Printed by Universitetsservice AB

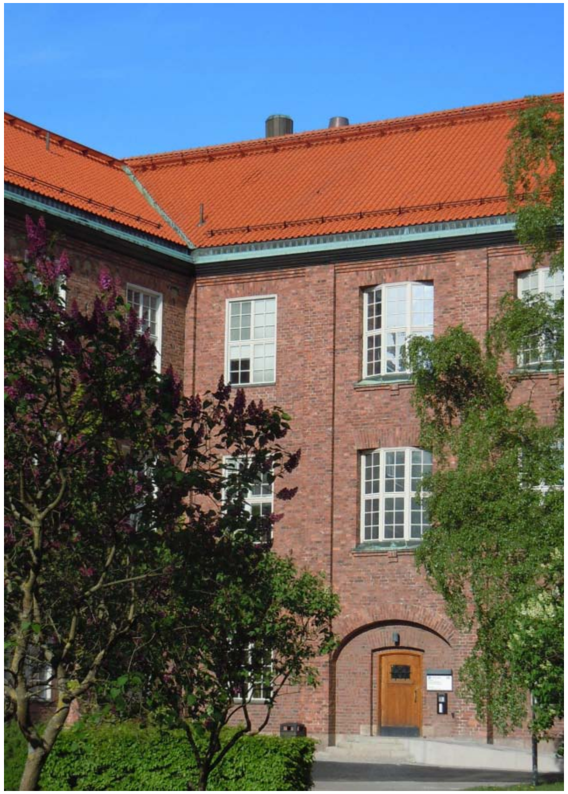
Stockholm 2011

Fonetik 2011

KTH, Speech, Music and Hearing

TMH-QPSR Vol. 51

Fonetik 2011



Speech, Music and Hearing
TMH-QPSR Vol. 51

KTH Computer Science
and Communication

Speech, Music and Hearing

Quarterly Progress and Status Report

TMH-QPSR, Volume 51, 2011

Proceedings from

Fonetik 2011

June 8 – June 10, 2011

Organized by
Department of Speech, Music and Hearing,
and Centre for Speech Technology, CTT,
KTH, Stockholm

Organizing Committee:

Björn Granström
David House
Daniel Neiberg
Sofia Strömbergsson

Front and back cover:
The meeting venue at KTH
Photos by Björn Granström

Published by Department of Speech, Music and Hearing, KTH, Stockholm

TRITA-CSC-TMH 2011:1
ISSN 1104-5787
ISRN KTH/CSC/TMH--11/01--SE

Printed by Universitetsservice AB, Stockholm
2011

Table of Contents

Author index

Preface

This volume of QPSR, the 51th in a long series of KTH publications, contains the 31 contributions to Fonetik 2011, the annual Swedish Phonetics Conference. It has been organised since the mid 1980's by different university departments involved in phonetics. This time the Department of Speech, Music and Hearing at KTH hosted the conference which was held on June 8 – June 10, 2011 at KTH

The conference was attended by close to 75 participants, mainly from Sweden and the other Nordic countries. Fonetik 2011 displays a variety of topics reflecting the wide range of activities in this field.

We thank all the contributors for their co-operative work to make this volume available in time for the conference. The conference activities and the printing of this volume were economically supported by Fonetikstiftelsen (the Swedish Phonetics Foundation) and by CSC (the School of Computer Science and Communication) at KTH, which we gratefully acknowledge.

The contributions in this volume are also published on the web, as are the previous 50 QPSR volumes - http://www.speech.kth.se/qpsr/

The Fonetik 2011 organisers

Björn Granström *David House* *Daniel Neiberg* *Sofia Strömbergsson*

Previous Swedish Phonetics Conferences (from 1986):

1986	Uppsala University
1988	Lund University
1989	KTH Stockholm
1990	Umeå University (Lövånger)
1991	Stockholm University
1992	Chalmers/Göteborg University
1993	Uppsala University
1994	Lund University (Höör)
1995	(XIIIth ICPhS in Stockholm)
1996	KTH Stockholm (Nässlingen)
1997	Umeå University
1998	Stockholm University
1999	Göteborg University
2000	University of Skövde
2001	Lund University (Örenäs)
2002	KTH Stockholm
2003	Umeå University
2004	Stockholm University
2005	Göteborg University
2006	Lund University
2007	KTH Stockholm
2008	Göteborg University
2009	Stockholm University
2010	Lund University

There is a web-page maintained by H. Traunmüller with links to all previous phonetics conferences at:
http://www2.ling.su.se/fon/fonkonfer.html

An acoustic analysis of lion roars. I: Data collection and spectrogram and waveform analyses

Robert Eklund [1,2,3], Gustav Peters [4], Gopal Ananthakrishnan [5] & Evans Mabiza [6]
[1] Voice Provider, Stockholm, Sweden
[2] Department of Cognitive Neuroscience, Karolinska Institute, Stockholm, Sweden
[3] Department of Computer Science, Linköping University, Linköping, Sweden
[4] Forschungsinstitut Alexander Koenig, Bonn, Germany
[5] Centre for Speech Technology, Royal Institute of Technology, Stockholm, Sweden
[6] Antelope Park, Gweru, Zimbabwe

Abstract

This paper describes the collection of lion roar data at two different locations, an outdoor setting at Antelope Park in Zimbabwe and an indoor setting at Parken Zoo in Sweden. Preliminary analyses of spectrographic and waveform data are provided.

Introduction

Felids are one of the most successful carnivore families ever to exist, and within the 35–40 different cat species that exist today several different vocalizations can be found, with different functions, ranging from the well-known purring to the most impressive sound of them all: roaring of lion (*Panthera leo*) fame. This paper focuses on the impressive lion roaring, and highlights methodological problems associated with the collection of animal vocalizations data.

Roaring: a primer

For a human observer the roaring of a lion – even more so that of a whole pride – certainly is one of the most impressive vocalizations in the animal kingdom. In its complete form lion roaring is a species-specific series of calls with a fairly regular structure of the single calls composing it and the series itself, in the latter in terms of the sequence of call types, their change of intensity in the course of the series, the temporal sequencing of the calls and their relative duration and that of the intervals between them. A typical lion roaring can last for more than a minute, usually starting off with a few low-intensity moan-like calls, then progressively increasing in intensity and duration of the calls, and in approaching the intensity climax of the series the calls become shorter again and harsher. After the climax follows a series of short harsh calls, in the beginning uttered at fairly monotonic intensity and brief intervals between the calls, then towards the end of the series gradually decreasing in intensity and with increasing interval duration (called “outro” in this paper).

Given the fact that the colloquial term ‘roar’ is commonly used for various intense animal vocalizations it is not surprising that even in the lion it has been applied to vocalizations which are definitely different from roaring as dealt with here. Early attempts at characterizing it in a more technical manner were published by e.g. Leyhausen (1950), Hemmer (1966) and Schaller (1972). More recent studies of lion roaring include Peters (1978), Peters & Hast (1994), and Pfefferle et al. (2007) .

Weissengruber et al. (2002:208) extended the definition of roaring in a general vertebrate vocalization context suggesting that lion roaring “has two distinct physiological and acoustic components:

1 a low fundamental frequency, made possible by long or heavy vocal folds, which lead to the low pitch of the roar;

2 lowered formant frequencies, made possible by an elongated vocal tract, which provide the impressive baritone timbre of roars.”

(See also Frey & Gebler, 2010).

In this paper, we studied lion roaring ‘proper’ as outlined at the start, in respect of the fine acoustic structure of its component single calls, the structural changes they undergo in the course of the roaring series and possible physiological mechanisms underlying these changes, considering the definition suggested by Weissengruber et al. (2007).

On the function of roaring

The function of lion roars has been discussed extensively in the literature, and several hypotheses have been suggested. Pfefferle et al. (2007:3952) concluded that the “primary function of roars is the advertisement and defense of territory”. In support of this hypothesis, it has been shown that lionesses can estimate the number of individuals roaring, and that they are less likely to approach foreign roars when they are outnumbered (McComb, Packer & Pusey, 1994).

Besides territorial defense, an additional function might also be coordination of hunting (Grinnell & McComb, 2001; McComb, Packer & Pusey, 1994; Schaller, 1972).

Method

The following sections describe the data collection, data processing and analysis tools.

Data collection

The data analyzed in this paper were recorded at two different locations, one outdoor and one indoor setting. Recording details are given below.

Antelope Park, Gweru, Zimbabwe

The first set of lion roar recordings was obtained at Antelope Park lion rehabilitation and release into the wild facility at Gweru, Zimbabwe, by the first and last author. Antelope Park presently holds a population of around 100 African (Zimbabwean) lions.

The recordings were made on 23 November 2010, between 0400 and 0600 hours in the morning at the main enclosure centre. This meant that at least 50 lions were within close earshot, and that most of the other 50 lions were also within hearing range, given that lion roars can be heard by humans at a distance of at least 8 kilometers (Sunquist & Sunquist, 2002:294). Estimated distance between the microphone and the lions varied from about four meters to several hundred meters, although the latter roars appeared as fairly weak signals.

The lions that were closest to the microphone were nine males, most of whom were born in 2006. Also close were seven other males with ages between seven and eight years old. Relatively close were another five males who are seven and eight years old, and also a number of females.

As it was more or less pitch-black during the recording it was impossible to know exactly what lion produced exactly what roar, or whether the roars were produced by a male or a female, although the former is more likely. Besides, there were considerable overlap between the roars of several lions (often more than a dozen at a time).

The equipment used was a Canon HG-10 HD camcorder with a clipon DM50 electret stereo condenser shotgun microphone with a 150–15,000 Hz frequency range and a sensitivity of –40 dB. The microphone was directed towards the lions that roared for the moment, and thus its position varied.

Other than slight contamination with morning bird chirping, the soundscape was relatively calm.

The recording location, with setup indications, is shown in *Plate 1*.

Parken Zoo, Eskilstuna, Sweden

Parken Zoo is a wildlife facility about an hour’s distance from Stockholm and holds a wide number of exotic animals, including several species of felids. There are presently three Asiatic (Gir) lions there: *Sarla*, a female born in 1997 (estimated 165 kilos); *Ishara*, another female born in 2007 (estimated 165 kilos); and *Kaya*, a male born in 1999 (estimated 180 kilos).

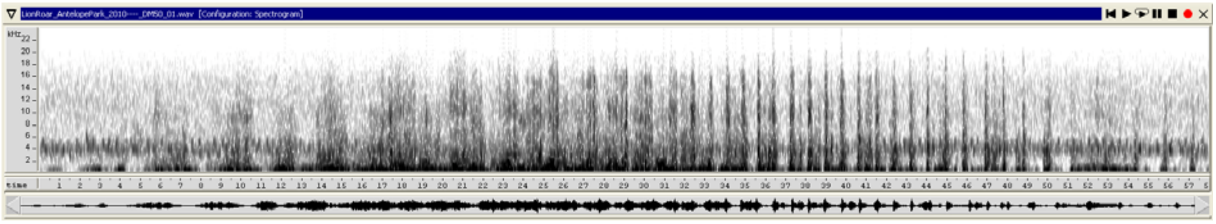
The recordings were made on 7 April 2011, between 0800 and 1000 hours in the morning. The recordings were made indoors to ensure that the lions remained in close proximity to the microphones – in their outdoor enclosure the lions would likely have walked off (far from the microphones). The cameras/microphones were set up by the first author. All three lions were at a distance from the microphones that varied between about one meter to around five meters. Since the recordings were made indoors, there were some echo effects. There was considerable contamination of the soundscape with bird chirps, emanating from a few birds perched somewhere in the enclosure.

The recordings were made with two Canon HG-10 HD camcorders. One camera used the same clipon microphone as is described above, while the other camera used an external professional high-fidelity Audiotechnica AT813 cardoid-pattern, condenser mono microphone, with a frequency range of 30–20,000 Hz and a sensitivity of –44 dB. The two cameras were placed so that, between them, they would cover as much of the enclosure as possible, with the hope of catching roaring sequences on film.

The recording location, with setup indications, is shown in *Plate 2*.

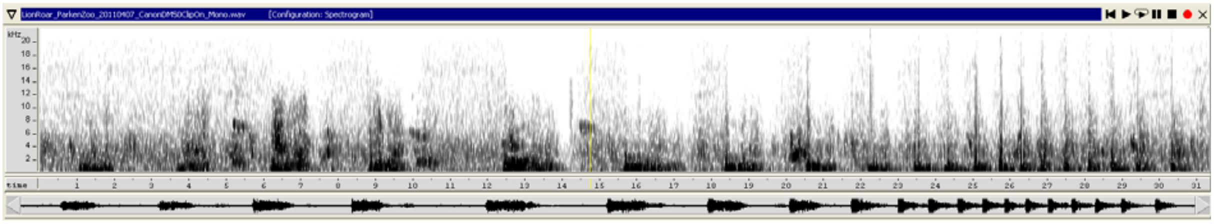
Plates 1 and 2. Recording setups at Antelope Park, Gweru, Zimbabwe (left) and Parken Zoo, Eskilstuna, Sweden (right). Left plate: Orange dots indicate approximate positions of roaring lions while the white dot indicates the most frequent position of the DM50 stereo microphone, ~250 cm above the ground. Right plate: White arrow indicates position of DM50 stereo microphone; orange arrow indicates position of AT813 mono microphone.

Plates 3 and 4. Roaring sequences caught on film at Parken Zoo, Eskilstuna, Sweden. Film captures from the two cameras lifted from the roaring sequences analyzed in this paper. Note that all film/sound files obtained at Antelope Park were recorded in complete darkness (to humans; the lions saw the authors quite well).



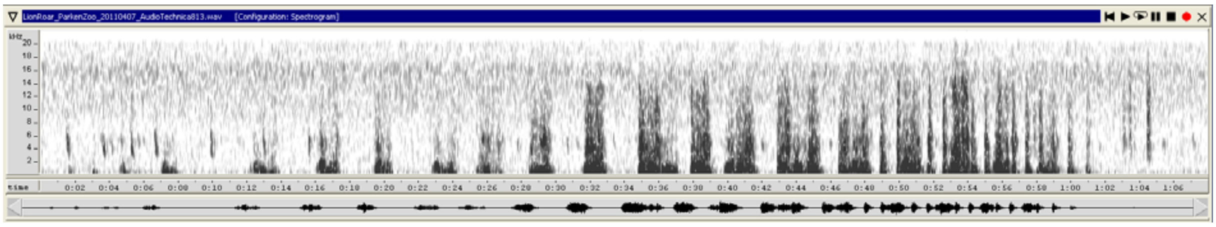
LionRoar_AntelopePark_2010---_DM50_01.wav [Configuration: Spectrogram]

Figure 1. Spectrogram and waveform (excerpt) of multiple lions roaring sequence recorded at Antelope Park, Gweru, Zimbabwe. Canon DM50 clipon stereo microphone. Duration: 58 seconds.



LionRoar_ParkenZoo_20110407_CanonDM50ClipOn_Mono.wav [Configuration: Spectrogram]

Figure 2. Spectrogram and waveform (excerpt) of lion roaring sequence recorded at Parken Zoo, Eskilstuna, Sweden. Canon DM50 clipon stereo microphone. Duration: 31 seconds.



LionRoar_ParkenZoo_20110407_AudioTechnica813.wav [Configuration: Spectrogram]

Figure 3. Spectrogram and waveform (excerpt) of lion roaring sequence recorded at Parken Zoo, Eskilstuna, Sweden. Audiotechnica AT813 external mono microphone. Duration: 67 seconds.

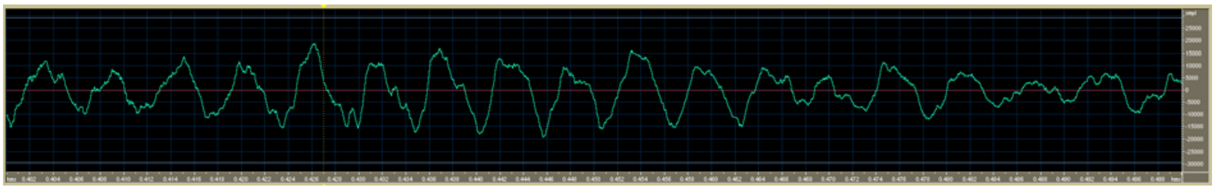


Figure 4. Waveform of lion roaring sequence ("outro" phase) recorded at Parken Zoo, Eskilstuna, Sweden. Audiotechnica AT813 external mono microphone. 18 distinct peaks – in a 100 ms sequence – give an estimated fundamental frequency of about 180 Hz.

Data post-processing

Audio tracks were extracted and converted into wav files (44.1 kHz, 16 bit, mono) with TMPGEnc 4.0 Xpress.

Analysis tools

Spectrogram and waveform analyses were carried out with WaveSurfer and Cool Edit.

Results

The film clips recorded at Parken Zoo resulted in two passages where the lions were caught on film while roaring; see *Plate 3* and *Plate 4*. This enabled comparison between acoustic and visual data (see Ananthakrishnan et al., 2011).

Spectrographic analysis

The three spectrograms shown in *Figure 1*, *Figure 2* and *Figure 3* all reveal the periodic phase characteristics of the roaring sequences. Despite the different acoustic characteristics between the microphones and the different recording setting, all three spectrograms reveal both low frequency components and a higher frequency component around 4 kHz.

Fundamental frequency analysis

A waveform passage is shown in *Figure 4*, and as is clearly seen there are 18 distinct peaks in the 100 ms long window. This gives an approximate fundamental frequency (F_0) of about 180 Hz, which is in accordance with the results reported by Pfefferle et al. (2007:3950), where mean F_0 in males was 194.55 Hz and 206.57 in females.. Naturally, further analyses are required in order to will reveal what degree of variation and range that occur in lion roars.

Discussion

The primary goal of this paper is to provide information about data collection issues associated with animal sounds, highlighting the difficulties involved when trying to obtain controlled high fidelity recordings of animal vocalizations. Future research will focus on more detailed acoustic analyses on the data obtained, and we hope to complement our data with additional high fidelity recordings of uncontaminated recordings of individual lions, in order to facilitate e.g. vocal tract estimation studies (see Ananthakrishnan et al., 2011).

Acknowledgements

Thanks to Jennie Westander, Conny Gärskog and Helena Olsson at Parken Zoo. Thanks to Jacqui Kirk at ALERT. Also thanks to Miriam Oldenburg for help with the "roarcordings".

References

Ananthakrishnan, G., R. Eklund, G. Peters & E. Mabiza (2011). An acoustic analysis of lion roars. II. Vocal tract characteristics. *Proceedings of Fonetik 2011*, 8–10 June 2011, Royal Institute of Technology, Stockholm, Sweden. [This volume.]

Frey, R. & A. Gebler (2010). Mechanisms and evolution of roaring-like vocalization in mammals. In: S. M. Brudzynski (ed.): *Handbook of Mammalian Vocalization – An Integrative Neuroscience Approach*. Amsterdam: Elsevier Academic Press, 439–450.

Grinnell, J. & K. McComb (2001). Roaring and social communication in African lions: the limitations imposed by listeners. *Animal Behavior* 62:93–98.

Hemmer, H. (1966). Untersuchungen zur Stammesgeschichte der Pantherkatzen (Pantherinae) Teil I. *Veröffentlichungen der Zoologischen Staatssammlungen München* 11:1–121.

Leyhausen, P. (1950). Beobachtungen an Löwen-Tiger-Bastarden, mit einigen Bemerkungen zur Systematik der Großkatzen. *Zeitschrift für Tierpsychologie* 7:46–83.

Peters, G. (1978). Vergleichende Untersuchung zur Lautgebung einiger Feliden (Mammalia, Felidae). *Spixiana* (Suppl.) 1:1–283.

Peters, G. & H. H. Hast (1994). Hyoid structure, laryngeal anatomy, and vocalization in felids (Mammalia: Carnivora: Felidae). *Zeitschrift für Säugetierkunde* 59:87–104.

Pfefferle, D., P. M. West, J. Grinnell, C. Packer & J. Fischer (2007). Do acoustic features of lion, *Panthera leo*, roars reflect sex and male condition? *Journal of the Acoustical Society of America* 121:3947–3953.

Schaller, G. B. (1972). *The Serengeti Lion – A Study of Predator-Prey Relations*. Chicago: Chicago University Press.

Sunquist, M. & F. Sunquist (2002). *Wild Cats of the World*. Chicago: University of Chicago Press.

Weissengruber, G., G. Forstenpointner, G. Peters, A. Kübber-Heiss & W. T. Fitch (2002). Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris* f. *catus*). *Journal of Anatomy* 201:195–209.

An acoustic analysis of lion roars. II: Vocal tract characteristics

G. Ananthakrishnan [1], Robert Eklund [2,3,4], Gustav Peters [5] & Evans Mabiza [6]

[1] Centre for Speech Technology, KTH, Stockholm, Sweden
[2] Voice Provider, Stockholm, Sweden
[3] Department of Cognitive Neuroscience, Karolinska Institute, Stockholm, Sweden
[4] Department of Computer Science, Linköping University, Linköping, Sweden
[5] Forschungsinstitut Alexander Koenig, Bonn, Germany
[6] Antelope Park, Gweru, Zimbabwe

Abstract

*This paper makes the first attempt to perform an acoustic-to-articulatory inversion of a lion (*Panthera leo*) roar. The main problems that one encounters in attempting this, is the fact that little is known about the dimensions of the vocal tract, other than a general range of vocal tract lengths. Precious little is also known about the articulation strategies that are adopted by the lion while roaring. The approach used here is to iterate between possible values of vocal tract lengths and vocal tract configurations. Since there seems to be a distinct articulatory changes during the process of a roar, we find a smooth path that minimizes the error function between a recorded roar and the simulated roar using a variable length articulatory model.*

Introduction

The roar is a distinct mammalian vocalization made by only five species of *Felidae*. Researchers suggest that the ability to roar is made possible due to the specialized hyoid apparatus present in these mammals (Weissengruber et al., 2002). Acoustic-articulation modeling has been applied on several mammalian vocalizations in order to estimate the approximate vocal tract length of the animal producing the sound (Hauser, 1993; Taylor and Reby, 2010). The purpose has often been to correlate the estimated length of the vocal tract to the size of the animal to see if larger vocal tract lengths meant relative size dominance. The estimates were further correlated with the social behavior and mating roles of these vocalizations. Most of these methods applied the source-filter theory (Fant, 1970; Titze, 1994) to obtain inferences regarding the vocal tract characteristics. Here the properties of the larynx control the source signal characteristics, while the vocal tract configuration controls the filter characteristics. Since articulation data for mammals have not been very easy to obtain, most of these methods assume a uniform vocal tract for the mammals when they produce the sound and use the formant dispersion method (Titze, 1994; Fitch, 1997).

The lion roaring sequence usually consists of three different phases (Peters, 1978). The first phase is a series of low-intensity calls similar to 'mews'. The second phase, builds up to the climax with calls of increasing duration (shortening again towards the climax). Finally the sequence ends with a series of 'grunt' like sounds. In this study, we are interested in the second phase which is tonal in nature and has the maximum intensity in the entire sequence. Henceforth we only refer to the second phase by the word 'roar'.

Figure 1 shows the spectrogram of a prototypical roar of a female lion. It is clear that there is change in the formant structure also illustrated in *Figures* 2 and 3, showing the Spectral Envelopes varying over time and the average spectral slices for the two parts of a single roar respectively. This change in formant structure indicates that there is a corresponding change in the vocal tract dimensions during the process of producing the roar. Change in the quality of vocalizations have also been observed in other animals to where the vocalization includes protrusion of lips or jaw movement (e.g., Harris et al. 2006). Some species of fallow deer (*Dama dama*) are known to lower their larynx during the call (Vannoni et al., 2005).

Given this observation of changing formant structure during the roar, the uniform tube assumption can no longer be valid. One can suppose that that the filter (vocal tract) undergoes

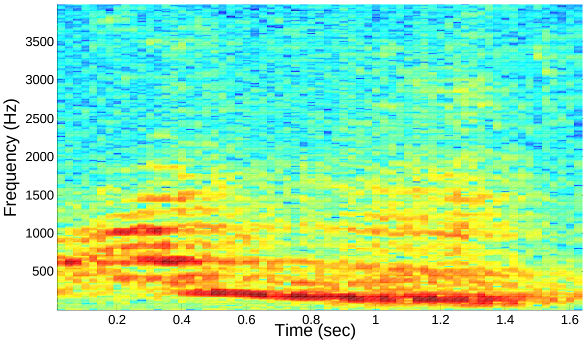


Figure 1: The spectrogram of a typical lion roar (in this case, a female lion's).

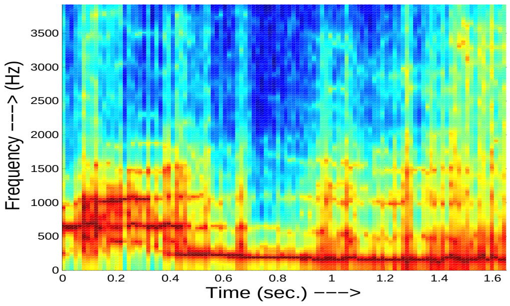


Figure 2: Illustration of the temporal changes in the formant structure, and therefore vocal tract configuration.

some change. However, one does not know what kind of change the vocal tract undergoes, whether it is the lowering of the larynx or changing of the vocal tract area function, or a combination of both.

Theory and Methods

The method proposed in this paper uses a Variable Linear Articulatory Model (VLAM) which allows the articulatory synthesizer developed by Maeda (1979) to be operated at different vocal tract lengths. Although this synthesizer has been designed for human-voices, the source-filter theory as shown previously by Taylor and Reby (2010) can be applied to other mammal vocalizations too. However, since the vocal tract area functions of a lion are largely unknown, we iterate over a range of values and select a configuration which best matches the spectral envelope of the recording of a lion roar. The several steps in the process are described below

1. The lion roar signal is segmented into overlapping windows, using the 'Hann' window function. Each window length is 30 ms in duration and successive windows are 5 ms apart.

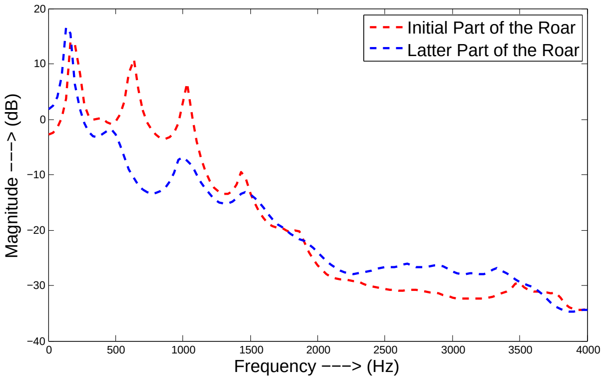


Figure 3: The Spectral Envelope, estimated using LPC analysis, from the beginning and the ending of one lion roar. This indicates that there is some change in the vocal tract configuration during the roar.

2. Linear Prediction Coefficients (LPC) were calculated for each window and then a Fast Fourier Transform (FFT) was applied, to the calculated transfer function so as to obtain the spectral envelope. The number of LPC parameters was set to 21, so as to obtain around 9 to 11 formant peaks within 4000 Hz. This was estimated based on the approximate dimensions of the Vocal-Tract Length (VTL) of a Lion, which is around 35 to 40 cm.
3. The spectral envelope for each window was converted to the decibel (dB) scale and normalized so as to limit the largest formant peak to 0 dB. We also subtracted the mean spectral slope from detected formants, so as to remove the effect of voicing in the estimates of the vocal tract shape.
4. We divided the vocal tract into three equal regions called the Jaw Section, Oral Section and the the Pharyngeal Section. The cross-sectional areas of the three sections were called JawSec, OralSec and PharSec respectively. We performed smoothing and linear interpolation on the three sections in order to approximate a 40 cylindrical tube model.
5. Using the VLAM simulations, we simulated the spectral transfer function, given different combinations of values for the four parameters VTL, JawSec, OralSec and PharSec. The spectral transfer function for each configuration was compared with the spectral envelope of the waveform for each time window to find the Euclidean distance between the two spectra.

6. Since several combinations of VTL and area functions can contribute to largely similar spectral characteristics Atal et al. (1978), we apply a smoothing function on the estimated vocal tract parameters. The movement being a muscular motion, a minimum jerk trajectory is the expected type of movement (at least for humans) Viviani and Terzuolo (1982). We thus apply a minimum jerk smoothing with multiple hypotheses Ananthakrishnan and Engwall (2011). The hypotheses are the 10 vocal tract configurations with minimum estimation error for each frame. These hypotheses are weighted by the inverse of the estimation error.

Data and Experiments

The data we used were recordings of lion roars made at two locations, namely, at the Antelope Park (Gweru, Zimbabwe), and Parken Zoo (Eskilstuna, Sweden). The equipment used at the Antelope Park was a DM50 electret stereo condenser shotgun microphone with a 150–15,000 Hz frequency range and a sensitivity of -40 dB. The estimated distance between the microphone and the lions varied from about four meters to ten meters, with the microphone pointing towards the general direction of a group of nine male lions (most of them born in 2006) in an open enclosure. Although there were other roars, we only considered the loudest roars which we assumed to be from the nine males mentioned above. The recordings at the Parken Zoo were made with two Canon HG-10 HD camcorders. One camera used the same microphone (DM50) as described above, while the other camera used an Audiotechnica AT813 cardoid-pattern, condenser mono microphone, with a frequency range of 30–20,000 Hz and a sensitivity of -44 dB. There where three lions, one male and two females. The male was 12 years old and weighed around 180 kilograms, while the females weighed around 165 kilograms was were around 14 years old. Further details of the data collected are mentioned in Eklund et al. (2011).

The waveforms were initially sampled at 44100 Hz, but were later sub-sampled to 8000 Hz to ensure compatibility with the VLAM model which estimated the vocal tract spectral transfer function in the frequency range of 0 to 4000 Hz. The waveforms were manually segmented to extract the second part of the roaring sequence, i.e. the tonal roar. We used a range of possible vocal tract lengths, ranging from 16 cm to 54 cm. The area functions for the three vocal tract sections were iterated between 8 to 24 sq. cm. These estimates were then compared with the videos sequences wherever available.

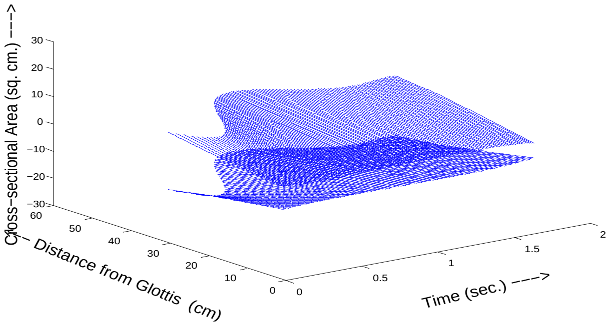
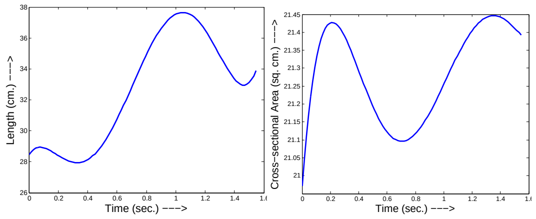
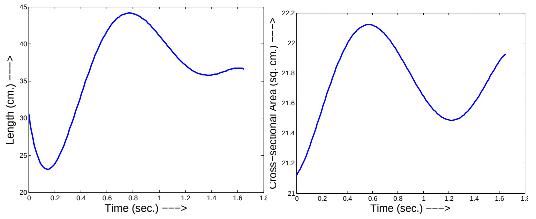


Figure 4: Illustration of how the vocal tract area function changes with respect to time during the course of a roar.



(a) Vocal Tract Length, Male (b) Jaw Section, Male



(c) Vocal Tract Length, Female (d) Jaw Section, Female

Figure 5: Illustration of how the vocal tract length and Jaw cross-sectional areas change with respect to time during the course of a roar.

Results and Conclusions

Figures 4 and 5 indicate the estimated vocal tract shapes and VTLs over time for the female lion. This shows that vocal tract of a lion, approximates a frustum of a cone, rather than a uniform cylinder. The plots also indicate that, the roar involves a lengthening of vocal tract and then then a stabilization during the course of the roar. The range of variation is from 28 cm to 38 cm for the male lion and from 25 cm to 45 cm for the female lion. This may be effected by lowering the larynx, achieved during lifting up of the head. The female lion shows a larger variation in VTL during the course of the roar. The results

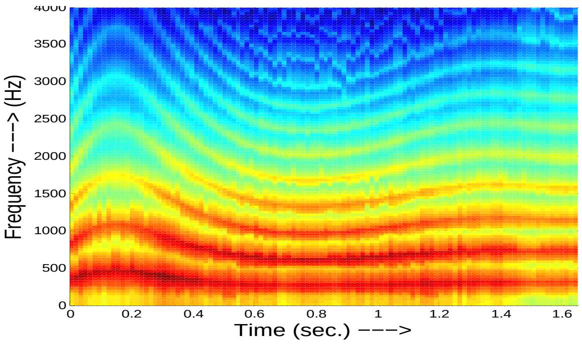


Figure 6: Illustration of the estimated spectral envelope of the lion roar.

also shows a slight decrease in the jaw area, especially for the female. In the videos that were recorded, the lions lifted their head up each time they roared. The jaw saw an increased opening followed by a reduction in the opening during the roars. The prediction from the estimates fit well with the observation about the VTL and the JawSec.

Dynamic analysis of animal vocalizations in order to extract the vocal tract characteristics is a very preliminary attempt in this paper. Some interesting observations have been uncovered in this study. The first being, the general shape of the vocal tract being more conical rather than cylindrical. Secondly, there seems to be a clear indication of larynx lowering, which is similar to the observations on fallow deer vocalizations (Vannoni et al., 2005). However, the female vocal tract is expected to be smaller than the male vocal tract given the differences in overall sizes. The mean VTL of the female lion's is estimated to be around 36 cm and is longer than the male lion's, estimated to be around 32 cm, which is rather unintuitive. Anatomical evidence for a male lion's vocal tract suggests a length of 38 cm Weissengruber et al. (2002). Estimating the mean VTL obscures the fact that the change in VTL for the female lion is also larger than the male lion's. This does not give any indication of what the static and normal lengths would be. Although some observations can be verified using video sequences, other observations need further data and analysis before make strong conclusions.

Future work would include analyzing physical, biological and ecological reasons for this type of motion during the roar, as well as other acoustic properties of the roar. Initial observations point to an increased roughness in the latter part of the roar, likely to be influenced by the voice source. This would also be an interesting investigation.

Acknowledgement

The authors would like to thank Jennie Westander and Conny Gärskog at Parken Zoo. Also thanks to Miriam Oldenburg for help with the "roarcordings" at Zimbabwe and Eskilstuna.

References

Ananthakrishnan G and Engwall O (2011). Mapping between acoustic and articulatory gestures. *Speech Communication*, 53(4):567–589.

Atal B S, Chang J J, Mathews M V and Tukey J W (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. *The Journal of the Acoustical Society of America*, 63(5):1535–1555.

Eklund R, Peters G, Ananthakrishnan G and Mabiza E (2011). An acoustic analysis of lion roars. I: Data collection and spectrogram and waveform analyses. In *Proc. Fonetik 2011*, this volume. Stockholm, Sweden.

Fant G (1970). *Acoustic theory of speech production with calculations based on X-ray studies of Russian articulations*. Mouton De Gruyter.

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.

Harris T, Fitch W, Goldstein L and Fashing P (2006). Black and white colobus monkey (*Colobus guereza*) roars as a source of both honest and exaggerated information about body mass. *Ethology*, 112(9):911–920.

Hauser M D (1993). The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight and social context. *American Nature*, 142:528542.

Maeda S (1979). An articulatory model of the tongue based on a statistical analysis. *J of Acous Soc Am*, 65:S22.

Peters G (1978). *Vergleichende Untersuchung zur Lautgebung einiger Feliden (Mammalia, Felidae)*. Zoologische Staatssammlung M ”unchen.

Taylor A M and Reby D (2010). The contribution of sourcefilter theory to mammal vocal communication research. *Journal of Zoology*, 208:221–236.

Titze I R (1994). *Principles of vocal production*. Englewood Cliffs: Prentice-Hall.

Vannoni E, Torriani M and McElligott A G (2005). Acoustic signaling in cervids: a methodological approach for measuring vocal communication in fallow deer. *Cognition, Brain, Behavior*, IX:551–565.

Viviani P and Terzuolo C (1982). Trajectory determines movement dynamics. *Neuroscience*, 7(2):431–437.

Weissengruber G, Forstenpointner G, Peters G, Kübber-Heiss A and Fitch W T (2002). Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris f. catus*). *Journal of anatomy*, 201(3):195–209.