.0	5.1				
Mean% int. male = actor and with VO	4.6	11.6	20.1	29.0	43.0	58.2	58.7	69.8	80.9	25.9				
Mean% int. male = reactor and with VO	1.1	2.6	5.2	6.8	7.1	3.9	3.4	4.2	1.5	4.4				
Mean% int. male react to VO with VO	0.8	1.4	2.4	3.5	4.7	2.6	0.8	3.1	0.0	2.4				
Mean number of VO per seq.	0.039	0.214	0.084	0.163	0.456	0.604	0.789	1.360	1.003	0.345				
Mean number of VO as actor per seq.	0.010	0.203	0.072	0.147	0.428	0.574	0.786	1.340	0.994	0.325				
Mean number of VO as reactor per seq.	0.029	0.010	0.013	0.016	0.028	0.030	0.002	0.020	0.010	0.020				
Mean number of bouts per seq.	0.039	0.426	0.160	0.315	0.907	1.084	1.548	1.680	1.502	0.637				
Mean number of bouts as actor per seq.	0.010	0.410	0.140	0.289	0.879	1.066	1.547	1.660	1.486	0.615				
Mean number of bouts as reactor per seq.	0.029	0.017	0.020	0.025	0.028	0.018	0.002	0.020	0.017	0.022				

summarize. Overall spectral shape might be informative (Owren & Linker, 1992), so we calculated some summary measures of it

5. Variables associated with internal bout structure. Measured from waveforms and spectrograms (Table 2)

Unless specified otherwise, we always measured variables at the bout level, and then calculated average values for vocalizations and for individual males for further analysis.

#### *Estimation of formants*

Many different methods have been proposed to estimate formants, such as visual assessment of spectrograms and spectra (Insley, 1992), Linear Predictive Coding (LPC; Owren & Bernacki, 1988; Owren & Linker, 1992; Riede & Fitch, 1999; Fitch & Reby, 2001; Reby & McComb, 2003) and custom methods applied to specific vocalizations (Darden et al., 2003). LPC has been used increasingly (Reby & McComb, 2003; Fischer et al., 2004), but we decided not to apply it in this study because, to properly apply LPC, a good knowledge of the vocal tract of elephant seals would have been required. Unfortunately, the knowledge of vocal tract anatomy for this species is almost non-existent. Moreover, male elephant seals have a proboscis that affects vocal production (Laws, 1953; Sanvito et al., 2006b) and presumably alters sound properties relative to the single resonator mechanism of sound production in other mammals (Fitch, 1997; Riede & Fitch, 1999; Fitch & Reby, 2001; Sanvito et al., 2006b). Even in humans, nasal sounds complicate evaluation of formants using LPC (Monsen & Engebretson, 1983; Miller, 1991). Therefore, we preferred to estimate formants by the usual FFT non-parametric technique, which requires less stringent assumptions about the mechanism of sound production.

#### *Modelling of acoustic variation with age*

Visual inspection of scatter plots of some acoustic features versus age with fitted LOWESS smoother (Trexler & Travis, 1993) showed a clear bending point in some cases. Therefore, we fitted the following piecewise linear model with unknown bending point (Muggeo, 2003; Sanvito et al., 2006b):

$$\text{Trait} = a_0 + (b_1 * \text{Age}) + (b_d * (\text{Age} - \text{BP}) * (\text{Age} > \text{BP})),$$

where BP is the bending point,  $a_0$  is the intercept of the regression line before the bending point,  $b_1$  is the slope of the regression line before the bending

point,  $b_d$  is the difference in the slope of the regression line before and after the bending point,  $\text{Age} > \text{BP}$  is a logic condition, returning 0 for ages  $\leq \text{BP}$  and 1 for ages  $> \text{BP}$ . Therefore, the slope of the regression line after the bending point ( $b_2$ ) is equal to  $b_d + b_1$ . We tested the difference between the slope before and after the bending point, by testing the null hypothesis  $b_d = 0$  (for which  $b_1 = b_2$ ). The piecewise models were fitted using the non-linear least squares module of Systat v. 11 software. Models were compared using AIC. A difference in AIC  $\geq 2$  gives moderate evidence that the model with the lower AIC should be preferred, and a difference  $\geq 7$  gives compelling evidence that the model with the lower AIC should be preferred (Burnham & Anderson, 1998).

#### *Measurement error and repeatability of acoustic features*

Choice of starting and ending points of each bout was performed manually, hence was associated with measurement error. To estimate this error, we choose 25 males at random, and then we chose one vocalization for each of these males, again at random. Each vocalization of this set of 25 was measured three times, in random order. Measurement error was then estimated from variance components of a Model II ANOVA (Bailey & Byrnes, 1990). Percentage measurement error was calculated as the percentage of within-bout variance over the total variance, where the total variance was the within-bout variance plus the among-bout variance.

We calculated repeatability ( $R =$  intra-class correlation coefficient) of all variables as the proportion of the among-individuals variance on the total variance, using a Model II ANOVA (Lessells & Boag, 1987). Confidence limits of repeatability were calculated using a jack-knife delete-one procedure, and its difference from zero was tested using randomization (10 000 re-samplings).

#### *Statistics: general*

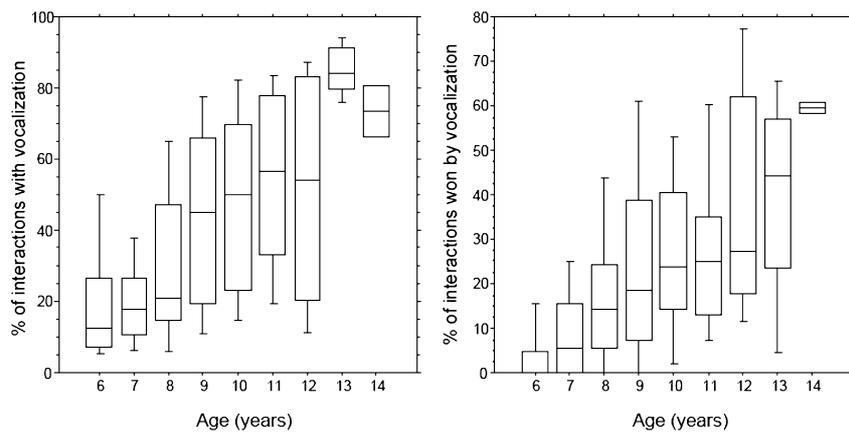
We present statistics as mean  $\pm$  standard deviation (SD) and least-squares estimates as estimate  $\pm$  standard error (SE). As a relative variability measure we calculated the coefficient of variation (CV) of each variable for each male. We tested normality using the Shapiro–Wilk test, homogeneity of means of two groups (i.e., ‘old’ and ‘young’) using a  $t$ -test with randomization, and homogeneity of variances using the Brown–Forsythe test with

randomization. In box plots, the upper hinge of the box was the 75% percentile and the lower was the 25%; the line within the box was the median; the upper whisker was 10% percentile and the lower was the 90%. For variables with non-normal asymmetric distributions, we calculated standard errors and confidence limits using a jack-knife delete-one procedure (Mitchell-Olds & Shaw, 1987), and conducted significance tests with randomization (Manly, 1997). The number of re-samplings used in randomization tests is stated as a subscript of the  $p$  label. In case of multiple non-independent tests, we calculated adjusted probabilities using a sequential Bonferroni procedure (Hochberg & Benjamini, 1990). In all cases in which the data structure was clearly longitudinal, we fitted random regression models, where the identity of the males was the random effect. General statistics procedures were carried out with STATA v. 9 (Stata, College Station, TX, USA).

## Results

### *Development of vocal behaviour*

The mean of most measures of agonistic and vocal behaviour increased with age (Table 2; Figure 2). For most variables, the regression model that best fitted the data was the negative binomial (Table 3), confirming the presence of over-dispersion. The models including age produced a significant improvement of the fit with respect to the null (intercept only) model (all  $l_r$  test



**Figure 2.** The use and effectiveness of vocalizations in interactions between male southern elephant seals increased with age.

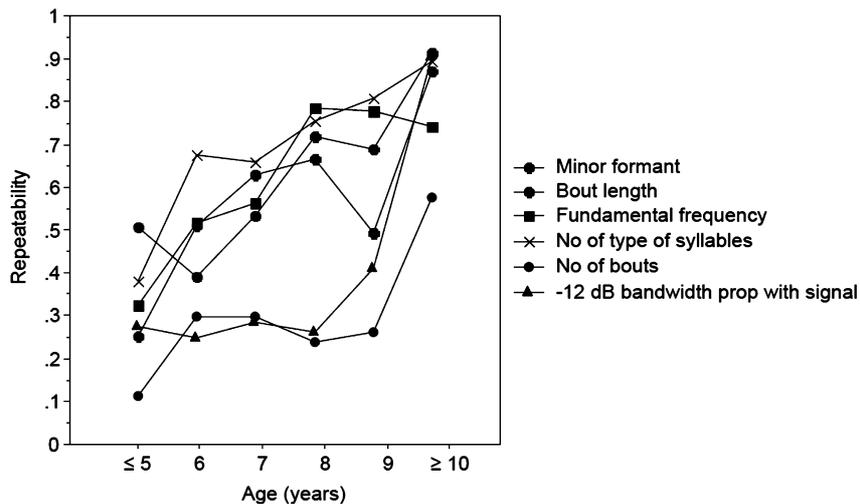
**Table 3.** Comparison of count regression models applied to the behavioural measures of social interaction and vocalization plus parameters and tests of the best fitting model. p = Poisson regression, nb = negative binomial regression, zinb = zero-inflated negative binomial regression, AIC = Akaike information criterion, Model = model chosen using AIC difference; lr test = likelihood ratio test of the model including age versus the null (intercept only) model (df = 1), \* = significant at alpha = 0.05 (sequential Bonferroni correction, Holm method), b = regression coefficient, SE(b) = robust standard error of the regression coefficient (see Methods), 95% CI(b) = 95% confidence interval of the regression coefficient; % unit = percent change in the variable due to a unit (one year) change in age; % SD = percentage change in the variable due to a one standard deviation change in age. int. = interaction, trans. = behavioural transition, vo = vocalization.

Variable	p vs nb	nb vs zinb	Model	lr test	b	SE(b)	95% CI(b)	% unit	% SD
Total number of interaction	27.161	no 0	nb	$\chi^2 = 37.6, p < 0.0001^*$	0.075	0.017	0.042, 0.108	7.8	14.4
Bi-directional interactions	1.773	0.014	p	$\chi^2 = 65.3, p < 0.0001^*$	-0.083	0.010	-0.104, -0.063	-8.0	-14.0
Male is actor	30.797	0.007	nb	$\chi^2 = 215.5, p < 0.0001^*$	0.403	0.027	0.349, 0.456	49.6	106.8
Male win	31.187	0.044	nb	$\chi^2 = 208.6, p < 0.0001^*$	0.421	0.027	0.346, 0.477	52.3	113.7
Interaction is ritualized	21.674	-0.008	nb	$\chi^2 = 34.5, p < 0.0001^*$	0.116	0.022	0.072, 0.159	12.2	23.2
Interaction settled by vo	22.053	0.000	nb	$\chi^2 = 35.4, p < 0.0001^*$	0.127	0.024	0.080, 0.175	13.6	25.8
Male vocalize	26.168	0.000	nb	$\chi^2 = 117.7, p < 0.0001^*$	0.502	0.053	0.398, 0.606	65.2	147.5
Vocalization is bi-directional	0.143	0.036	p	$\chi^2 = 3.0, p = 0.0858$	-0.060	0.046	-0.150, 0.031	-5.8	-10.2
Male replays with vo to vo	0.001	0.000	p	$\chi^2 = 2.1, p = 0.1505$	-0.071	0.054	-0.178, 0.035	-6.9	-12.1
Male act (initiate) with vo	26.475	0.001	nb	$\chi^2 = 117.0, p < 0.0001^*$	0.534	0.057	0.422, 0.645	70.5	172.1
Male react with vo	0.241	0.042	p	$\chi^2 = 2.9, p = 0.0878$	0.477	0.033	-0.017, 0.112	4.9	9.0
Number of vo in trans.	26.280	0.105	nb	$\chi^2 = 42.5, p < 0.0001^*$	0.862	0.127	0.613, 1.110	136.7	373.8
Num. of vo in trans. as actor	25.945	0.101	nb	$\chi^2 = 43.7, p < 0.0001^*$	0.900	0.139	0.629, 1.171	146.0	408.1
Number of bouts in vo	54.947	0.106	nb	$\chi^2 = 37.2, p < 0.0001^*$	0.912	0.147	0.623, 1.200	148.9	418.8
Num. of bouts in vo as actor	54.604	0.105	nb	$\chi^2 = 37.6, p < 0.0001^*$	0.948	0.165	0.625, 1.271	158.1	453.7

$p < 0.0001$ ). In particular, a significant increasing trend with age was observed for: the number of interactions in which the male is the actor (percent change per one year change in age = 49.6%); the number of interactions won (52.3%); the number of interactions settled by ritualized displays (= without direct contact, chase or fight; 12.2%); the number of interactions settled by vocalization (13.6%); the number of interactions in which the male uses vocalization (65.2%); the number of interactions in which the male uses vocalization to initiate the interaction (70.5%); the number of behavioural transitions involving vocal modules (136.7%); the number of behavioural transitions in which the male is the initiator and the behavioural module is the vocalization (146.0%); the number of bouts emitted (148.9%); the number of bouts emitted when the male is the initiator of the behavioural transition (158.1%).

*Measurement error, repeatability and seasonal changes of acoustic features*

Percent measurement error for start and end time of bouts was about 1%. Repeatability of different variables averaged  $0.513 \pm 0.155$  (range 0.210–0.861). Only 21% of the repeatabilities were below 0.400 and 8% below 0.300. Highest values were for structural features and bout duration. Repeatability increased with age for many variables (Figure 3) but did not vary



**Figure 3.** Repeatability of some variables of vocalizations of southern male elephant seals increased with age.

seasonally (based on visual inspections of scatter plots with fitted LOWESS smoother of acoustic features versus the day of the breeding).

#### *Acoustic features and age*

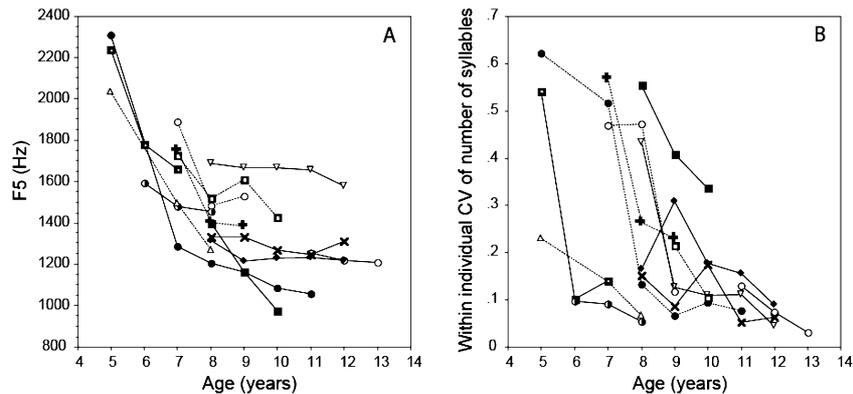
Preliminary analyses revealed many differences between vocalizations of ‘young’ and ‘old’ males. Measures of temporal and structural variables were greater and most frequency measures were lower for ‘old’ males (Table 4). Analyses incorporating true age revealed: (i) little age-specific variation and a clear trend over age for just few variables; (ii) a high within-age variation (see Appendix 1). To consider the longitudinal structure of our data, we ran a preliminary random-effect regression analysis for all acoustic variables, with male identity as the random effect. Only formant frequencies showed clear (decreasing) trends with age, in particular for upper formants (Table 5; Figure 4).

**Table 4.** Vocalizations differed between ‘young’ and ‘old’ male southern elephant seals (only significant differences shown). Mean  $\pm$  SD ( $N$ ) are shown.  $p_{10k}$  = probability by randomization test with 10 000 replicates ( $*p < 0.05$ ; sequential Bonferroni correction, Holm method).

Acoustic variable	Young ( $\leq 8$ years)	Old ( $\geq 9$ years)	$p_{10k}$
Fundamental frequency (Hz)	29.7 $\pm$ 14.0 (233)	27.1 $\pm$ 7.8 (132)	0.0442
Max frequency at $-12$ dB (Hz)	793 $\pm$ 421 (234)	880 $\pm$ 366 (132)	0.0466
12 dB bandwidth (Hz)	687 $\pm$ 442 (234)	780 $\pm$ 383 (132)	0.0415
3 dB bandwidth proportion	0.91 $\pm$ 0.14 (234)	0.88 $\pm$ 0.13 (132)	0.0236
6 dB bandwidth proportion	0.88 $\pm$ 0.14 (234)	0.85 $\pm$ 0.13 (132)	0.0343
F1 (Hz)	273 $\pm$ 40 (220)	259 $\pm$ 28 (130)	0.0003*
F2 (Hz)	619 $\pm$ 130 (220)	549 $\pm$ 83 (130)	0.0001*
F3 (Hz)	962 $\pm$ 187 (220)	799 $\pm$ 118 (129)	0.0001*
F4 (Hz)	1298 $\pm$ 271 (220)	1067 $\pm$ 146 (130)	0.0001*
F5 (Hz)	1619 $\pm$ 308 (219)	1341 $\pm$ 195 (130)	0.0001*
Formant dispersion (Hz)	336 $\pm$ 73 (220)	271 $\pm$ 48 (130)	0.0001*
Minor formant (Hz)	87 $\pm$ 26 (204)	77 $\pm$ 17 (111)	0.0001*
No. of bouts per vocalization	1.9 $\pm$ 1.0 (246)	2.3 $\pm$ 1.1 (156)	0.0001*
Vocalization duration (s)	8.22 $\pm$ 4.78 (246)	10.43 $\pm$ 6.03 (156)	0.0000*
Relative peak time	0.569 $\pm$ 0.164 (246)	0.616 $\pm$ 0.149 (156)	0.0038
Peak intensity predominance	0.821 $\pm$ 0.029 (246)	0.832 $\pm$ 0.041 (156)	0.0013*
Spectrum max/4000 Hz slope	$-0.006 \pm 0.001$ (234)	$-0.007 \pm 0.001$ (132)	0.0110
Spectrum max/ $-24$ dB slope	$-0.036 \pm 0.027$ (234)	$-0.030 \pm 0.021$ (132)	0.0289
No. of syllables per bout	5.75 $\pm$ 3.00 (233)	6.44 $\pm$ 3.34 (132)	0.0460
No. types of syllables per bout	2.17 $\pm$ 0.64 (233)	2.41 $\pm$ 0.67 (132)	0.0008*

**Table 5.** Relationships of acoustic variables to age, for vocalizations of male southern elephant seals. Results of random-regression model analysis are shown for variables that showed trends (see Methods).  $R^2$  = coefficient of determination;  $a$  and  $b$  = model parameters (intercept and slope);  $SE(b)$  = standard error of  $b$  calculated with a jackknife delete-one procedure;  $95\% CI(b)$  =  $95\%$  confidence interval for  $b$ ;  $p_{10k}$  = significance of regression coefficient calculated with randomization;  $p$  = probability of Breusch–Pagan test on significance of random effect (i.e., male identity) (\*  $p < 0.05$ ; sequential Bonferroni correction, Holm method).

Acoustic variable	$R^2$	$a$	$b$	$SE(b)$	95% $CI(b)$	$p_{10k}$	$p$
F1	0.0857	321.751	-6.633	1.12249	-9.804, -5.388	0.0001*	0.0000
F2	0.1320	805.589	-25.719	3.50418	-35.042, -21.258	0.0001*	0.0000
F3	0.2543	1314.608	-49.657	4.86313	-62.534, -43.405	0.0001*	0.0000
F4	0.2774	1828.181	-74.142	7.88737	-95.003, -63.977	0.0001*	0.0000
F5	0.2925	2278.525	-92.987	9.45475	-121.042, -83.851	0.0001*	0.0000
Formant dispersion	0.2775	489.097	-21.567	2.19934	-27.994, -19.343	0.0001*	0.0000
Minor formant	0.0672	108.177	-2.997	0.60713	-4.170, -1.780	0.0001*	0.1877
Fundamental frequency	0.0131	35.100	-0.755	0.31716	-1.503, -0.256	0.0250*	0.0002
No. of bouts per vocalization	0.0564	0.925	0.132	0.02492	0.093, 0.191	0.0001*	0.0000
Vocalization duration	0.0733	2.126	0.825	0.13668	0.688, 1.226	0.0001*	0.0000
Relative peak time	0.0382	0.477	0.013	0.00452	0.002, 0.019	0.0001*	0.0237
No. of types of syllables per bout	0.0596	1.529	0.088	0.01836	0.057, 0.129	0.0001*	0.0000



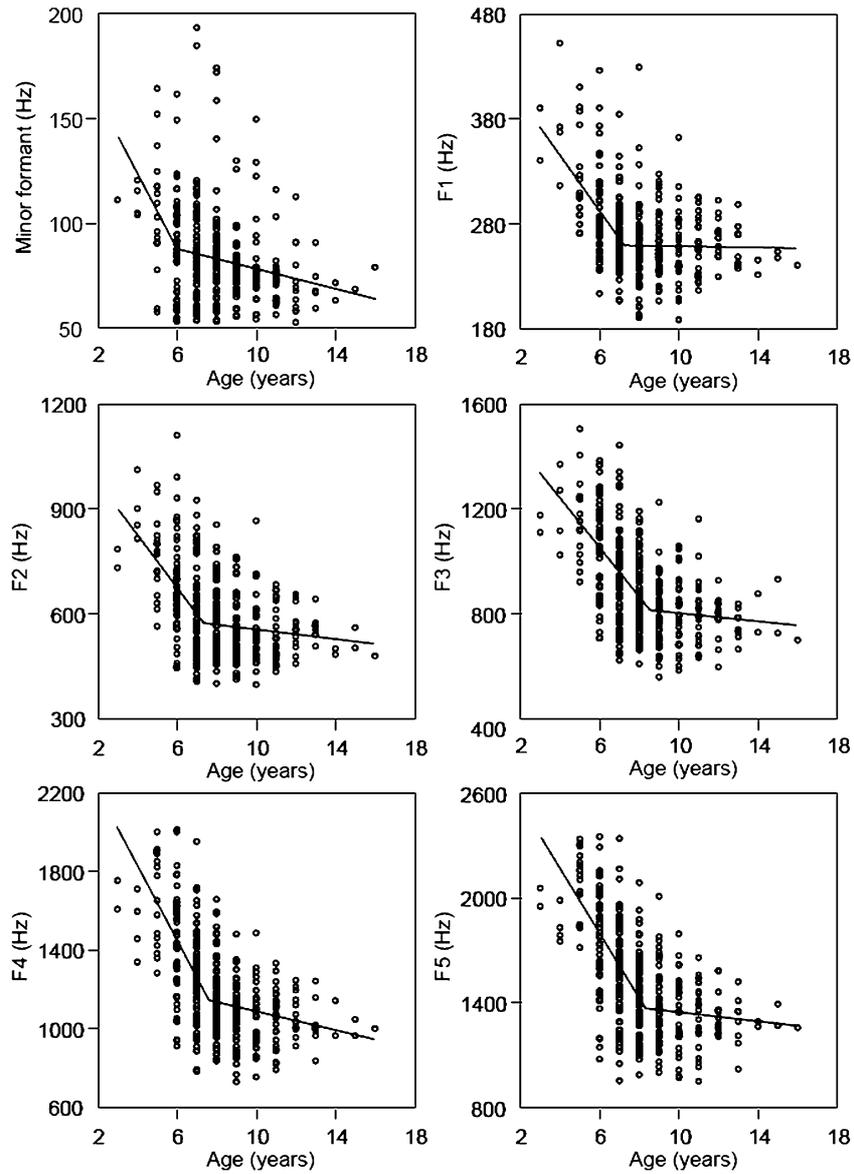
**Figure 4.** Individual variation of acoustic parameters with age. Data are plotted for males of known age recorded over at least three breeding seasons: developmental changes in (A) the fifth formant (F5) and (B) the relative variation in the number of syllables. Different symbols in the figures are for different individual males.

A few temporal and structural features also showed weak trends of change with age, in particular vocalization duration and number of syllable types per bout increased with age, but data dispersion was high, hence the percentage of explained variance was small. In all cases, except for the minor formant, Lagrange multiplier tests revealed significant within-individual effects (Table 5).

All formant frequencies appeared to decrease with age, especially upper formants and formant dispersion. Examination of LOWESS smoother suggested the presence of a threshold effect in relationships of formants to age. Formant frequencies and formant dispersion decreased sharply and almost linearly until approximately 7–9 years of age, after which the rate of decline slowed or stopped. Piecewise regressions with unknown bending point provided better fits than simple linear models (Table 6, Figure 5). In all cases, the regression lines after the bending point were significantly less steep than before. Examination of residuals confirmed the better fit of the piecewise model. In linear analysis, there were excess positive residuals for extreme ages and excess negative ones for middle ages; in piecewise regressions, residuals were homogeneously distributed around zero. No effect of the longitudinal component of data was detected: slopes from random regression models applied before and after the bending point for each variable did not differ significantly from slopes from the linear piecewise model. Percent-

**Table 6.** Formant characteristics changed with age in vocalizations of male southern elephant seals: results of piecewise linear models of formant frequencies versus age (see Methods). The  $t$ -test is for  $H_0: b_{\text{dif}} = 0$ , i.e., significance of change in slope after bending point  $BP$ .  $R^2 =$  coefficient of determination; 95% CI = confidence intervals for  $a_1$  and  $b_1$ ;  $p =$  significance of  $b_1$  ( $* p < 0.05$ ; sequential Bonferroni correction, Holm method).

Acoustic variable	$R^2$	BP	$a_0$	95% CI( $a_0$ )	$b_1$	95% CI( $b_1$ )	$b_{\text{dif}}$	95% CI( $b_{\text{dif}}$ )	$b_2$	$t$	$p$
Minor formant	0.0860	6.00	195.8	126.96, 264.62	-18.04	-20.19, -15.89	15.66	13.13, 18.19	-2.39	12.175	0.0001*
F1	0.1840	7.23	450.9	399.3, 502.5	-26.50	-34.47, -18.53	26.20	17.75, 34.64	-0.31	6.102	0.0001*
F2	0.2010	7.33	1126.7	961.8, 1291.7	-75.51	-100.98, -50.04	68.62	41.64, 95.60	-6.89	5.002	0.0001*
F3	0.3300	8.52	1616.7	1465.2, 1768.3	-94.26	-115.88, -72.64	86.45	58.92, 113.99	-7.81	6.176	0.0001*
F4	0.3800	7.60	2588.1	2274.4, 2901.8	-189.92	-238.37, -141.47	166.10	114.76, 217.43	-23.83	6.364	0.0001*
F5	0.4000	8.30	2908.4	2670.1, 3146.6	-185.71	-219.71, -151.71	172.69	129.44, 215.93	-13.02	7.854	0.0001*
Formant dispersion	0.3720	8.37	626.4	568.3, 684.5	-41.73	-50.02, -33.45	38.58	28.03, 49.13	-3.15	7.193	0.0001*



**Figure 5.** Formant frequencies decrease substantially with age, with an inflection point between 6 and 8 years of age. Piecewise regression lines are shown (see Methods). Points represent individual males.

ages of variance in the formant frequencies explained by age were higher for upper formants (Table 6).

#### *Within-male variation and age*

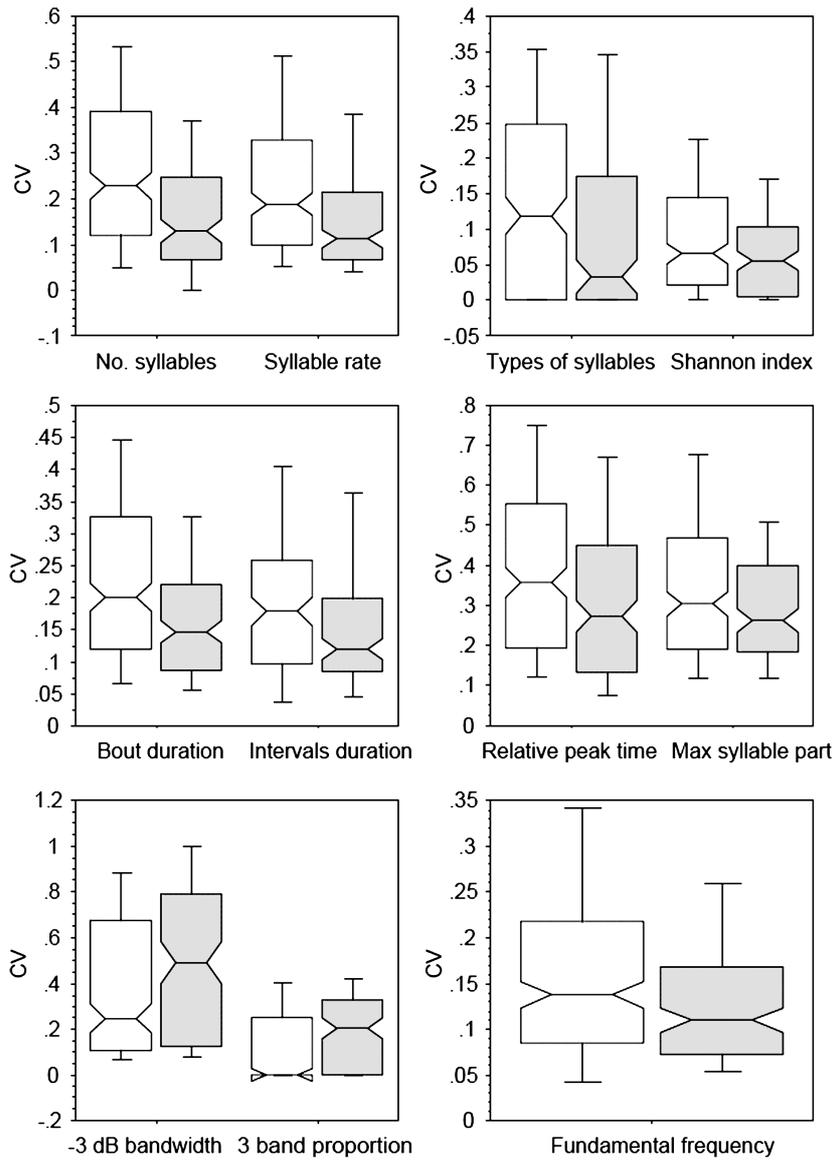
We previously showed that general structure of vocalizations becomes more stereotyped with age (Sanvito et al., 2006a). To determine whether this is true also for detailed acoustic features of vocalizations, we investigated the relationship of within-male variation in acoustic variables to age. ‘Young’ and ‘old’ animals differed significantly in within-male CV for many variables ( $t$ -tests with randomization;  $p_{10k} < 0.05$  in all cases), including macrostructure (syllables, syllable rate, number of syllable types of syllables, Shannon index); temporal features (bout duration, interval between bouts, relative peak time, max syllable part); and only few frequency traits (fundamental frequency,  $-3$  dB bandwidth, 3 bandwidth proportion). No formant frequencies differed significantly in within-male CV between ‘young’ and ‘old’ animals. In all cases, except for  $-3$  dB bandwidth and 3 bandwidth proportion, within-male variation was higher for ‘young’ animals (Figure 6). Overall, temporal and structural variables varied most within males, and intensity and frequency variables varied least.

A random regression model with male identity as the random effect suggested that mean individual CVs decreased with age for different variables, in particular for some temporal and structural variables (Table 7; Figure 4), but the dispersion was very high within each age class, and overall the model fit the data poorly.

On the whole dataset, temporal and structural parameters showed the highest average within-male CVs, in particular for measures of the whole vocalization (vocalization duration and number of bouts, respectively with 0.555 and 0.413), bout duration (0.207) and syllable structure features (0.220 for syllable rate and 0.238 for number of syllables per bout). On the contrary, intensity and frequency features showed the lower within individual variations (0.016 for the peak intensity predominance and 0.024 for the relative peak intensity; 0.069 for F1; 0.150 for dominant frequency and 0.157 for fundamental frequency).

#### *Among-male variation and age*

For most acoustic variables, among-male variation was lower in older animals, and in particular for formant frequencies and several variables related



**Figure 6.** Levels of within-individual variation in acoustic traits differ between ‘young’ (white) and ‘old’ (grey) animals, shown as box plots. The notches of the boxes represent 95% confidence intervals around the median.

**Table 7.** Acoustic within-individual variation declines with age in vocalizations of male southern elephant seals. Mean within-male CVs for different age classes are shown, for variables that showed trends. Sample sizes shown in parentheses.  $p_{10k}$  = significance of regression coefficient (see Methods). (\* $p < 0.05$ ; sequential Bonferroni correction, Holm method).

Acoustic variable	≤5 (18)	6 (55)	7 (94)	8 (78)	9 (60)	≥10 (96)	$p_{10k}$
Bout duration	0.252	0.249	0.237	0.216	0.199	0.148	0.0001*
No. of bouts per vocalization	0.243	0.374	0.389	0.447	0.457	0.429	0.0257
Interval between bouts duration	0.251	0.227	0.182	0.197	0.205	0.135	0.0019*
Relative peak time	0.459	0.477	0.426	0.329	0.418	0.274	0.0002*
Syllable part with max peak pressure duration	0.380	0.369	0.361	0.336	0.332	0.286	0.0164
No. of syllables per bout	0.455	0.301	0.268	0.223	0.208	0.160	0.0001*
Syllable rate	0.363	0.271	0.250	0.197	0.216	0.150	0.0001*
No. of types of syllables per bout	0.233	0.154	0.142	0.116	0.107	0.102	0.0001*
Shannon index	0.183	0.105	0.093	0.068	0.069	0.071	0.0002*
Spectrum max amplitude/4000 Hz slope	-0.159	-0.160	-0.160	-0.153	-0.151	-0.132	0.0278
Fundamental frequency	0.176	0.198	0.167	0.152	0.137	0.135	0.0030*

to -6, -12 and -18 dB bandwidths (Table 8). The among-male variation of only one temporal feature (relative peak intensity) differed between 'young' and 'old' seals. The other simple temporal (e.g., bout duration) or structural features (related to syllable structure) had homogenous among-male variation in the two age groups.

## Discussion

### *Development of vocal behaviour*

Vocalizations are a crucial component of male agonistic behaviour of southern elephant seals, being present in most interactions between males for access to breeding females (Sandegren, 1976; McCann, 1981; Sanvito et al., 2006a). SES males at Sea Lion Island showed a clear increase with age in the involvement in social interactions, in the use of ritualized forms of agonistic

**Table 8.** Among-male variation in acoustic features differed between ‘young’ and ‘old’ male southern elephant seals. Statistics are mean  $\pm$  SE of CVs ( $N$ ).  $p_{10k}$  = probability of randomization test on difference between the mean CVs (Sokal & Braumann, 1980). (\* $p < 0.05$ ; sequential Bonferroni correction, Holm method).

Acoustic variable	Young ( $\leq 8$ years)	Old ( $\geq 9$ years)	$p_{10k}$
Min frequency at $-6$ dB	$0.372 \pm 0.027$ (234)	$0.281 \pm 0.019$ (132)	0.0426
Min frequency at $-12$ dB	$0.561 \pm 0.046$ (234)	$0.438 \pm 0.033$ (132)	0.0041*
18 dB bandwidth proportion	$0.124 \pm 0.008$ (234)	$0.095 \pm 0.006$ (132)	0.0116
F1	$0.146 \pm 0.01$ (220)	$0.110 \pm 0.007$ (130)	0.0375
F2	$0.200 \pm 0.014$ (220)	$0.152 \pm 0.01$ (130)	0.005
F3	$0.188 \pm 0.013$ (220)	$0.149 \pm 0.01$ (130)	0.0251
F4	$0.202 \pm 0.014$ (220)	$0.140 \pm 0.01$ (130)	0.0018*
F5	$0.178 \pm 0.012$ (220)	$0.145 \pm 0.01$ (130)	0.0391
Minor formant	$0.322 \pm 0.027$ (204)	$0.230 \pm 0.019$ (111)	0.0499
Spectrum 0/4000 Hz slope	$-0.185 \pm 0.012$ (234)	$-0.136 \pm 0.009$ (132)	0.0068
Spectrum max/4000 Hz slope	$-0.193 \pm 0.013$ (234)	$-0.159 \pm 0.01$ (132)	0.0485
Relative peak time	$0.299 \pm 0.018$ (246)	$0.242 \pm 0.015$ (156)	0.0264
Peak intensity predominance	$0.039 \pm 0.002$ (246)	$0.049 \pm 0.003$ (156)	0.0392

behaviour, in the use of vocalizations and in the effectiveness of its use in the resolution of agonistic contests. Competition between elephant seal males is intense and costly, in term of energetic expenditure and physical damage. The absolute energetic cost of competition of elephant seals is the highest observed in vertebrates (Deutsch, 1990; Galimberti et al., in press). The cost due to wounds and physical damage is more difficult to assess (Geist, 1974). In elephant seal males, wounding due to fights is frequent (Deutsch, 1990; Braschi, 2004), and even non-lethal wounds may affect the breeding success of males, reducing their capability to compete (e.g., due to infections). The tendency towards the ritualization of agonistic behaviour and the increase in the use and effectiveness of vocalizations can reduce both these costs.

#### *Development of acoustic features of vocalizations*

Vocal ontogeny can result from two main processes, vocal learning of acoustic features and maturation in the structural phenotype that affects the structures used to produce the sounds. Vocal learning should be focused on acoustic features that are not strictly and directly constrained by structural

phenotype, whereas maturation should be prevalent for acoustic features that depend on growth.

The source-filter model of sound production, originally proposed and developed for human speech (Fant, 1960), was recently extended to all mammal vocalizations (Fitch, 1994, 2000). According to this model, most mammal vocalizations are produced by the larynx (the periodic or quasi-periodic source) and then filtered by the vocal tract (the filter), before being emitted through the nostrils and lips. The rate of oscillation of the source is responsible for the pitch of the sound (i.e., its fundamental frequency,  $F_0$ ). The resonances of the vocal tract act as a filter on the original sound, selectively amplifying some frequencies by damping others, thus producing what are known as formant frequencies, i.e., frequencies in the amplitude spectrum of a sound that seem to be reinforced by resonance of the vocal tract (Fry, 1979; Fitch, 1994; Miller & Murray, 1995). Overall, formant frequencies and their spacing should decrease with increase in vocal tract length (Fry, 1979; Fitch, 1994), whereas  $F_0$  should decrease with increasing size and mass and decreasing tension of the vocal folds. Even though mammals seem able to change their vocal fold tension and length over a quite large range by the action of the laryngeal muscles (hence changing their  $F_0$ ), their vocal tract length depends on the structural phenotype, being constrained by the bones of the skull (Fitch, 2000; Fitch & Hauser, 2002). Therefore, the formants are expected to be strictly dependent on the animal's morphology, whereas temporal and structural features of the vocalizations are expected to be almost free from physical constraint, hence can be the target of vocal learning.

#### *Variation in the mean of acoustic features*

In our study, formants showed the strongest trend of change with age, with a linear decrease in frequency up to age 7–8, and increasing strength of the trend for upper formants. Older males had lower frequency of formants than younger males, although the relationship of formant frequencies to age was not homogeneous across the whole age span, as shown by the better fit obtained using piecewise regression models. Formants are rather strictly related to the structural phenotype of the male, because they are determined and constrained by the size and shape of the vocal tract. In contrast, temporal features and the syllable structure of the sounds showed no clear trend of change with age, and even the differences between 'young' and 'old' males were blurred.

Moreover, the frequency features that are not related to the vocal tract length (e.g.,  $F_0$ ) showed no trend. All these findings are in accordance with our initial hypotheses about the role of vocal learning and physical maturation in the development of different acoustic features. Contrary to these expectations, in some species (yellow baboon, *Papio cynocephalus*: Fischer et al., 2004) temporal components of the vocalizations, and in particular the duration and repetition rate of calls, increase with age. In these species, sound emission seems to be a significant cost for males, and older males seem to be able to sustain this cost. On the contrary, in SES males, the cost of vocalizations is likely to be negligible if compared to the whole breeding effort (Deutsch, 1990; Galimberti et al., in press) and, therefore, a relationship between temporal aspects of vocalizations and age due to a cost reduction strategy of younger males is not expected.

#### *Variability of acoustic features*

The interest in variability of behavioural patterns dates back to the very beginning of ethology (Barlow, 1977), and the coefficient of variation has been frequently used as a measure of stereotypy (Bekoff, 1977). Change in variability of acoustic features is another important cue to discriminate between vocal learning and physical maturation.

In our study, the within-male variability of formants, which are the most structurally constrained of all acoustic features, showed no change with age and no increase in stereotypy. These features depend directly on emitter morphology. On the contrary, most acoustic features related to the time domain and all features related to the syllable structure of the bouts showed a decrease with age in the within-male variability (i.e., an increase in stereotypy), which is well explained by a learning process, by which males increase their vocal competence. Moreover, we expected the whole variability, pooled across ages, of constrained acoustic features to be lower than other features. In accordance with this hypothesis, the relative variability of frequency features was lower than the variability of structural and temporal features, as in other species (Miller, 1991; Insley, 1992).

As for within-male variation, we expected structurally constrained features to mimic the variation in structural phenotype, hence to display greater variability among young growing males, which vary greatly in size at each age (Clinton, 1994). Indeed, CVs for formants was significantly greater

in 'young' than 'old' males. On the contrary, features that are not constrained by structural phenotype should show no systematic difference between 'young' and 'old' males. In accordance with this hypothesis, there was no significant difference between CVs of 'young' and 'old' males in temporal or structural acoustic features.

Overall, age-specific variation of mean and variability (within and among-individuals) of acoustic features is in accordance with the presence of two simultaneous processes, maturation and vocal learning. The first one affects the frequency component of sounds that depend on the vocal tract size and shape, and the second one works on temporal and structural features of sounds that are rather free to vary unconstrained by the individual phenotype.

#### *Development, honest signalling and individual recognition*

The relationship between formants and age confirms that the formants are constrained by the structural phenotype and can, in principle, be honest signals of age (Fitch & Hauser, 2002). Another acoustic feature of SES male vocalizations, the source level, has been shown to be related to the structural phenotype and, hence, to be a good candidate as honest signal (Sanvito & Galimberti, 2003). The relationship between formants and age was only moderately strong, and the percentage of variance explained by the relationship was rather low, around 40% for the highest formant. These results are in line with some recent findings obtained in another polygynous mammal, with a similar mating system, the red deer (*Cervus elaphus*, Reby & McComb, 2003).

We have shown previously (Sanvito et al., 2006a) that young males tend to mimic and adopt the vocal type of older, more successful males. This result is confirmed by the development trend found for the temporal and structural features of their vocalizations. Young elephant seals seem to be able to produce the same pulse trains that are the building blocks of the vocalizations of mature males, but their vocalizations are initially extremely variable and not structured (Sanvito et al., 2006a). The increase in structuring of the vocalizations is likely to be related to an increase in competence to assemble the pulse trains in a structured, complex vocalization (Shipley et al., 1986), i.e., a vocal learning process. The result is the emission by adults of strongly structured vocalizations, specific for each individual and stereotyped, and different between individuals, with no clear constraint due

to the structural phenotype. The reduction in intra-individual variability of the acoustic features that are not constrained by structural phenotype, which is likely a product of vocal learning, may improve individual recognition (Insley, 1992; Insley et al., 2003).

### Acknowledgements

We thank: M. and R. Sanvito, C. and A. Galimberti for their long-lasting support of our research; Dr. H. Paddock for enlightening discussions about similarities between human and seal 'speech'; the Falkland Islands Government for permission to conduct field research in the Falkland Islands; the Falkland Islands Development Corporation for permission to conduct field work on Sea Lion Island; D. Gray, J. Luxton and Strachan Visick Ltd for help with field logistics; and the many people who helped us in the field — too many to be named here. Our research at Sea Lion Island was funded in part by Earthwatch Institute grants and Strachan Visick Ltd. Support for the first author was provided by Memorial University of Newfoundland.

### References

- Aitkin, L., Nelson, J. & Shepherd, R. (1996). Development of hearing and vocalization in a marsupial, the northern quoll, *Dasyurus hallucatus*. — *J. Exp. Zool.* 276: 394-402.
- Bailey, R.C. & Byrnes, J. (1990). A new, old method for assessing measurements error in both univariate and multivariate morphometric studies. — *Syst. Zool.* 39: 124-130.
- Barlow, G.W. (1977). Modal action patterns. — In: *How animals communicate* (Sebeok, T.A., ed.). Indiana University Press, Bloomington, IN, p. 98-134.
- Bartholomew, G.A. & Collias, N.E. (1962). The role of vocalization in the social behaviour of the northern elephant seal. — *Anim. Behav.* 10: 7-14.
- Bekoff, M. (1977). Mammalian dispersal and the ontogeny of individual behavioural phenotypes. — *Am. Nat.* 111: 715-732.
- Braschi, C. (2004). Confronti agonistici nell'elefante marino del sud: fattori coinvolti e ruolo delle asimmetrie tra due maschi (Male agonistic behavior in southern elephant seals: factors and asymmetries involved in conflict). — PhD thesis, Università La Sapienza di Roma, Rome.
- Burnham, K.P. & Anderson, D.R., eds (1998). *Model selection and inference: a practical information-theoretic approach*. — Springer, New York, NY.
- Cameron, A.C. & Trivedi, P.K., eds (1998). *Regression analysis of count data*. — Cambridge University Press, Cambridge.
- Chaadaeva, E.V. (2002). Postnatal ontogeny of vocalization in the racoon dog, *Nyctereutes procyonoides* (Carnivora, Canidae). — *Zool. Zh.* 81: 234-241.
- Chaadaeva, E.V. & Sokolova, N.N. (2005). Development of vocal repertoire in kittens of *Felis libyca* and *F. catus* (Carnivora, Felidae). — *Zool. Zh.* 84: 1402-1415.
- Charif, R.A., Mitchell, S. & Clark, C.W. (1995). *Canary 1.2 user's manual*. — Cornell Laboratory of Ornithology, Ithaca, NY.

- Cheung, Y.B. (2002). Zero-inflated models for regression analysis of count data: a study of growth and development. — *Stat. Med.* 21: 1461-1469.
- Clinton, W.L. (1994). Sexual selection and growth in male Northern Elephant Seals. — In: Elephant seals. Population ecology, behavior, and physiology (Le Boeuf, B.J. & Laws, R.M., eds). University of California Press, Berkeley, CA, p. 154-168.
- Cox, C.R. (1981). Agonistic encounters among male elephant seals: frequency, context, and the role of female preference. — *Am. Zool.* 21: 197-209.
- Darden, S.K., Pedersen, S.B. & Dabelsteen, T. (2003). Methods of frequency analysis of a complex mammalian vocalisation. — *Bioacoustics* 13: 247-263.
- Deutsch, C.J. (1990). Behavioral and energetic aspects of reproductive effort of male northern elephant seals (*Mirounga angustirostris*). — PhD thesis, University of California, Berkeley, CA.
- Diggle, P.J., Kung-Yee, L. & Zeger, S.L., eds (1994). Analysis of longitudinal data. — Clarendon Press, Oxford.
- Egnor, R.S.E. & Hauser, M.D. (2004). A paradox in the evolution of primate vocal learning. — *Trends Neurosci.* 27: 649-654.
- Fant, G., ed. (1960). Acoustic theory of speech production. — Mouton, The Hague.
- Fischer, J., Kitchen, D.M., Seyfarth, R.M. & Cheney, D.L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. — *Behav. Ecol. Sociobiol.* 56: 140-148.
- Fitch, W.T. (1994). Vocal tract length perception and the evolution of language. — PhD thesis, Brown University, Providence, RI.
- 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 phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. — *Phonetica* 57: 205-218.
- Fitch, W.T. & Hauser, M.D. (2002). Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals. — In: Acoustic communication (Simmons, A.M., Popper, A.N. & Fay, R.R., eds). Springer, New York, NY.
- Fitch, W.T. & Reby, D. (2001). The descended larynx is not uniquely human. — *Proc. Roy. Soc. Lond. B Biol.* 268: 1669-1675.
- Fry, D.B. (1979). The physics of speech. — Cambridge University Press, Cambridge.
- Galimberti, F. & Boitani, L. (1999). Demography and breeding biology of a small, localized population of southern elephant seals (*Mirounga leonina*). — *Mar. Mammal Sci.* 15: 159-178.
- Galimberti, F., Boitani, L. & Marzetti, I. (2000). The frequency and costs of harassment in southern elephant seals. — *Ethol. Ecol. Evol.* 12: 345-365.
- Galimberti, F., Fabiani, A. & Boitani, L. (2003). Socio-spatial levels in linearity analysis of dominance hierarchies: a case study on elephant seals. — *J. Ethol.* 21: 131-136.
- Galimberti, F., Fabiani, A. & Sanvito, S. (2002). Measures of breeding inequality: a case study in southern elephant seals. — *Can. J. Zool.* 80: 1240-1249.
- Galimberti, F. & Sanvito, S. (2001). Modeling female haul out in southern elephant seals (*Mirounga leonina*). — *Aquat. Mammal.* 27: 92-104.
- Galimberti, F., Sanvito, S., Braschi, C. & Boitani, L. (in press). The cost of success: reproductive effort in male southern elephant seals. — *Behav. Ecol. Sociobiol.*
- Gardner, W., Mulvey, E.P. & Shaw, E.C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. — *Psychol. Bull.* 118: 392-404.

- Gautier, J.P. (1998). Vocal communication in primates. — *Bull. Soc. Zool. France* 123: 239-253.
- Geist, V. (1974). On fighting strategies in animal combat. — *Nature* 250: 354.
- Hochberg, Y. & Benjamini, Y. (1990). More powerful procedures for multiple significance testing. — *Stat. Med.* 9: 811-818.
- Insley, S.J. (1992). Mother-offspring separation and acoustic stereotypy: A comparison of call morphology in two species of pinnipeds. — *Behaviour* 120: 103-122.
- Insley, S.J., Phillips, A.V. & Charrier, I. (2003). A review of social recognition in pinnipeds. — *Aquat. Mammal.* 29: 181-201.
- Job, D.A., Boness, D.J. & Francis, J.M. (1995). Individual variation in nursing vocalizations of Hawaiian monk seal pups, *Monachus schauinslandi* (Phocidae, Pinnipedia), and lack of maternal recognition. — *Can. J. Zool.* 73: 975-983.
- Khan, C.B., Markowitz, H. & McCowan, B. (2006). Vocal development in captive harbor seal pups, *Phoca vitulina richardii*: Age, sex, and individual differences. — *J. Acoust. Soc. Am.* 120: 1684-1694.
- Killebrew, D.A., Mercado, E., Herman, L.M. & Pack, A.A. (2001). Sound production of a neonate bottlenose dolphin. — *Aquat. Mammal.* 27: 34-44.
- Krebs, C.J. (1989). *Ecological methodology*. — Harper & Row, New York, NY.
- Laws, R.M. (1953). The elephant seal (*Mirounga leonina* Linn.). I. Growth and age. — *Sci. Rep. Falkland Is. Depend. Surv.* 8: 1-62.
- Laws, R.M. (1956). The elephant seal (*Mirounga leonina* Linn.). II. General, social, and reproductive behaviour. — *Sci. Rep. Falkland Is. Depend. Surv.* 13: 1-88.
- Lessells, C.M. & Boag, P.T. (1987). Unrepeatable repeatabilities: a common mistake. — *Auk* 104: 116-121.
- Long, J.S. (1997). *Regression models for categorical and limited dependent variables*. — Sage, Thousand Oaks, CA.
- Long, J.S. & Freese, J., eds (2001). *Regression model for categorical dependent variables using STATA*. — STATA, College Station, TX.
- Mandelli, M.J. & Sales, G. (2004). Ultrasonic vocalizations of infant short-tailed field voles, *Microtus agrestis*. — *J. Mammal.* 85: 282-289.
- Manly, B.F.J. (1997). *Randomization, bootstrap and Monte Carlo methods in biology*. — Chapman & Hall, London.
- McCann, T.S. (1981). Aggression and sexual activity of male Southern elephant seals, *Mirounga leonina*. — *J. Zool. Lond.* 195: 295-310.
- Miller, E.H. (1991). Communication in pinnipeds, with special reference to non-acoustic signalling. — In: *The behaviour of the Pinnipeds* (Renouf, D., ed.). Chapman & Hall, London, p. 128-235.
- Miller, E.H. & Murray, A.V. (1995). Structure, complexity, and organization of vocalizations in harp seals (*Phoca groenlandica*). — In: *Sensory systems of aquatic mammals* (Kastelein, R.A., Thomas, J.A. & Nachtigall, P.E., eds). De Spil, Woerden, p. 237-264.
- Mitchell-Olds, T. & Shaw, R.G. (1987). Regression analysis of natural selection: statistical inference and biological interpretation. — *Evolution* 41: 1149-1161.
- Monsen, R.B. & Engebretson, A. (1983). The accuracy of formant frequency measurements: a comparison of spectrographic analysis and linear prediction. — *J. Speech Hear. Res.* 26: 89-97.
- Muggeo, V.M. (2003). Estimating regression models with unknown break-points. — *Stat. Med.* 22: 3055-3071.

- Owren, M. & Linker, C. (1995). Some analysis methods that may be useful to acoustic primatologists. — In: Current topics in primate vocal communication (Elke, Z., ed.). Plenum Press, New York, NY, p. 1-28.
- Owren, M.J. & Bernacki, R.H. (1988). The acoustic features of vervet monkey alarm calls. — *J. Acoust. Soc. Am.* 83: 1927-1935.
- Phillips, A.V. & Stirling, I. (2000). Vocal individuality in mother and pup South American fur seals, *Arctocephalus australis*. — *Mar. Mammal Sci.* 16: 592-616.
- Pistorio, A.L., Vintch, B. & Wang, X. (2006). Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*). — *J. Acoust. Soc. Am.* 120: 1655-1670.
- Reby, D. & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. — *Anim. Behav.* 65: 519-530.
- Riede, T. & Fitch, T. (1999). Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). — *J. Exp. Biol.* 202: 2859-2867.
- Sandegren, F.E. (1976). Agonistic behavior in the male Northern Elephant seal. — *Behaviour* 57: 136-158.
- Sanvito, S. & Galimberti, F. (2000). Bioacoustics of southern elephant seals. I. Acoustic structure of male aggressive vocalizations. — *Bioacoustics* 10: 259-285.
- Sanvito, S. & Galimberti, F. (2003). Source level of male vocalizations in the genus *Mirounga*: repeatability and correlates. — *Bioacoustics* 14: 45-57.
- Sanvito, S., Galimberti, F. & Miller, E.H. (2006a). Observational evidences of vocal learning in southern elephant seals: a longitudinal study. — *Ethology* 113: 137-146.
- Sanvito, S., Galimberti, F. & Miller, E.H. (2006b). Vocal signalling of male southern elephant seals is honest but imprecise. — *Anim. Behav.* 73: 287-299.
- Schassburger, R.M. (1993). Vocal communication in the timber wolf, *Canis lupus*, Linnaeus. Structure, motivation, and ontogeny. — *Adv. Ethol.* 30: 1-84.
- Seyfarth, R.M. & Cheney, D.L. (1986). Vocal development in vervet monkeys. — *Anim. Behav.* 34: 1640-1658.
- Shipley, C., Hines, M. & Buchwald, J.S. (1986). Vocalizations of northern elephant seal bulls: development of adult call characteristics during puberty. — *J. Mammal.* 67: 526-536.
- Siegel, S. & Castellan, N.J. (1988). Nonparametric statistics for the behavioral sciences. — McGraw Hill, New York, NY.
- Snowdon, C.T. & Hausberger, M. (1997). Social influences on vocal development. — Cambridge University Press, Cambridge.
- Sokal, R.R. & Braumann, C.A. (1980). Significance tests for coefficients of variation and variability profiles. — *Syst. Zool.* 29: 50-66.
- Trexler, J.C. & Travis, J. (1993). Nontraditional regression analysis. — *Ecology* 74: 1629-1637.
- Tyack, P.L. (1997). Development and social functions of signature whistles in bottlenose dolphins *Tursiops truncatus*. — *Bioacoustics* 8: 21-46.
- Van Parijs, S.M. & Corkeron, P.J. (2002). Ontogeny of vocalizations in infant black flying foxes, *Pteropus alecto*. — *Behaviour* 139: 1111-1124.
- Vater, M., Kossel, M., Foeller, E., Coro, F., Mora, E. & Russel, I.J. (2003). Development of echolocation calls in the mustached bat, *Pteronotus parnelli*. — *J. Neurophysiol.* 90: 2274-2290.
- Williams, R.L. (2000). A note on robust variance estimation for cluster-correlated data. — *Biometrics* 56: 645-646.

Zuberbuhler, K., Noé, R. & Seyfarth, R.M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. — *Anim. Behav.* 53: 589-604.

## Appendix

**Table A1.** Average values of frequency domain parameters for different age classes. Sample size in parentheses below the column headings.

Acoustic parameter	≤5 (18)	6 (55)	7 (95)	8 (78)	9 (60)	≥10 (96)
Dominant frequency (Hz)	375	342	297	301	310	321
Fundamental frequency (Hz)	31.2	30.3	29.4	29.1	28.0	26.6
Peak frequency (Hz)	402	387	308	329	319	349
3 dB bandwidth (Hz)	187	192	134	196	168	211
6 dB bandwidth (Hz)	331	354	252	340	305	385
12 dB bandwidth (Hz)	766	797	560	733	756	796
18 dB bandwidth (Hz)	1362	1342	1059	1345	1373	1357
Min frequency at -3 dB (Hz)	292	236	229	223	236	235
Min frequency at -6 dB (Hz)	243	187	187	180	189	183
Min frequency at -12 dB (Hz)	146	99	107	99	102	98
Min frequency at -18 dB (Hz)	70	42	50	49	46	43
Max frequency at -3 dB (Hz)	478	427	363	419	404	446
Max frequency at -6 dB (Hz)	574	541	439	520	495	568
Max frequency at -12 dB (Hz)	912	897	667	833	858	894
Max frequency at -18 dB (Hz)	1433	1384	1109	1394	1419	1400
3 dB bandwidth proportion	0.95	0.90	0.94	0.87	0.90	0.86
6 dB bandwidth proportion	0.92	0.86	0.91	0.85	0.87	0.83
12 dB bandwidth proportion	0.85	0.82	0.87	0.86	0.83	0.86
18 dB bandwidth proportion	0.86	0.88	0.85	0.84	0.85	0.88
F1 (Hz)	322	289	266	259	259	258
F2 (Hz)	759	668	600	575	554	546
F3 (Hz)	1168	1050	945	872	810	792
F4 (Hz)	1667	1431	1264	1159	1086	1055
F5 (Hz)	2086	1757	1593	1444	1364	1326
Formant dispersion (Hz)	441	367	332	297	276	268
Minor formant (Hz)	106	88	85	85	78	76

**Table A2.** Average values of temporal, intensity and structural parameters for different age classes. Sample sizes in parentheses below the column headings.

Acoustic parameter	≤5 (18)	6 (55)	7 (95)	8 (78)	9 (60)	≥10 (96)
No. of bouts per vocalization	1.4	1.6	1.9	2.2	2.1	2.4
Bout duration (s)	3.39	3.64	3.57	3.84	3.69	3.48
Interval between bouts (s)	2.20	2.15	2.11	2.21	2.16	2.04
Vocalization duration (s)	5.29	6.96	8.20	9.81	9.45	11.04
Duty cycle in bout	0.671	0.727	0.724	0.726	0.732	0.730
Relative peak time	0.513	0.543	0.567	0.601	0.582	0.637
Max syllable part duration (s)	0.63	0.50	0.51	0.60	0.51	0.61
Inst. relative peak intensity	0.75	0.75	0.76	0.76	0.76	0.75
Peak intensity predominance	0.82	0.81	0.83	0.82	0.83	0.84
Spectrum total slope	-0.00184	-0.00193	-0.00185	-0.00189	-0.00192	-0.00192
Spectrum 0/4000 Hz slope	-0.00658	-0.00635	-0.00678	-0.00649	-0.00663	-0.00686
Spectrum ascending slope	0.06174	0.06183	0.07511	0.07454	0.07005	0.06752
Spectrum max/4000 Hz slope	-0.00668	-0.00612	-0.00637	-0.00628	-0.00645	-0.00674
Spectrum max/-24 dB slope	-0.03817	-0.03216	-0.03810	-0.03435	-0.03053	-0.02882
No. of syllable per bout	5.12	5.44	5.98	5.87	6.35	6.50
Syllable rate	1.76	1.68	1.78	1.74	1.91	1.95
No. of types of syllables per bout	1.90	2.07	2.20	2.28	2.37	2.43
Shannon index	0.817	0.814	0.816	0.822	0.801	0.776