



HAL
open science

Understanding the origins of vocal complexity in terrestrial mammal vocalizations

Lucie Barluet de Beauchesne

► **To cite this version:**

Lucie Barluet de Beauchesne. Understanding the origins of vocal complexity in terrestrial mammal vocalizations. Biologie animale. Université Jean Monnet - Saint-Etienne, 2023. Français. NNT : 2023STET0003 . tel-04540626

HAL Id: tel-04540626

<https://theses.hal.science/tel-04540626>

Submitted on 10 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2023STET003

**THESE de DOCTORAT DE L'UNIVERSITE JEAN
MONNET SAINT-ETIENNE
MEMBRE DE L'UNIVERSITE DE LYON**
opérée au sein de
Université Jean Monnet

**Ecole Doctorale N° 488
Sciences Ingénierie Santé**

Spécialité de doctorat :
Discipline : Bioacoustique

Soutenue publiquement le 12/01/2023, par :
Lucie Barluet de Beauchesne

**Understanding the origins of vocal
complexity in terrestrial mammal
vocalizations**

Devant le jury composé de :

Adam, Olivier, Professeur, Université Paris-Saclay

Président

Gamba, Marco, Professor, Université de Turin

Rapporteur

Adam, Olivier, Professeur, Université Paris-Saclay

Rapporteur

Briefer, Elodie, Associate professor, Université de Copenhague

Examinatrice

Sèbe, Frederic, Maître de conférence, Université de Saint-Etienne

invité

Reby, David, Professeur, Université de Saint-Etienne

Directeur de thèse

Mathevon, Nicolas, Professeur, Université de Saint-Etienne

Co-directeur de thèse

Au Professeur Joël Gilbert

Acousticien/Bioacousticien en herbe/Collaborateur en or



Avant de continuer, je souhaite écrire un petit mot pour la famille de Joël Gilbert, à qui je dédie ce mémoire. Joël m'a épaulé dès le début de ma thèse, participant à tous mes comités de suivi de thèse avec passion, enthousiasme et un soutien indéfectible. C'était une personne formidable et un collaborateur hors pair sans qui je n'aurais pu réaliser le travail présenté dans ce mémoire. C'est un honneur d'avoir pu le connaître et je souhaite faire de cette thèse un hommage à ce grand chercheur.

Remerciements/Aknowledgement

To me, this is the most important part of the thesis, the one that allows to acknowledge all the people that participated in this accomplishment. Sorry in advance for the long list but it was important to not forget anyone.

First, I thank my jury members **Marco Gamba, Olivier Adam, Elodie Briefer** and **Frédéric Sèbe** for the time you are dedicating to my work, for your valuable feedback and for our future discussions.

Merci à **David Reby**, mon directeur de thèse, qui m'a permis de grandir professionnellement et personnellement au travers de cette thèse. Merci pour tes précieux conseils et tes idées, a priori insensées, qui ont donné un coté exceptionnel à cette thèse. Je souhaite également remercier mon co-directeur de thèse **Nicolas Mathevon** pour son soutien, ses conseils et son accueil au sein de l'ENES. Merci d'avoir fait de ce laboratoire un lieu convivial, propice à l'entraide et aux nouvelles rencontres toutes plus enrichissantes les unes que les autres.

Je souhaite remercier mes collaborateurs **Yann Locatelli** et **Matthieu Keller** de m'avoir ouvert les portes de l'INRAE et de la Haute-Touche. Merci pour vos feedbacks et votre confiance, je n'aurais pas pu récolter toutes ces données sans votre aide. Merci **Frédéric Sèbe** pour tes précieux conseils, ton partage de connaissances et ta joie de vivre qui a fait de ces petits moments, partagés entre deux missions, de supers souvenirs. Merci à toute l'équipe de l'INRA et notamment **François Lecompte, Didier Chesneau, Chantale Porte** et **Maxence** pour leur aide précieuse, lors de mes venues. Thank you, **David Whitby**, for your warm welcoming at Petworth Park it was a great pleasure to meet you and your crazy pets, I wish I would have more time to discuss with you and your badger. Merci également à **l'équipe vétérinaire du zoo de Beauval** qui m'a ouvert les portes de leur salle de scan. Un grand merci à **tous les soigneurs de la Haute-Touche** qui font un travail formidable.

Un *special thanks* à **Léo Papet**, le *grand léo* ou également appelé mon *stagiaire premium* ;P Merci pour ces discussions, ton partage d'expérience, ton aide précieuse sur le terrain des brebis et pour la réalisation du set up de notre salle. Merci de n'avoir jamais rechigné à passer des heures dans cette fameuse salle et pour cette idée d'huile essentielle de menthe dans le masque qui a certainement contribué à la sauvegarde de mes bulbes olfactifs !

Merci à mes précieuses stagiaires **Nolwenn Risser** et **Emmalie Alvarado** qui ont fait un travail formidable. Je suis persuadée que vous êtes destinées à un grand avenir dans la recherche. Ne perdez jamais confiance en vous, *quand on veut on peut* et ce mémoire en est la preuve !

Je ne peux écrire cette partie de remerciements sans avoir un petit mot pour **Charlotte Curé**, chercheuse talentueuse et soutien psychologique de taille. J'ai fait mes premiers pas dans la bioacoustique grâce à toi qui a su me faire confiance dans toutes les tâches que tu as pu me confier depuis 7 ans. J'ai énormément appris à tes côtés et notamment à ne rien lâcher. Je ne serais certainement pas arrivée jusque-là sans ton soutien et ton talent pour remotiver les troupes !

En parlant de troupe, la *team* ENES en est une sacrée bonne. Merci aux permanents **Vincent Médoc**, **Florence Levréro**, **Marilyn Beauchaud** de participer à faire de l'ENES un lieu propice aux échanges et à la convivialité. Un merci particulier à **Joël Attia**, dit *Jojo*, pour sa sagesse (;P) et ces heures passées à discuter d'avenir. **Emmanuelle Combes**, *Manue*, merci pour ton aide, ta réactivité malgré des demandes souvent express. **Nicolas Boyer**, *Neeko*, merci pour toute l'aide que tu m'as apporté au cours de cette thèse, ton ingéniosité pour construire des supports et trouver du matériel adapté, je n'aurais pas pu réaliser toutes ces manip sans toi. Merci pour tes bons petits plats et surtout ta joie de vivre, tu es pour moi un pilier dans ce labo. **Aurélien Pradeau**, tu es devenue très vite essentielle à ce labo et une véritable amie à qui on peut se confier quand on a « un petit coup de moins bien » mais également avec qui on peut partager d'énormes moments de rigolade. Reste comme tu es ne change rien tu es parfaite !

Évidemment l'ENES ne serait pas l'ENES sans l'ambiance et la solidarité entre *jeunes*. **Kasia Pisanski**, thank you for your kindness your support and proofreading. Merci aux *anciens* **Thibaut Marin-Cudraz**, **Juliette Linossier**, **Laura Chabrolles**, **Aourell Lanfrey**, **Daria Valente**, **Livio Favaro** ainsi qu'aux actuels post-doctorants **Elisa Demeru**, **Clément Cornec** et **Loïc Prosnier** pour leur partage d'expériences, leurs précieux conseils et les excellents moments passés au bar. Thank you, **Alexandra Green**, *Ale*, for all crazy moments we had together, I promise that I will come to visit your wild and dangerous Australia, no worries mate ;) Je voulais aussi remercier ceux qui ont été adoptés par la *team* **Leny Lego**, **Coralynn Pierrat** et **Maxime Coëffé** pour la bonne humeur qu'ils apportent. Merci aux p'tits nouveaux **Floriane Fournier**, **Siloé Corvin**, **Naïs Caron Delbosc**, **Wenjing Wang**, **Théophile Turco**, **Léo Perrier** (*p'tit Léo* :P) pour tous ces petits moments partagés, la relève est bien assurée (oui Léo je reconnais que c'est bien une phrase de vieux ça...). Merci **Anna Terrade** d'avoir fait du

super boulot en tant que stagiaire. Je suis super fière et heureuse que tu sois maintenant doctorante à l'ENES, crois toujours en toi ;) Évidemment merci aux doctorants avec qui j'ai partagé une grande partie de ma thèse **Arthur Guibard** et **Julie Thévenet** merci pour tous ces bons moments. Merci **Mathilde Massenet** d'être, mon amie et mon binôme depuis plus de 5 ans. Merci pour ces souvenirs inoubliables sur le terrain que ce soit aux Açores, à Petworth ou dans un élevage canin. En tant que coloc' de bureau depuis tout ce temps, merci pour tous ces instants de gros délire, surtout en fin de journée, de craquage complet à s'envoyer des papiers et se faire des grimaces. Ces moments vont clairement me manquer mais ce ne seront pas les derniers. **Emilie Rojas** bien plus qu'une amie, tellement de fous rires passés ensemble, tu es devenue une petite sœur (*petite sœur je connais ta peine et ta douleur, toi et moi petite sœur...* juste pour le plaisir de te mettre la chanson dans la tête). Même ma famille t'a adopté ! Toujours présente, encore aujourd'hui je rédige ces remerciements alors que tu es dans l'autre pièce de mon appart. Un soutien psychologique de taille même dans les coups durs, tu es une personne formidable. Tic et tac n'ont pas fini de vivre des aventures extraordinaires :D

Merci **Romain Lefèvre** d'avoir été là, de m'avoir épaulée, remotivée et aidée à reprendre confiance en moi.

Merci aux amis qui me soutiennent depuis longtemps voire très longtemps malgré la distance qui nous sépare, **Morgane Raynaud, Diego Ramassamy, Emilie Marotte, Allan Giovagnoli.**

Je tiens évidemment à remercier **ma famille**, sans qui rien de tout ça n'aurait eu lieu. Merci de m'avoir poussée à aller à la fac pour *essayer et ne pas regretter*, finalement ce n'était pas une mauvaise idée :) Merci de m'avoir épaulée toutes ces années et de m'avoir inculquée les valeurs que j'ai aujourd'hui. **Clara** petite sœur et coloc' au top du top, merci pour toutes ces petites anecdotes aux heures des repas, pour les soirées télé et les bons délires. Et, parce qu'ils font partie de ma petite famille, merci à mes lapins Paping et Olaf de m'avoir poussée à faire de la soudure pour réparer mon chargeur d'ordinateur (une nouvelle passion en devenir), merci à **Pantaï** mon petit toutou qui me pousse tous les jours à avoir un rythme de vie plus sain en me sortant de la maison et en me poussant à me socialiser un peu plus.

Abstract

Human speech relies on our unique ability to produce and control a rapid flow of phonetic sounds supported by complex articulatory gestures, such as tongue, mandible or lips movements that allow to engage or disengage oral or nasal vocal tract. The evolutionary origins and evolution of these complex articulatory gestures can be studied using the comparative approach, to identify when, how and why basic articulatory gestures have evolved in other species. Yet, the lack of studies on vocal production in non-primate species limits our knowledge on articulatory gestures in mammals. In this thesis, I first investigate the hormonal factors affecting the development of the vocal apparatus in two closely related *cervinae* species: sika deer which possess a fixed larynx in a high resting position and red deer which possess a mobile larynx in a low resting position. Then, I study basic articulatory gestures in two species: mouth opening in sheep and laryngeal movements in fallow deer.

In the first chapter, I hypothesize that testosterone level accentuates the laryngeal development in species with high vocal sexual dimorphism (i.e., red deer) compared to species with low sexual dimorphism (i.e., sika deer). I show that at the age of 18 months, testosterone level only had marginal effects on the laryngeal development in both deer species, suggesting that testosterone is not responsible of the sexual dimorphism in their laryngeal development, independently of their high or low larynx resting position. Then, I investigate the role of mouth opening in ewes using acoustic recordings and post-mortem investigations. I demonstrate that ewes' vocalizations encode for arousal level in the degree of mouth opening in context of mother-young communication. My results also suggest that the production of these calls involves both oral and nasal tracts and induces the modulation of acoustic parameters (i.e., the first and the fifth formants). In the last chapter, I study the mechanisms underlying the correlation between the laryngeal position and formant modulation in fallow deer bucks' vocalizations. I use post-mortem investigations to document in details how the elongation of both oral and nasal tracts during laryngeal descent induces the lowering of all formants. Finally, to test whether the formant modulation speed has a function in the communication of the vocalizer's motivational state, I conducted playback experiments of synthetic groans of fallow deer bucks. I did not identify an effect of the formant lowering speed on the behavior of rival males, suggesting that this parameter does not transmit motivational information.

My findings do not only provide the first investigation of the effect of androgens on the ontogeny sexually dimorphic vocal apparatuses, but they also highlight the importance of basic articulation gestures in the natural vocal communication of species. As such, they provide a useful background for our understanding of the evolutionary origins of human vocal anatomy and vocal production.

Table of contents

INTRODUCTION.....	- 1 -
Human vocal production	- 2 -
<i>History of vocal production: from Aristotle to Fant</i>	- 2 -
<i>Laryngeal anatomy</i>	- 4 -
<i>Production of the fundamental frequency (f_0)</i>	- 5 -
<i>Articulation and the human vocal tract</i>	- 6 -
Non-human mammal vocal production	- 10 -
<i>Fundamental frequency in non-human mammals</i>	- 10 -
<i>Formants in non-human mammals</i>	- 10 -
<i>Laryngeal descent and speech production abilities: P. Lieberman's theory</i>	- 13 -
<i>Influence of the laryngeal resting position on articulation</i>	- 14 -
Thesis objectives	- 15 -
.....	- 19 -
CHAPTER I: HORMONAL EFFECT ON THE LARYNGEAL DEVELOPMENT.	- 19 -
ARTICLE 1: IMPACT OF ANDROGENS ON LARYNGEAL DEVELOPMENT IN TWO SEXUALLY DIMORPHIC DEER SPECIES.	22
Abstract	23
Introduction	24
Material and methods	26
Ethical statement.....	26
Subjects.....	26
Castration procedure.....	26
Blood sample and injection of testosterone.....	26
CT-scan protocol.....	27
CT-scan measurements.....	27
Statistical analysis.....	30
Results	31
Testosterone level.....	31
Effect of castration in red and sika deer males.....	33
Effect of supplementation in castrated red deer males.....	35
Discussion	35
Influence of testosterone on the larynx growth.....	36

Influence of testosterone Vs body size on the laryngeal descent.....	36
Acknowledgements	38
CHAPTER II: MOUTH OPENING AND EXPRESSION OF AROUSAL IN SHEEP .	40
ARTICLE 2: MOUTH OPENING AS AN INDICATOR OF THE LEVEL OF AROUSAL IN EWES....	42
Abstract.....	43
Introduction.....	44
Material and methods.....	48
Ethics.....	48
Separation experiment	49
<i>Subjects</i>	<i>49</i>
<i>Protocol of the experiment.....</i>	<i>49</i>
<i>Evaluation of the arousal level</i>	<i>51</i>
<i>Selection of bleats</i>	<i>52</i>
<i>Extraction of the observed formants and of the fundamental frequency</i>	<i>52</i>
Post-mortem investigation	53
<i>Subjects</i>	<i>53</i>
<i>Resonances predicted from CT-scan</i>	<i>53</i>
<i>Post-mortem recordings.....</i>	<i>56</i>
<i>Measures from impedance sensor.....</i>	<i>57</i>
Statistical analysis.....	58
Results	58
Influence of arousal level on vocal production.....	58
Effect of mouth opening on the fundamental frequency and on the formants in natural bleats	61
Vocal tract geometry involved in the production of closed and open mouth bleats....	62
Discussion.....	66
Supplementary	70
Acknowledgements	- 71 -
CHAPTER III: LARYNGEAL DESCENT IN FALLOW DEER	- 73 -
ARTICLE 3: MODELLING THE LARYNGEAL DESCENT USING POST-MORTEM APPROACHES. -	75 -
Abstract.....	- 76 -
Introduction.....	- 77 -

Material and methods	- 78 -
Dead specimens	- 78 -
CT-scan.....	- 79 -
Prediction of vocal tract resonances	- 80 -
Post-mortem recordings.....	- 81 -
Results	- 82 -
Formants predicted from CT-scan	- 82 -
Formants observed in post-mortem recordings.....	- 84 -
Correlation between observed and predicted formants.....	- 87 -
Discussion	- 88 -
Acknowledgements	- 90 -
ANNEX TO CHAPTER III: PLAYBACK EXPERIMENT TO TEST THE INFORMATION ENCODED BY THE LARYNGEAL DESCENT OF FALLOW DEER	- 91 -
Introduction	- 91 -
Material and methods	- 91 -
Playback experiment.....	- 91 -
Measure of the behavioral response.....	- 93 -
Statistical analysis.....	- 94 -
Results	- 94 -
Conclusion	- 96 -
DISCUSSION AND PERSPECTIVES	- 98 -
Selective forces driving the evolution of the vocal apparatus	- 100 -
The low resting position of the larynx	- 100 -
Mouth opening in sheep.....	- 101 -
Laryngeal descent in fallow deer	- 102 -
Effect of the source on the filter	- 102 -
New tools for studying the mammal vocal production	- 104 -
Behavioral and acoustic recordings	- 104 -
CT-scan to assess the anatomy of vocal apparatus	- 105 -
Post-mortem recordings	- 106 -
Impedance sensor.....	- 106 -
Conclusion	- 108 -
REFERENCES	- 109 -

Introduction

The main objective of my thesis is to determine the anatomical and developmental mechanisms responsible for the large diversity of sounds in terrestrial mammal species in order to provide a comparative background for the origins and evolution of the complex speech production mechanism that characterize humans. Vocal production in humans has been widely studied and the foundation of our understanding of the vocal production in mammals came from studies on humans. Thus, in this introduction I first review our knowledges on human vocal production, before describing the mechanisms behind the vocal production and possible articulatory gestures in non-human mammals.

Human vocal production

History of vocal production: from Aristotle to Fant

The mechanisms underlying human vocal production have fascinated humans since the Greek antiquity. For example, Aristotle (350 BC) highlighted the importance of air coming from the lungs during vocal production. Galien, in 170, compared the vocal apparatus to a flute, where the glottis was described as the mouthpiece and the trachea as the flute's body. This idea was reversed by d'Acquapendente in 1537, who was the first to consider the pharyngo-buccal cavities as part of the vocal apparatus instead of the trachea. In 1741, Ferrein compared human voice to a new instrument, between a flute and a violin. The "glottal lips" (now known as the vocal folds) were compared to violin strings, the air coming from lungs to a bow, and the chest and lungs efforts to the hand controlling the bow (Ferrein, 1741). He tested his hypothesis using larynges of several species including humans into which he blew air from the trachea (Ferrein, 1741). He showed that the air flow induced vibration of the vocal folds producing sounds that could be modulated by changes to the glottis opening and the air pressure (Ferrein, 1741).

The myoelastic-aerodynamic theory (hereafter MEAD), first evoked by Johannes Müller in the mid-19th century and more explicitly described by Van den Berg (1958), posits that the fundamental frequency of the voice (hereafter f_0) depends on the subglottal pressure and the tension of the vocal folds controlled by the crico-thyroid muscles. When the glottis is closed by the vocal folds, the subglottal pressure increases until it becomes stronger than the closing forces, causing the glottis to open and release the air, thus decreasing the subglottal pressure. The elastic forces of the vocal folds bring them back together until the glottis is again closed. This cycle is repeated a certain number of times per second, which corresponds to the f_0 .

The importance of the vocal tract in speech production was first demonstrated by Helmholtz (1863). He explained the principle of the resonator function of the pharyngo-buccal cavities. In 1960, Fant highlighted the importance of both vocal folds and vocal tract in human sound production in an extremely influential theory called the source-filter theory of speech production. The source-filter theory is based on the assumption that the source (i.e., the larynx) and the filter (i.e., the vocal tract) contribute independently to the production of vocal signals (Taylor and Reby, 2010). Indeed, the air coming from the lungs passes through the glottis and vibrates the vocal folds located at both side of the glottis. This vibration creates a sound source signal characterized by the f_0 and harmonics (corresponding to integer multiples of the f_0). Before its radiation through the mouth and nostrils into the external environment, this source signal is propagated into the vocal tract (i.e., all the supra-laryngeal cavities from the glottis to the lips/nostrils) which filters the sound by attenuating specific frequencies while highlighting broadband frequencies, described as resonant frequencies of the vocal tract and called formants (Figure 1).

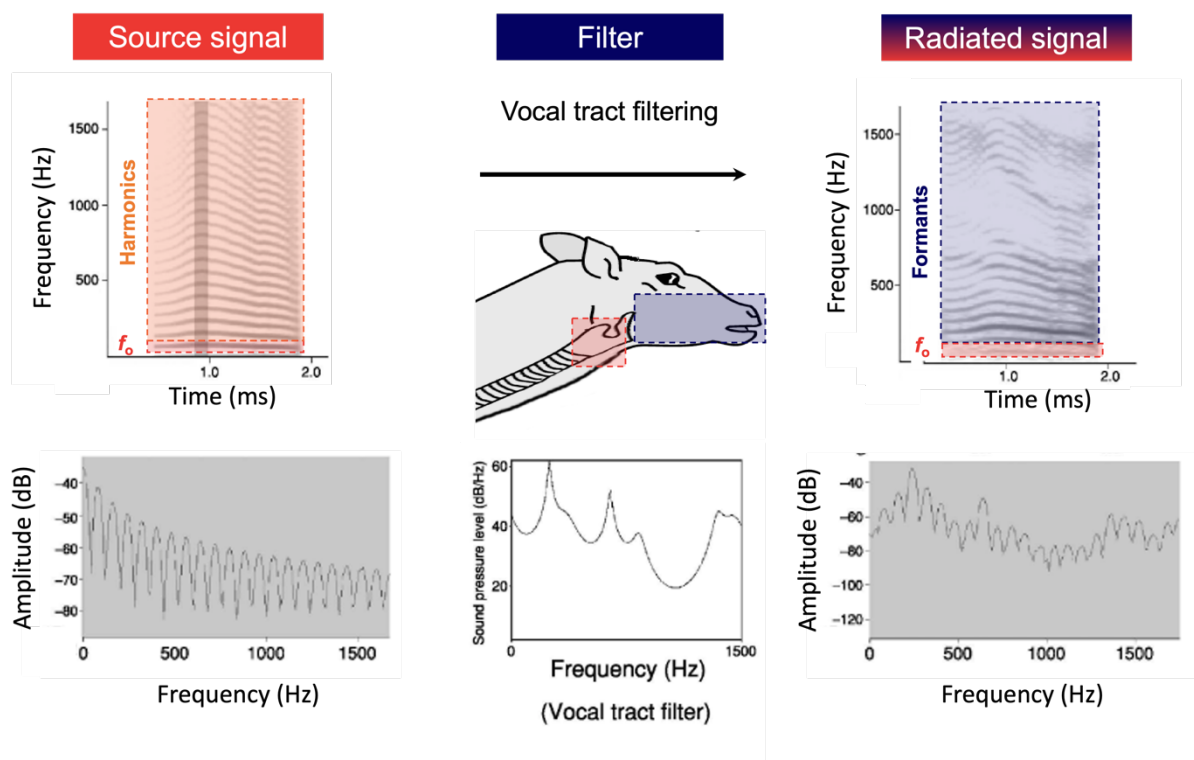


Figure 1. The principle of the source-filter theory and its consequences visible on spectrograms (above) and frequency spectrum (below). The vibration of vocal folds in the larynx (red box) creates the sound source (here represented by a synthesized signal) characterized by a fundamental frequency and harmonics. The vocal tract (blue box) filters this sound by attenuating certain frequency bands relative to others to create formants (broadband peaks of energy). The source-filter theory was first used to explain sound production in humans and later generalized to other mammal species. Adapted with permission from Reby & McComb, (2003b).

In the following sections, I review the anatomy of vertebrates' vocal apparatus, starting with the source, the larynx.

Laryngeal anatomy

The larynx is located at the crossroad of the pharynx, the trachea and the esophagus, which make it a crucial organ implicated in three vital functions: swallowing, respiration and vocal production. It is constituted of five main cartilages: the epiglottis, the cricoid, the thyroid and the two arytenoids, articulated by the muscles and ligaments present in the laryngeal region (Figure 2).

The cricoid corresponds to the larynx bottom and is attached to the top of the trachea by fibroelastic membranes. The thyroid cartilage is the largest part of the larynx, and is attached to the hyoid bone by the thyrohyoid ligament, which itself corresponds to the front of the larynx. The cricothyroid muscle is responsible for the back-and-forth movements of the thyroid cartilage. This movement stretches the vocal ligaments and increases the tension applied on the vocal folds. The two arytenoids are structures on which vocal folds are attached and ride atop the back of the cricoid cartilage. The posterior cricoarytenoid muscle articulates the arytenoids and is responsible for the opening of the glottis whereas the lateral cricoarytenoid and the interarytenoid muscles (transverse arytenoid and oblique arytenoid) are involved in the adduction of the arytenoids and thus the closing of the glottis. The thyroarytenoid muscles, attached along the vocal folds, are responsible for loosening the vocal ligaments and thus decreasing vocal fold tension. Finally, the epiglottis is located at the top of the larynx attached to the inner part of the thyroid cartilage by ligaments. It is responsible for swallowing and for respiratory functions by closing or opening the top of the larynx.

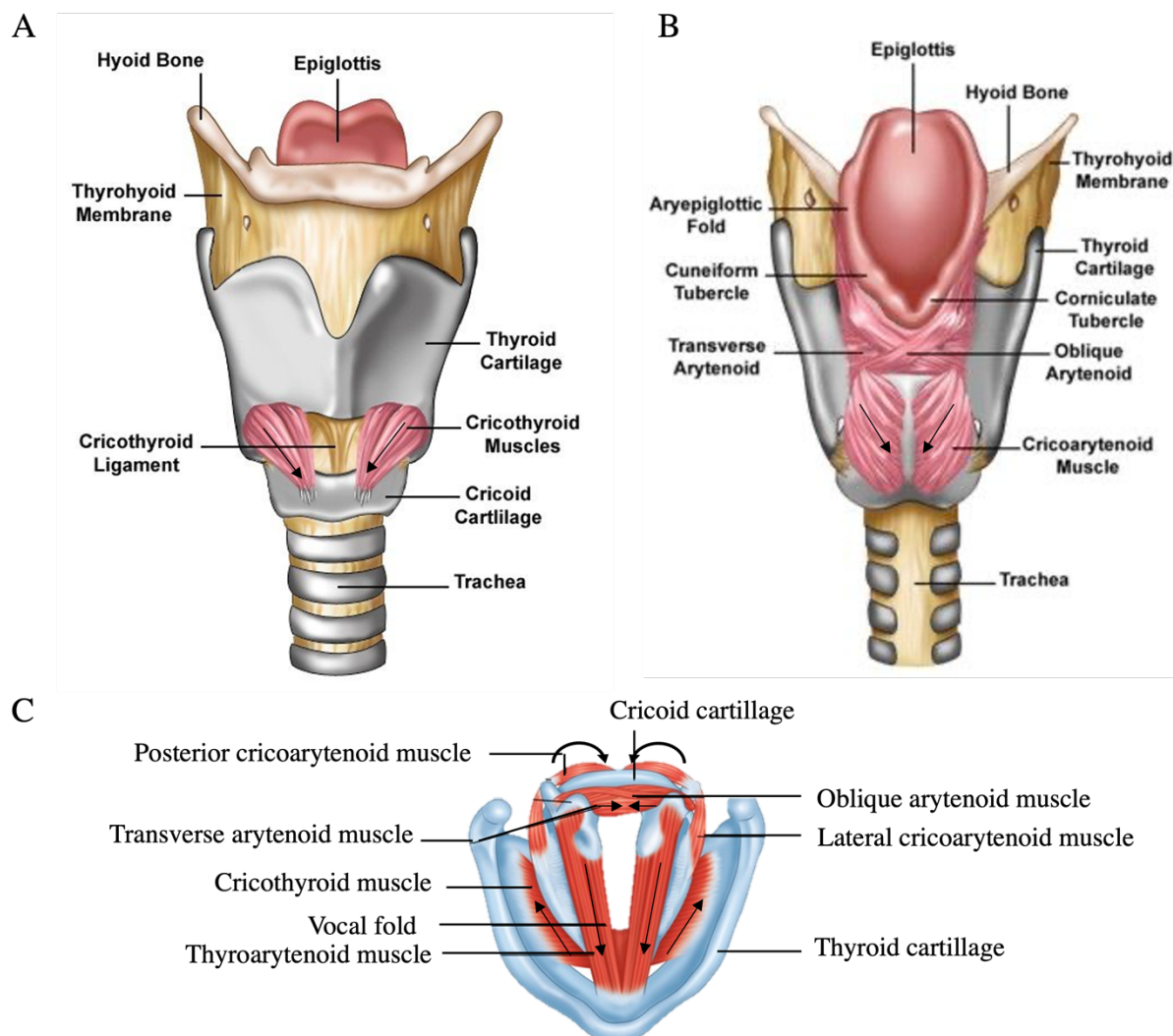


Figure 2. Human laryngeal anatomy showing the cartilages and muscles from a front view (A), internal view (B) and a zoom on the vocal fold region (C). The black arrows show the contracting movements of the muscles. (Stojanović et al., 2013).

Production of the fundamental frequency (f_0)

The fundamental frequency is one of the main components of vertebrate vocalizations and with its harmonics constitutes, in a perceptual point of view, the voice pitch. Four key physiological parameters can be explored to understand the production of fundamental frequency: subglottal pressure, vocal fold tension, vocal fold mass, and vocal fold length (Titze, 1994).

Subglottal pressure is the pressure below the glottis that is generated by the air flow coming from the lungs. The pressure typically achieved during speech is between 1.3 and 2.2 kPa (Kent et al., 1987). A positive correlation between subglottal pressure and fundamental frequency has been demonstrated in humans (Titze, 1989).

The tension of the vocal folds is also involved in the modulation of the fundamental frequency (Titze, 1994). Indeed, vocal fold stiffness is largely dependent of the cricothyroid and thyroarytenoid muscles (Kempster et al., 1988; Titze et al., 1989). The contraction and relaxation of these muscles will impact both the tension of vocal folds and the vocal fold length (Kempster et al., 1988; Titze et al., 1989). Vocal fold mass and length contribute to the determination of the fundamental frequency (Fitch & Hauser, 2003). Vocal folds are constituted of three layers: the epithelium, the *lamina propria* and the thyroarytenoid muscle (Hirano, 1974).

In humans, sexual dimorphism in vocal folds appears at puberty simultaneously with an increase of circulating testosterone levels in males, suggesting a role of this hormone in laryngeal development (Fitch & Giedd, 1999). This hypothesis is reinforced by the presence of androgen receptors in the larynx (Aufdemorte et al., 1983; Boyd et al., 1999; Tuohimaa et al., 1981). However, while there are some evidence that f_0 is negatively correlated with testosterone levels in adult human males (e.g. Dabbs & Mallinger, 1999; Evans et al., 2008), none of these studies have demonstrated the mechanisms whereby testosterone may impact vocal folds development.

Articulation and the human vocal tract

Human speech is distinct from other animal vocalizations by its high modulation of the formant patterns linked with the high mobility of the vocal tract and articulators (e.g., lips, tongue, jaw). As explained earlier in this introduction, the vocal tract corresponds to the air cavities between the glottis and the lips/nostrils (Figure 1; Fant, 1960). It comprises three main parts: (1) the oral tract, which includes the oral cavity and the oropharynx, (2) the nasal tract, which includes the nasal cavities and the nasopharynx, (3) the laryngopharynx which are common components of both oral and nasal tracts. The nasal and oral tracts are separated by the hard palate and the soft palate. Both tracts join together at common point of the vocal apparatus called ostium (at the end of the soft palate) (Figure 3).

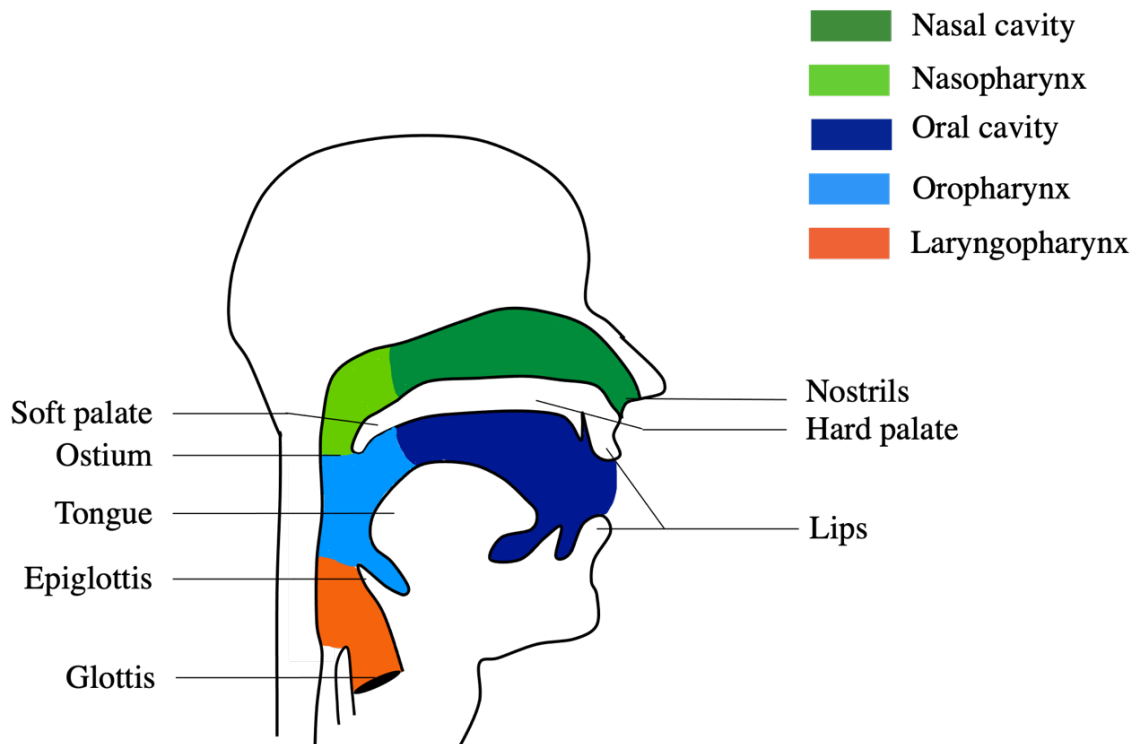


Figure 3. Human vocal tract corresponding to all the air cavities between the glottis and the lips/nostrils.

Humans possess articulators such as the lips, tongue, and mandible that can apply constrictions at several points of the oral vocal tract. These articulators are thus considered to play a key role in speech production as their associated constrictions along the vocal tract can indeed modulate the formant frequencies individually by inducing perturbations in the acoustic wave. Specifically, constrictions of the vocal tract at a point where the velocity is at the maximum and the pressure at the minimum lower the given formant frequency, whereas constrictions at a point where the velocity is at the minimum and the pressure is at the maximum raise the given formant frequency (see Figure 4) (Titze, 1994).

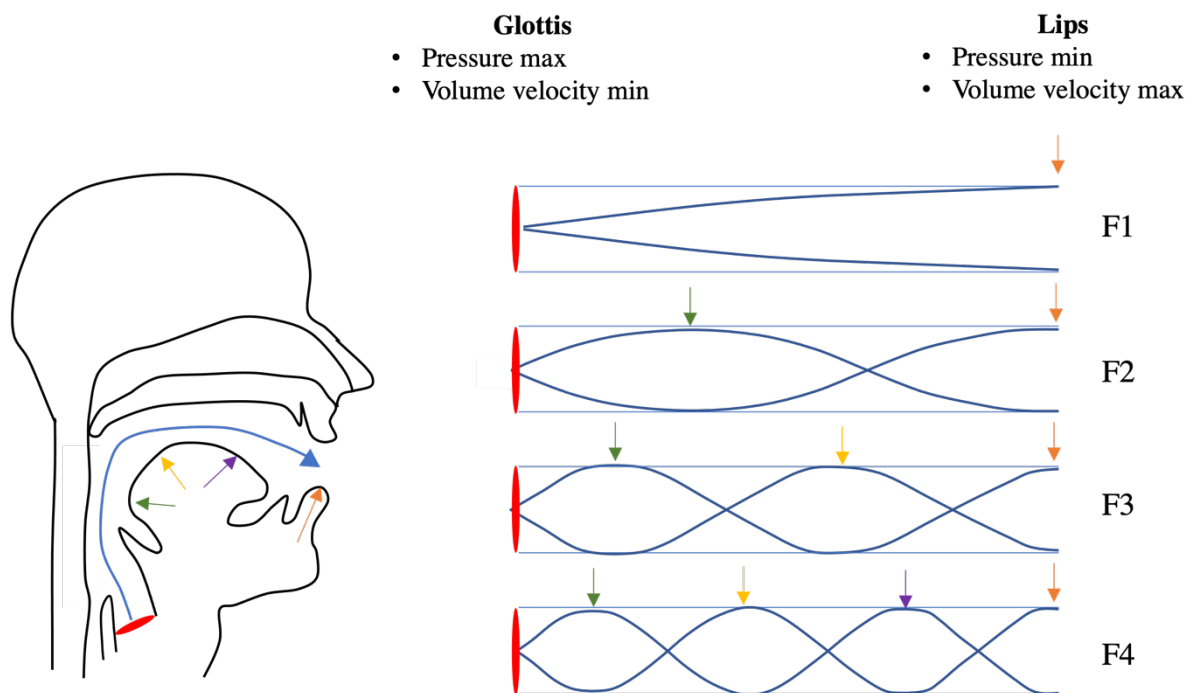


Figure 4. Formants (F1, F2, F3, F4) correspond to standing waves propagated in a cylindrical tube closed at one end (glottis) and open at the other end (lips). Arrows correspond to the area through the oral tract (left) where a constriction lowers the formant as it increases the pressure at an antinode of the standing waves (right). Each colored arrow represents a specific point of potential constriction in the oral tract that affect the standing waves. For example, F1 is mainly affected by a constriction at the level of the lips (orange arrow).

Mandibular movements play an important role in speech production as they are implicated in the production of both vowels and consonants (Stevens, 2000). Singers often open their mouth as wide as possible to produce loud high sound (Joliveau et al., 2004). The lowering of the mandible, increasing the mouth opening, tends to raise all formant frequencies but mostly the first formant characterizing human voice (Behrman, 2021; Joliveau et al., 2004). The mobility of the lips is mostly responsible for the production of bilabial (/p, b, m/) and labiodental (/f, v/) consonants in human speech. More specifically bilabial consonants are produced by the obstruction of the airway between the lips and labiodental consonants with the lower lip against the upper front teeth (Behrman, 2021).

The highly articulated human tongue possesses multiple muscles allowing it to move in all directions (Sanders & Mu, 2013). This articulator is highly involved in the production of vowels and its movements impact both the second and third formants (Figure 5). The second formant is modulated by movements of the posterior portion of the tongue and the third formant by the

tip of the tongue. The tongue movements consisting in constricting of the oropharynx tends to lowers the second formant while increasing the third formant (Behrman, 2021).

In addition to these complex articulatory maneuvers inducing constriction/ relaxation of the vocal tract at specific points, humans are also able to punctually elongate their vocal tract by moving their larynx upward and downward in their throat. As such, to produce the /u/ vowel (*IPA Chart with Sounds – International Phonetic Alphabet Sounds*), the larynx is lowered from its initial position whereas it is raised for the production of the vowel /i/ (Hoole & Kroos, 1998). This movement has also been described in singing humans (Sundberg & Rossing, 1990) and induces changes in the overall formant pattern. Indeed, lowering the larynx induces increases the vocal tract length and thus decreases formant spacing.

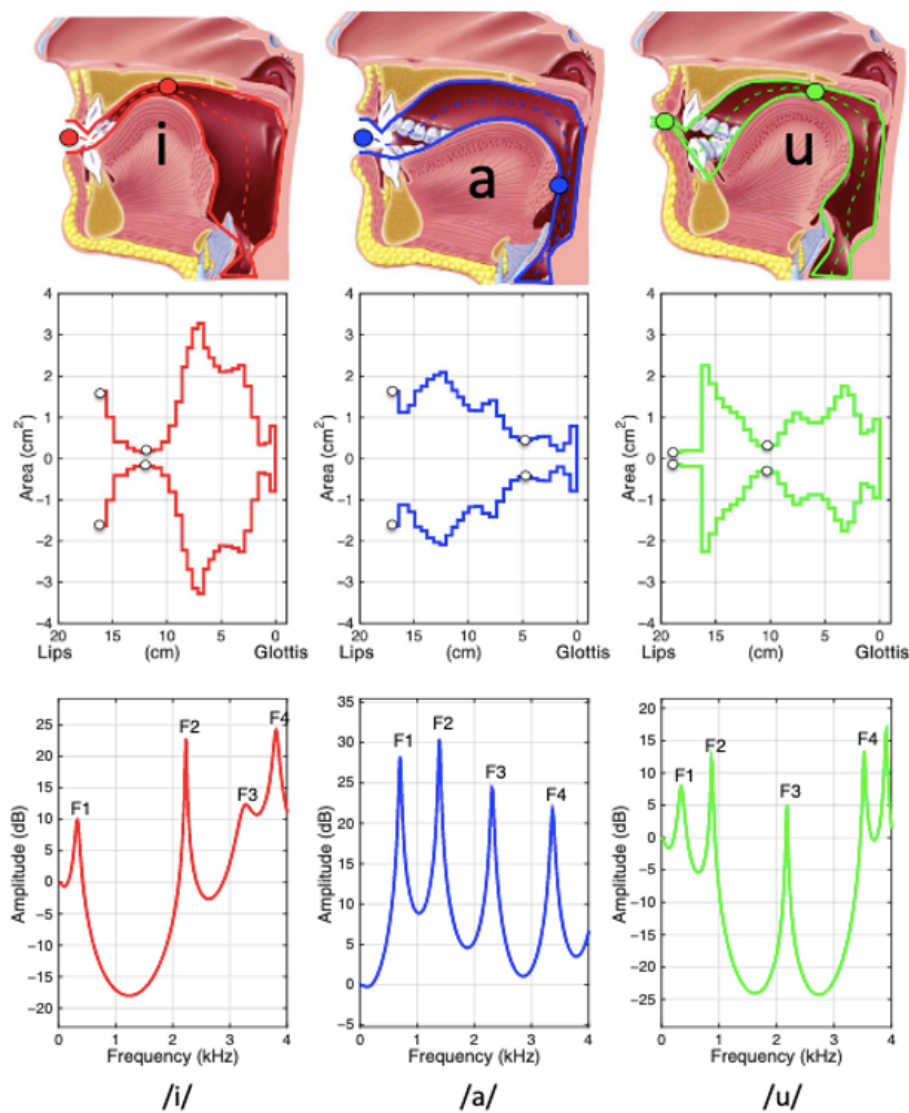


Figure 5. Production of vowels /i/, /a/, and /u/. Sagittal section (top) with dots representing the lips and the vowel’s constriction along the vocal tract. Area function (middle) with dots in corresponding locations. Acoustic transfer function, with the four first formants marked (F1 to F4). (Boë et al., 2019)

The mechanisms of vocal production in humans are thus well understood, however, to understand the evolutionary origins of human articulatory abilities the application of the comparative method is required. This method consists of studying the presence of similar mechanisms in multiple species that share basic foundations of human vocal production such as the other terrestrial mammals. Indeed, most terrestrial mammal species share similar mechanisms of sound production with those of humans, based on the source-filter theory (Figure 1; Fant, 1960).

Non-human mammal vocal production

Fundamental frequency in non-human mammals

Like humans, most mammal species are able to generate a sound source by the vibration of their vocal folds. The fundamental frequency (f_0), that characterized this sound source, can be a reliable cue to static information such as the sex of the caller. Indeed, in humans (Fitch & Giedd, 1999), and other species such as horses (Charuta et al., 2009) and several species of deer (Frey et al., 2021), it has been demonstrated that the larynx is sexually dimorphic with longer vocal folds in adult males than adult females. This induces the production of higher f_0 in females than in males.

Moreover, f_0 can also encode dynamic information such as motivation and level of stress (Feighny et al., 2006; J. Fischer et al., 2004; Taylor et al., 2009). Indeed, f_0 is highly variable and its mean or its modulation can vary as a function of the internal state of the vocalizer and/or the specific context, as shown in many species (goat: Briefer et al., 2015; wapiti: Feighny et al., 2006; baboon: Fischer et al., 2002; domestic dog: Taylor et al., 2009). This is not surprising as we know that affective states are mediated by physical changes influencing the muscular tension on the vocal folds (Bachorowski & Owren, 2008; Hauser, 2000; Scherer, 1986; Titze, 1994). Thus, as the Morton's motivation-structural code suggest (Morton, 1977), animals with aggressive motivation and dominant tend to produce lower-pitched vocalizations in order to larger the impression of size (Owings & Morton, 1998).

Formants in non-human mammals

The generalization of the source-filter theory to non-human mammal species emerged only in the 1990s (Fitch, 1997; Hauser et al., 1993; Riede & Fitch, 1999). Although few earlier studies applied it without making explicit reference to it (e.g. cats : Carterette et al., 1979). Carterette

et al. (1979) demonstrated that a simple uniform cylindrical tube closed at one end can be used to predict formant frequencies in cat vocalizations. This suggests that the main parameter influencing formant frequencies is the length of the oral tract which is closely related to individual body size (Taylor & Reby, 2010). To predict vocal tract length based on vocalizations, Fitch (1997) introduced for the first time a new acoustic parameter called “formant dispersion” defined as the averaged difference between successive formant frequencies. Then, more accurate method was defined by Reby & McComb, (2003), using a linear regression between the formant frequency and the formant spacing (see details Figure 7).

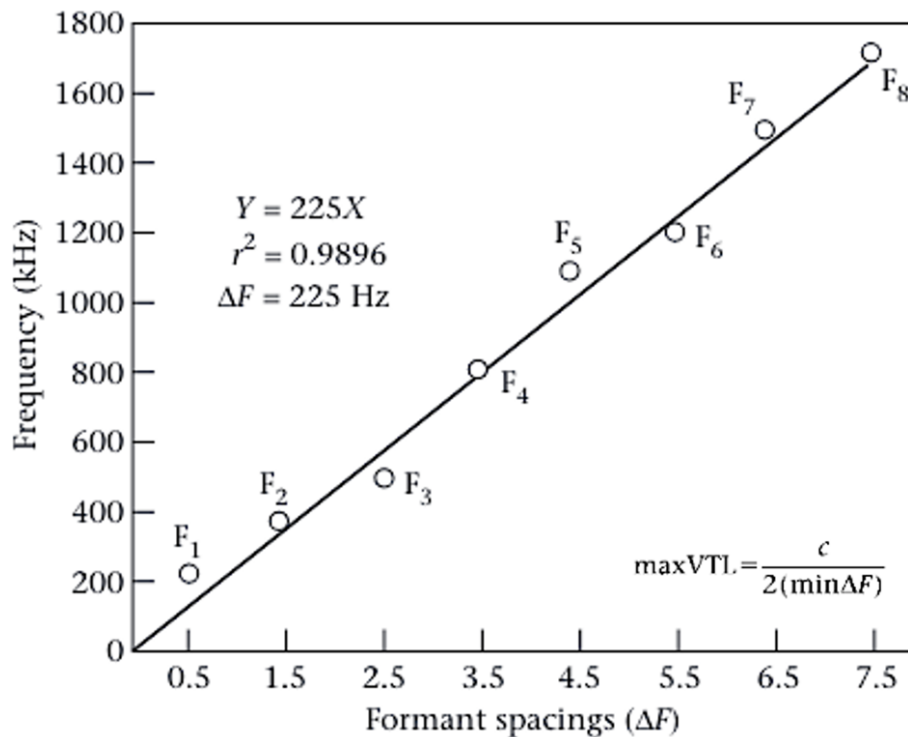


Figure 7. Regression method to estimate the maximal vocal tract length (maxVTL) based on formants frequency. $\min\Delta F$ is the minimum formant spacing and c is the speed of sound approximated as 350 m/s in a mammal's vocal tract. (Reby & McComb, 2003a)

Following these studies, formant frequencies received more attention and became a powerful tool to study the transmission of “honest cues to body size” in mammal vocalizations (Charlton & Reby, 2016; Riede & Fitch, 1999). It has been shown that formants can encode for the caller's identity (e.g., domestic cat (*Felis silvestris catus*), Scheumann et al., 2012). Formants can also encode social information related to individuals' dominant status, and therefore can play a key role in e.g. the context of male competition for female access to avoid potential agonistic interactions (Pitcher et al., 2015; Taylor et al., 2010). In the context mating, the perception of vocal cue to the caller's body size also appears important to assess male quality or formidability. For instance, red deer hinds and female koalas prefer calls given by relatively larger males (Charlton et al., 2007, 2012).

However, formants are not always shaped predominantly by the oral tract (Hattori et al., 1958), as in humans. Indeed, it has been demonstrated that the nasal tract of animals can also shape the formants in several species' vocalizations, such as saiga antelope (Frey, Volodin, et al.,

2007), elephants (Gilbert et al., 2014), and fallow deer (Reby et al., 2018). A possible function for the implication of the nasal tract along the oral tract in vocal production of non-human animals may be the exaggeration of perceived body size by adding additional resonator to the original oral vocal tract. Indeed, the combination of the nasal and oral tracts in groan production in fallow deer results in a higher formant density which may increase body size perception (Reby et al., 2018). In polygynous mammal species, strong selection pressure for the biggest male has shaped anatomical adaptations of the vocal apparatus toward an exaggeration of the caller's body size (Charlton & Reby, 2016). As an example, some species such as lions (Weissengruber et al., 2002), koalas (Charlton et al., 2011) and deer species (Fitch & Reby, 2001), evolved a low-resting position of the larynx, similar to that of the human larynx (Kahane, 1982). This low laryngeal position induces an elongation of the vocal tract (especially the pharyngeal region) and thus a reduction of formant spacing in vocalizations (Charlton & Reby, 2016; Fitch, 1997).

Laryngeal descent and speech production abilities: P. Lieberman's theory

Yet, our current knowledge indicates that humans appear to be the mammal species that possess the highest flexibility in their vocal tract shapes (Lieberman, 1991). One hypothesis to explain this unique characteristic is physical limitations in the vocal tract dynamics of non-human mammals. Using a computer program calculating predicted formant frequencies based on vocal tract anatomy, Lieberman et al., (1969) largely contributed to the initial acceptance of this hypothesis. Indeed, the conclusion of their research was that the lack of a pharyngeal region in most mammal species, caused by a larynx positioned relatively high in the throat (Figure 6), does not allow the high vocal tract dynamics needed for the production of vowels. However, more recent studies, using audio-video recordings and computational models, found that different lips configurations and mouth positions are associated with different call types in chimpanzees (*Pan troglodytes*: Bauer, 1987), rhesus monkeys, (*Macaca mulatta*: Hauser et al., 1993), ruffed lemurs (*Varecia variegata*: Gamba & Giacomini, 2006) and Diana monkeys (*Cercopithecus diana*: Riede et al., 2005). For example, the modulation of formants in Diana monkey's alarm calls has been associated with the modulation in the vocal tract where the lip aperture is narrowed and the mandible is raised (Riede et al., 2005). Moreover, recent findings show that the primate vocal tract could allow the production of higher ranges of formant patterns comparable to human vowels and consonants (Boë et al., 2017; Fitch et al., 2016).

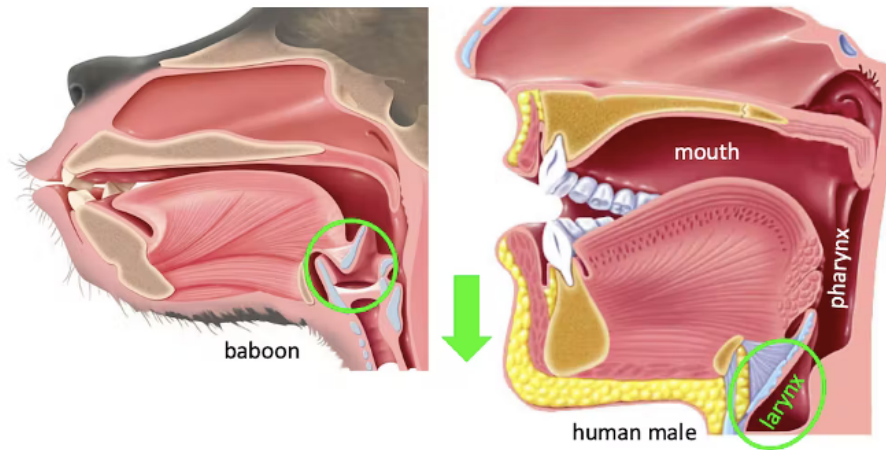


Figure 6. Vocal tract of a baboon (left) compared to the vocal tract of a human (right). The baboon has a larynx in a relatively high position compared to the human larynx which is low in throat, driving the hypothesis that low laryngeal position induced more flexibility of the upper part of the vocal tract necessary for speech articulation (Lieberman et al., 1969). (Laboratory of Cognitive Psychology, CNRS & Aix-Marseille University and GIPSA-lab, CNRS & University Grenoble-Alpes)

Hence, contrary to the P. Lieberman's theory, non-human primates are able to perform basic articulatory gestures even with a larynx in high position.

Influence of the laryngeal resting position on articulation

Regarding vocal tract articulatory gestures in non-primate mammal species, a pioneering study showed that different jaw configurations may encode arousal levels in cats (Shipley et al., 1991). In this study, the degree of mouth opening was measured while the cats produced vocalizations in a context of isolation. The authors showed that 1) the variability in the fundamental frequency (f_0) did not affect the frequency containing the greatest energy (f_{max}) suggesting that f_{max} is a formant, and thus independent from f_0 ; 2) the degree of mouth opening was highly correlated with f_{max} rather than f_0 . Thus, this study showed the role of simple articulation gestures in cats' vocal production. Since then, only sporadic studies have investigated the function of vocal tract articulation gestures in the sound production and most of them were focused on laryngeal movements. Using cineradiography, Fitch (2000) showed the presence of restricted laryngeal movements in several species with a laryngeal position high in the throat, such as monkeys, dogs, goats and pigs. The retraction of the larynx and the raising of the velum allows these species to disengage the nasal cavity during calls, which favors the involvement of the oral cavity over than the nasal cavity (Fitch, 2000). In dogs, laryngeal descent was enough to create a cavity comparable to the pharyngeal part found in species with

a low-resting larynx (Fitch, 2000). A mobile larynx was also found in impala (*Aepyceros melampus*: Frey et al., 2020) and in species with permanently descended larynges such as, the goitered gazelles (*Gazella subgutturosa*: Frey et al., 2011) and several polygynous deer species (fallow deer: McElligott et al., 2006; red deer: Reby & McComb, 2003a). For instance, the surprisingly high mobility of the male fallow deer's larynx allows males to increase their vocal tract length by an average of 52% when calling (McElligott et al., 2006).

Several examples of mammals mimicking human speech sounds have further revealed their abilities to modulate their vocal tract (e.g., Hoover the seal: Ralls et al., 1985; Koshik the elephant: Stoeger et al., 2012). However, such vocal imitation has been reported for very few human speech sounds and to the best of my knowledge, no animals perform such vocal tract articulation in their own and natural vocal system.

These studies show that some vocal tract dynamics exist even in species with a larynx in a high position and highlight the effects of these gestures on the vocalizations' acoustics. However, the presence of speech precursors of articulation has received little attention in non-primate species. To better understand the mechanisms underlying complex human articulation, it is thus necessary to apply a comparative approach by studying non-primate species that share basic articulation gestures with humans.

Thesis objectives

In this introduction, we have seen that while terrestrial mammals share common vocal production mechanisms, largely following the principles of the source-filter theory, their vocal behavior can be complex and diverse due to a range of anatomical and behavioral specificities. While the high position of the larynx has long been put forward to hastily exclude articulation in non-human mammals, it is now broadly accepted that the vocal apparatus of primates is not at the origin of their limited articulation abilities (Lieberman, 2012). Their vocal behavior actually shows evidence of articulation contributing to the diversity of their vocal repertoire (Lieberman, 2012). The aim of my thesis is to help to fill the gap in our understanding of articulation abilities in non-primate mammalian species, by studying the laryngeal development and mechanisms underlying articulation gestures in species whose vocal apparatus is characterized by a relatively high (sheep) or low (deer) laryngeal resting position.

The phonetic complexity and capacity for articulated speech in humans was typically attributed to the presence of a permanent descended larynx (Lieberman, 2012). However, this specificity has since been identified in several species that are phylogenetically distant from humans such as koalas or several polygynous deer (Charlton et al., 2011; Frey et al., 2021). Moreover, in these species, the larynx of males is bigger and characterized by a lower-resting position compared to females. As testosterone is involved in the development of most sexual secondary traits (Greenwood & Adams, 1987), I hypothesize here that the sexual dimorphism of the vocal apparatus of these species are under the influence of testosterone. The first objective is, thus, to study the role of postnatal testicular androgens on the development of the vocal apparatus. To do this, in the first chapter, I propose to contrast the effect of androgens on the development of the vocal apparatus in two closely related species that show major differences in their vocal anatomy and vocal behavior. Specifically, I compare two *cervinae* species: sika deer where males possess a small barely sexually dimorphic larynx in a high resting position, and red deer, whose males possess a highly sexually dimorphic larynx in a low resting position.

The second objective of my thesis is to test whether species with larynges resting in a high position can modulate their vocalizations by simple vocal tract movements, such as mouth opening. Ewes constitute a good model to investigate this, as they have a relatively simple vocal repertoire comprising two main types of vocalizations characterized by their mouth configuration (i.e., open mouth in high bleats and closed mouth in low bleats). Combining behavioral and postmortem methodologies, I investigate information encoded by the mouth opening and the mechanisms of sound production resulting from this articulatory gesture. I argue that identifying such basic precursors of articulatory gestures in species with a larynx in a high resting position contributes the diversification of articulation and thus to our understanding of phonetic complexity.

Similarly to humans, fallow deer possess a larynx in low resting position and is highly mobile when they vocalize. While the external observation of laryngeal descent in fallow deer has been correlated with the formant modulations observed in their groans (Fitch & Reby, 2001; McElligott et al., 2006), the effect of vocal tract modulation on formant frequency patterns has yet not been investigated.

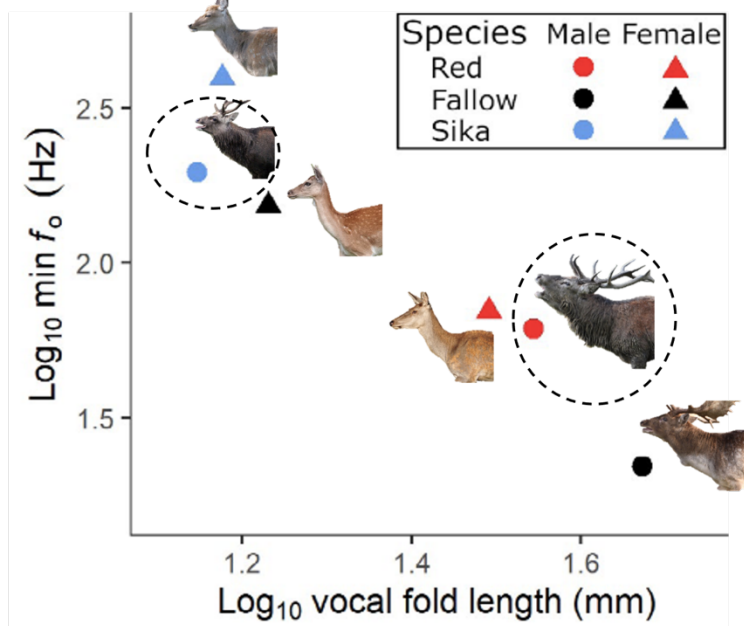
The last objective of my thesis is thus to determine the effects of vocal tract shape and vocal extension on formant patterns to understand the mechanisms underlying this complex mobility as found in fallow deer males. Using post-mortem specimens in chapter III, I demonstrate the implication of the different vocal tract shapes on the dynamic formant patterns found in fallow deer groans. Finally, in an exploratory experiment, I tested the possible communicative function of this formant dynamic using playback experiments on rival males.

Chapter I: Hormonal effect on the laryngeal development



To understand how the sexual selection pressure can drive a part of the larynx and vocal tract diversity, it is important to study the factors influencing the laryngeal development.

Species from subfamily of *Cervinae* are good models to study laryngeal development as there is a high variability in sexual dimorphism between species (Geist & Bayer, 1988). Indeed, fallow deer possess the strongest sexual dimorphism in their vocal fold length and their vocal tract length, followed by red deer whereas in sika deer both sexes present a larynx in high position with relatively small vocal folds (Frey et al., 2021; Wyman et al., in prep.). Testosterone, present in higher level in males than females, influences the development of most of the secondary sexual traits in deer species such as the antler growth (Greenwood & Adams, 1987). The presence of testosterone receptors in the laryngeal region let to hypothesize that testosterone level might play a role in the laryngeal development.



Relationship between the minimal fundamental frequency (f_0) and the vocal fold length in both sex of three deer species the red deer, the fallow deer and the sika deer. Circled in dot line: animals studied in chapter I. (Adapted from Frey et al. 2021).

Thus, in this chapter, I study the effect of testosterone on the laryngeal development in red and sika deer. The strong effect of castration in fallow deer males and the fragility of this species did not allow me to test the effect of testosterone on this species. I first compare the laryngeal growth and descent between intact and castrated groups of sika and red deer without supplementation in testosterone. Then, I focus on four red deer groups in which different treatments were applied: intact, castrated without supplementation and two groups of castrated with different level of testosterone supplementation. While my results confirm the species-specific differences in the laryngeal anatomy of red and sika deer males, crucially I show that testosterone has only a marginal role on laryngeal development in both species.

Article 1: Impact of androgens on laryngeal development in two sexually dimorphic deer species.

Lucie Barluet de Beauchesne¹, Yann Locatelli², Baptiste Mulo³, Nicolas Mathevon^{1,4} and David Reby^{1,4}

¹ ENES Bioacoustics Research Laboratory, CRNL, CNRS, Inserm, University of Saint-Etienne, France

² Réserve de la Haute Touche, Muséum National d'Histoire Naturelle, Obterre, France

³ ZooParc de Beauval & Beauval Nature, 41110, Saint Aignan, France

⁴ Institut universitaire de France

Abstract

In mammalian polygynous species, the fact that larger males typically have better access to reproduction can lead to the evolution of secondary sexual traits enabling the exaggeration of body size. For example, the males of many species of deer have evolved a disproportionately large vocal apparatus, allowing them to produce lower frequencies and thus sound bigger. While it is now well established that the development of secondary sexual traits, including the exaggerated low male voice of humans are influenced by the androgens, the role of testosterone on the development of the vocal apparatus remains unclear. To better understand the implication of testosterone on the growth and on the descent of the larynx, we studied two deer species characterized by low (i.e., sika deer (*Cervus nippon*)) or high (i.e., red deer (*Cervus elaphus*)) level of vocal sexual dimorphism. More specifically, we used computer tomography imaging to compare the development of the vocal apparatus in uncastrated males (hereafter intact males), castrated males without testosterone supplementation (hereafter *No_ST*), and castrated males with two levels of testosterone supplementation (in red deer only, supplementation at physiological level hereafter *Phy_ST* and at higher level hereafter *High_ST*). Our preliminary results confirmed that sika deer and red deer show strong developmental differences in laryngeal growth and descent. At 18 months of age, red deer males have achieved significantly greater laryngeal growth and descent than sika deer males. We did not find an effect of testosterone supplementation on the overall larynx measures of size and on the extent of laryngeal descent. However, analysis on individual laryngeal measurements showed that testosterone had an effect the height of the cricoid and on the maximal length of the thyroid lamina. These preliminary results thus suggest that testosterone has a marginal role on laryngeal development and that the sexual dimorphism observed in the vocal apparatus of red deer is mostly due to the sexual dimorphism in body size. We highlighted the importance of the sexual selection on the evolution of the sexual dimorphism observed in polygynous species.

Introduction

Sexual dimorphism is a widespread phenomenon in animal kingdom and can be expressed by different colors, shape or body size (Greenwood & Adams, 1987; Owens & Hartley, 1998). A common hypothesis on the cause of this sexual dimorphism is the sexual selection. Indeed, sexual selection resulting either from male-male competition or female mate choice induces the development of specific anatomical structures allowing to increase the chance to dominate the other males and/or to access to the reproduction (Clutton-Brock & Vincent, 1991).

The strong sexual selection pressures, characteristic to polygynous species, favor the selection of relatively big males, either to attract females or to deter male competitors (thus avoiding potentially costly agonistic interactions (Hunt et al., 2009; Smith & Price, 1973)). As such, size communication plays a key role in inter- and intra- sexual interactions. Among sensory channels, animals often use acoustic cues indexes of caller body size. Indeed, in accordance with the source-filter theory, mammalian vocalizations result from two stages of production process both in close relation with the body size. First, air is expelled from lung through the glottis containing the larynx with the vocal folds (described as the “source”). The rate at which vocal fold vibrate determine the fundamental frequency of acoustic signals. Then, this sound source propagates through the supra-laryngeal vocal tract where it is filtered before its radiation through the animals’ nostrils or mouth into the environment. The resonance frequencies of the vocal tract correspond formants of vocalizations (Fant, 1960; Taylor & Reby, 2010; Titze, 1994). Vocal folds are also constrained by the body size which can induce a negative correlation between the body size and the fundamental frequency (Evans et al., 2006; Morton, 1977; Pfefferle & Fischer, 2006; Titze, 1994). However, the lack of correlation between the body size and the fundamental frequency in a lot of species suggest that other parameters influence directly the vocal folds growth and the production of fundamental frequency (Reby & McComb, 2003a; Rendall et al., 2005). Formants are more directly dependent on vocal tract anatomy and a large number of studies have demonstrated a negative relationship between vocal tract length and formant spacing (Charlton et al., 2011; Fitch, 1997; Reby & McComb, 2003a; Riede & Fitch, 1999).

The presence of androgen receptors in the laryngeal region of multiple species suggests a direct effect of testosterone on the laryngeal development (Boyd et al., 1999; L. M. Fischer et al., 1995; Saez & Sakai, 1976). Indeed, testosterone can regulate the expression of genes via its interaction with androgen receptors (Brinkmann et al., 1999). A high level of testosterone

increases the chance to activate the androgen receptor and we know that testosterone present in higher quantity in male than in female, is implicated in most of secondary sexual characters observed in males (Greenwood & Adams, 1987; Owens & Hartley, 1998). Beckford et al., (1985) showed that the testosterone supplementation in castrated sheep increased the larynx growth and more specifically the thyroid and cricoid cartilages. They did not find an effect on the vocal fold length (Beckford et al., 1985). The larynx and fundamental frequency of sheep are not highly sexually dimorphic, thus testosterone may have a higher impact in species where males have higher larynx dimorphism. For example, in human, male developed bigger larynx and longer vocal folds than in female which is consistent with the sexual dimorphism observed in their fundamental frequency (Kahane, 1982). Several studies suggested, without demonstrating it, an effect of testosterone on the vocal folds development as they found negative correlation between the testosterone level and the human pitch (Dabbs & Mallinger, 1999; Moore et al., 2005).

Species from subfamily of *Cervinae* are good models to study the effect of testosterone on laryngeal development as there is a high variability in sexual dimorphism between species (Geist & Bayer, 1988). Red deer have an important sexual dimorphism especially in the resting position of their larynx which is much lower in males than in females (Frey et al., 2021). In contrast, in sika deer, both sexes possess a larynx characterized by a relatively high resting position (Frey et al., 2021). These two species are also dimorphic in their larynx size as red deer have a larger larynx than sika deer (Frey et al., 2021). These anatomical differences have an effect on deer vocalization as the red deer can produce very lower calls (mean f_0 : 106.9 Hz Reby & McComb, 2003) than sika deer (f_0 range: 196-1187 Hz ; Minami & Kawamichi, 1992). In our study, we test the effect of testosterone on the larynx growth and the laryngeal descent in these two species. We hypothesize that testosterone level accentuates laryngeal descent in species with a high sexual dimorphism (i.e., red deer) compared to species with low sexual dimorphism in laryngeal position (i.e., sika deer). We thus expect that testosterone would have a higher effect on the red deer larynx growth and specifically on vocal fold length. Anatomical changes will be measured using CT-scan (computer Tomography scanner) made during the development of the individuals at the age of 6,18 and 30 months. Here, we present the first results after the second CT-scan.

Material and methods

Ethical statement

The experiment was approved by the local and French national ethics committees (APAFIS#22166-2019092712174749v3). A longitudinal observation with a welfare scoring was applied to stop the experimentation in case of change in animal health (i.e., variation in the weight, physical appearance, posture).

Subjects

We used 18 red deer and 10 sika deer bred in the zoo parc of La Haute-Touche. Prior to experiments, individuals were trained to be manipulated to facilitate transport and injections/ blood samples. Habituation was performed twice a week and consisted in enclosure changes and exposures to livestock vehicles before feeding animals. The mineralized antlers were cut to avoid injuries between animals.

Castration procedure

Castration was performed in 6-month-old deer that were fasted 24h before the surgery. Just before the surgery, animals were anesthetized with an injection of Zolazepam/Tilétamine (Zoleti®) ~1 to 2 mg/kg and xylazine 1.5mg/kg. Animals were intubated to allow the injection of isoflurane in cases where it was necessary to prolongate the anesthesia. The scrotal skin and vaginal tunic were incised allowing testicle to be exteriorized. Clamps were positioned on the *vas deferens* and the vascular network before ligaturing them. After 5 minutes, the testicles were resected and the vaginal tunic and scrotal skin were sutured. Antibiotics were administered postoperatively. Awakening was triggered by intramuscular injection of atipamezole (ANTIDORM®).

Blood sample and injection of testosterone

In order to monitor testosterone levels, blood samples were taken from the jugular vein every 2 weeks: 5 ml in young animals (less than 10 months) and then, 10 ml / individual.

In red deer supplemented with testosterone, Testosterone Enanthate Long Action (Schering) (175 for sika deer and 275 mg/red deer) was administered every 2 weeks from castration until

antlers reached 5 cm length. These injections were then interrupted until the end of the summer of the following year, in order to induce new cycles of antler mineralization and fall.

CT-scan protocol

Animals were transported from La Haute-Touche Zoopark to the Beauval Zoopark to perform CT-scans. The Beauval Zoopark possess all facilities to take in charge the animals (e.g., indoor enclosure, facilities to access to the room with the scanner). Deer were anesthetized before CT-scan using same protocol as previously described (see *castration procedure* for more details). The first CT-scans were performed in December 2019, when animals were 6 months of age. The second CT-scans were performed in February 2021, i.e., a year after testosterone supplementation had started. Body weight of each individual was also measured while it was anesthetized.

After CT-scan, the animal was placed in sternal recumbency in its transport crate, then awakening was ensured by intramuscular injection of atipamezole (ANTIDORM®). At least two hours after waking up, the animal was transported back to La Haute Touche ZooPark and kept in an observation indoor enclosure for a period of 24 hours.

CT-scan measurements

For both CT-scan, at 6 months and 18 months, and for both species, red and sika deer, we used the same following protocol to measure vocal apparatus. All the CT-scan measurements were obtained using the Osirix software (version 12.0). The length of the mandibula and the distance between the angle of the stylohyoid bone and the anterior prominence of the thyroid cartilage (used to measure the laryngeal descent) were obtain from the *3D volume rendering*. The other measurements were obtain using the *region of interest (ROI)* function from the *3D Multi-Planar Representation* viewer which allows to modulate the axis to obtain either the mid-sagittal section at the laryngeal position or the cross section at the vocal folds position (Fig.1; Fig.2). We choose the *CT-Brain* pre-set contrast allowing a good contrast between tissue and cartilages constituting the larynx.

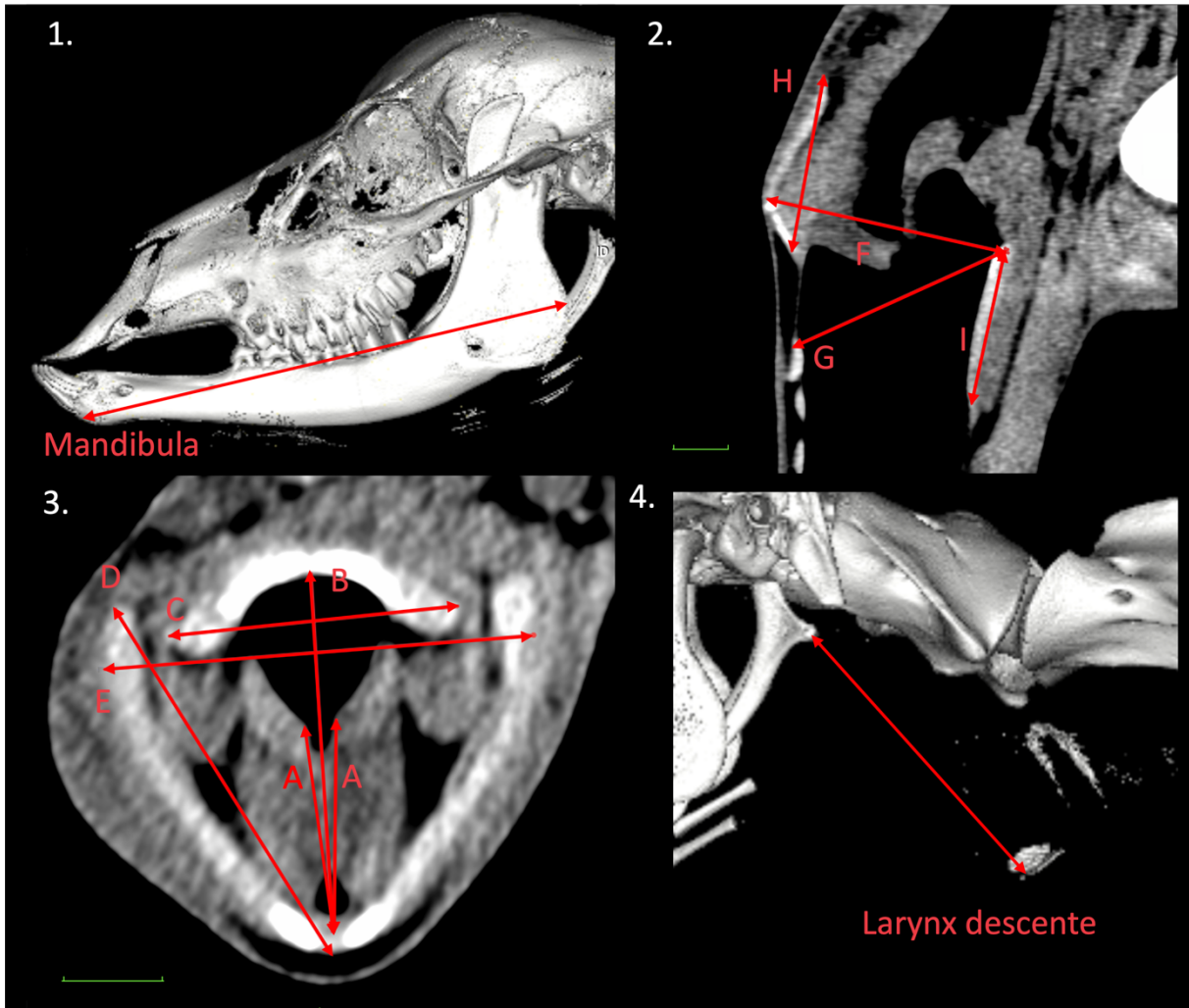


Figure 1. Anatomical measurements based on CT-scan of a red deer. 1) Left side view of cranial bones to measure the mandible length 2) sagittal section at the laryngeal position 3) cross section at the vocal folds position; 4) left side view of stylohyoid bone and larynx cartilages to measure the descent of larynx. Scale: 1 cm.

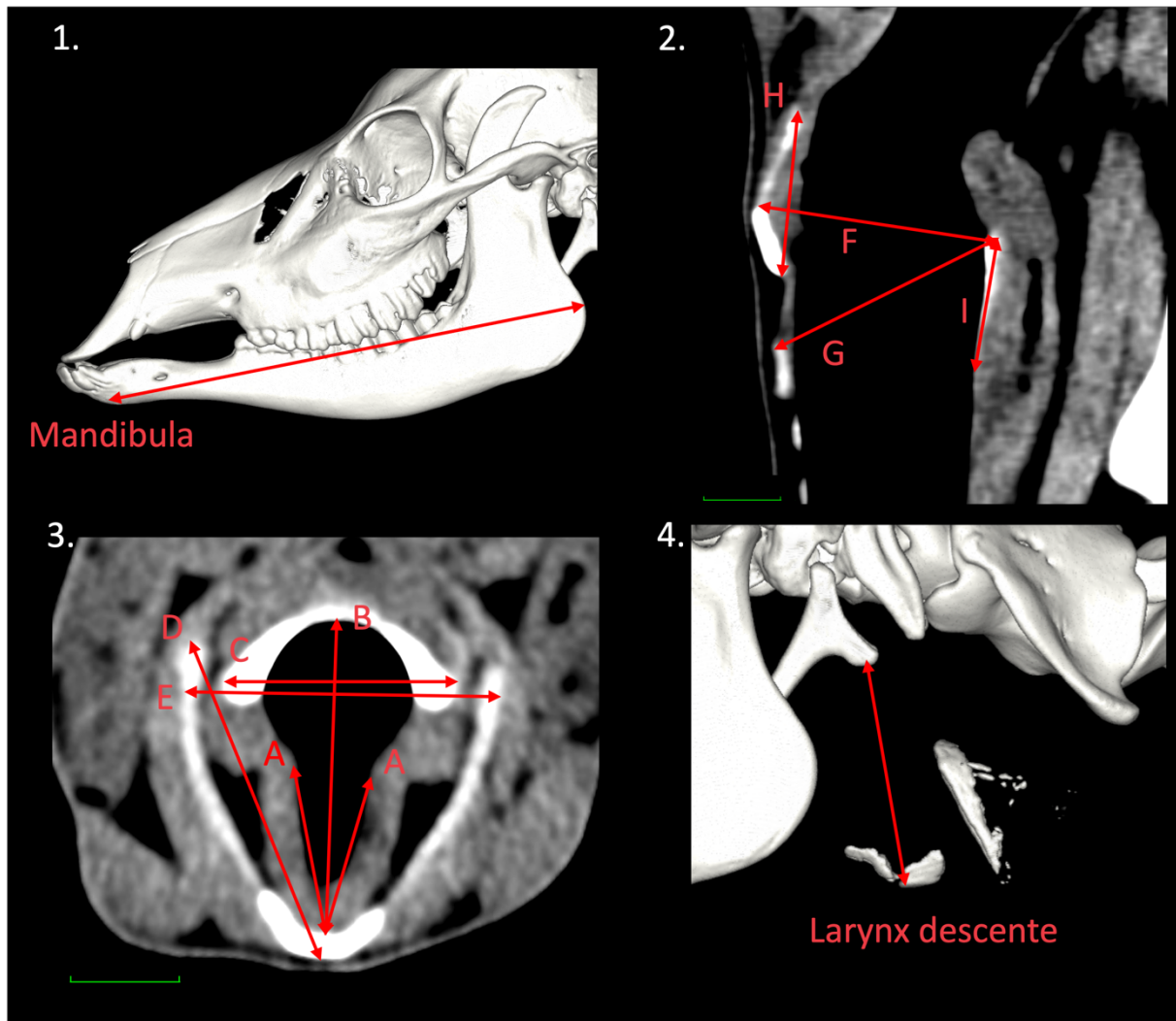


Figure 2. Anatomical measurements based on CT-scan of a sika deer. 1) Left side view of cranial bones to measure the mandible length 2) sagittal section at the laryngeal position 3) cross section at the vocal folds position; 4) left side view of stylohyoid bone and larynx cartilages to measure the descent of larynx. Scale: 1 cm.

For each CT-scan, we extracted the following measurements:

- Mandibula length, measured from the anterior mandibular border to the caudal mandibular border (Fig.1; Fig.2).
- Laryngeal descent, measured from the angle of the stylohyoid bone and the anterior prominence of the thyroid cartilage (Fig.1; Fig.2).
- Laryngeal measurements (Fig.1; Fig.2):
 - A: Vocal fold length left and right
 - B: Total glottal length measured from the interior cricoid cartilage to the interior thyroid cartilage
 - C: Maximal transverse width of cricoid
 - D: Thyroid lamina dorsoventral maximal length
 - E: Maximal width of thyroid
 - F: Dorso-ventral distance between thyroid and cricoid cartilages
 - G: Dorso-Ventral cricoid length

- H: Thyroid height
- I: Cricoid height

The mandibula length is commonly used as a proxy of animal size (Blant & Gaillard, 2004) to control for inter-individual size variation. To avoid using redundant measures, we calculated the mean between left and right vocal fold length (based on a reasonable assumption of lateral symmetry in vocal fold length).

Statistical analysis

To reduce our number of measurements to a smaller set describing the dimensions of the larynx, a Principal Component Analysis (PCA) was performed on the 9 laryngeal measurements. The coordinates on the axis that explained more than 40% of the variability were used as an index of laryngeal size (outcome variable called *laryngeal size*).

To test the effect of testosterone on each species, we used two Linear Mixt Model (LMM) to compare the laryngeal descent and laryngeal size in intact (*Intact*) and castrated without supplementation in testosterone (*No_ST*) males. Both LMM included the age (6 months or 18 months), the group type (*Intact* or *No_ST*), the species (sika deer or red deer) and all the interactions between the three variables, as fixed effect variables. The mandibula length was include in both models as a covariate variable.

To confirm the effect of testosterone in red deer males, we compared the laryngeal descent and laryngeal size in the four red deer groups (*Intact*, *No_ST*, *Phy_ST*, and *High_ST*). Both LMM included as fixed effect variables the age (6 months or 18 months), the group type (*Intact*, *No_ST*, *Phy_ST* or *High_ST*), all the interactions between these two variables and as a covariate variable the mandibula length.

All the statistical analyses were performed using SPSS for Mac version 28.0.1.0. Assumptions for normality of residuals and homogeneity of variances were verified using qqplot and the distribution of the residuals.

Results

After the first CT-scan at 6 months, deer were divided into 6 groups corresponding to the different treatments across species. Sika deer were divided in two groups: 5 uncastrated sika deer (*Intact*) and 5 castrated sika deer (*No_ST*). Red deer were divided into four groups: 5 uncastrated red deer (*Intact*); 5 castrated red deer without supplementation (*No_ST*); 3 castrated deer that received supplementation in testosterone at a physiological level (*Phy_ST*) and 4 castrated deer that receive testosterone supplementation at higher level than physiological level (*High_ST*).

One red deer stag from the *High_ST* group was diagnosed with a respiratory disease and died shortly after the second CT-scan and was not included in our results.

Testosterone level

Before treatments, at 6 months, male sika deer had an average testosterone level of 0.21 ± 0.03 ng/ml (mean \pm s.d.) (N = 10 sika deer) while male red deer had 0.88 ± 1.05 ng/ml (N = 16 red deer) of testosterone.

In both sika and red deer, castration inhibited the production of testosterone in non-supplemented castrated animals (*No_ST* groups) compared with intact animals (*Intact* groups) (Fig.3). In intact animals, testosterone increased by 0.98 ± 1.15 ng/ml in sika deer and 2.78 ± 2.4 ng/ml in red deer between the first and the second CT-scan.

The testosterone supplementation, which started in January 2020, in *Phy_ST* and *High_ST* red deer groups induced pics of testosterone (maximum of the mean pic 10 ± 0.14 ng/ml and 19.3 ± 4.25 ng/ml respectively) that decreased to similar testosterone levels measured in intact red deer.

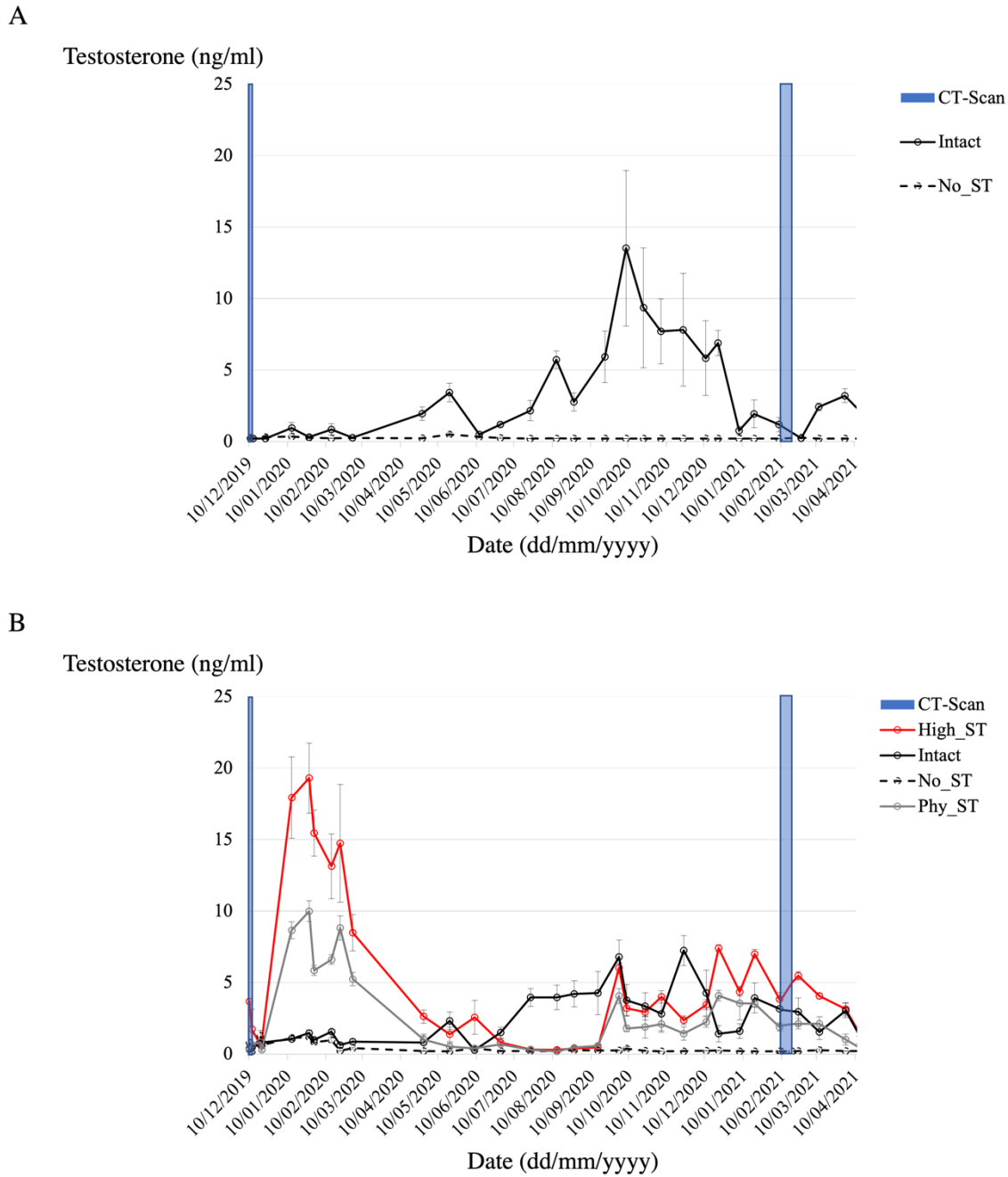


Figure 3. Testosterone level (mean \pm s.e.m) sampled from the 10 sika deer (A) and the 16 red deer (B) from the start of experiment the 10/12/2019 to the 10/04/2021. The first CT-scans were effectuated between the 10/12/2019 and the 13/12/2019. The second CT-scans were effectuated between 02/02/2021 and 18/02/2021. The supplementation in testosterone for the *Phy_ST* and *High_ST* red deer groups started in January 2020.

Effect of castration in red and sika deer males

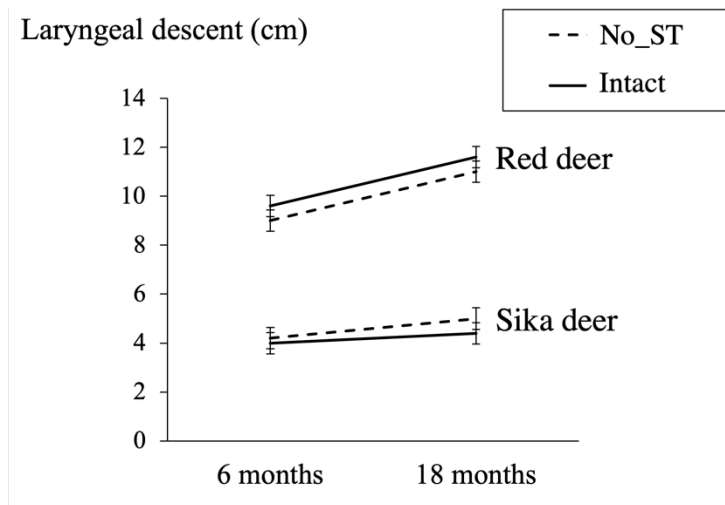


Figure 4. Laryngeal descent (mean \pm s.d.) in *Intact* and *No_ST* red deer and sika deer groups at 6 and 18 months. Laryngeal descent is the distance between the angle of the stylohyoid bone and the anterior prominence of the thyroid cartilage.

In order to test the effect of testosterone on laryngeal descent, we performed a LMM on the distance between hyoid bone and thyroid cartilage, measured from red deer and sika deer CT-scan at the age of 6 months (before treatment) and 18 months (one year after treatment). In this analysis only the *Intact* and *No_ST* groups were included to compare both species. LMM on the laryngeal descent showed a significant effect of the mandibula length ($F=5.16$; $p<0.05$) and a significant difference between species ($F = 5.78$; $p<0.05$). Red deer have a larynx more descended than sika deer (Fig.4).

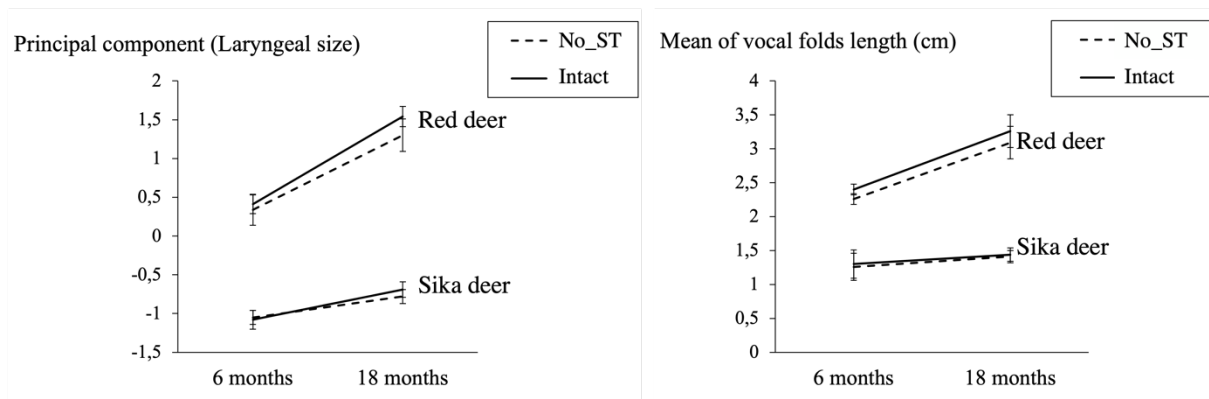


Figure 5. Laryngeal size (left) and vocal fold length (right) (mean \pm s.d.) in *Intact* and *No_ST* red deer and sika deer groups at 6 and 18 months. Laryngeal size is the coordinates from the x-axis of PCA on the 9 laryngeal measurements. Vocal fold length corresponds to the mean between the left and right vocal folds.

The effect of testosterone on laryngeal development was tested using the 9 laryngeal measurements measured from red deer and sika deer CT-scan. The PCA on the 9 laryngeal measurements indicated that the first axis of the PCA explained more than 98% of the total variability shown between individuals. All the variables were strongly correlated with this axis (>90%), high value on first PCA axis corresponds to high values of the 9 laryngeal measurements. To simplify the interpretation, the coordinates of the first axis were renamed *laryngeal size*. A LMM performed on the *laryngeal size* Principal component showed a significant effect of mandibula length ($F = 4.79$; $p < 0.05$) and of the interaction between species and age ($F = 34.75$; $p < 0.001$). After correcting for individual size, laryngeal growth between 6 months and 18 months is higher in red deer than in sika deer (Fig.5).

Vocal folds growth was significantly greater in red deer (at 6 months: 2.33 ± 0.23 cm; at 18 months: 3.27 ± 0.18 cm (mean \pm s.d.)) than in sika deer (at 6 months: 1.22 ± 0.11 cm; at 18 months: 1.49 ± 0.09 cm) (LMM; interaction species/age; $F = 33.14$; $p < 0.001$) however there was no effect of the interaction between group and age ($F = 0.05$; $p = 0.82$). Thus, the castration did not appear to affect the growth of the vocal folds.

In addition to the significant interaction effect of species and age reported for all of the laryngeal measurements, we found a significant effect of the interaction between groups and age on cricoid height (LMM; $F=6.27$; $p<0.05$) and maximal length of the thyroid lamina (LMM; $F=4.56$; $p<0.05$). Independently of the species, the growth of the cricoid height and the maximal length of the thyroid lamina were higher in the *Intact* groups than in the castrated *No_ST* groups (Fig.6).

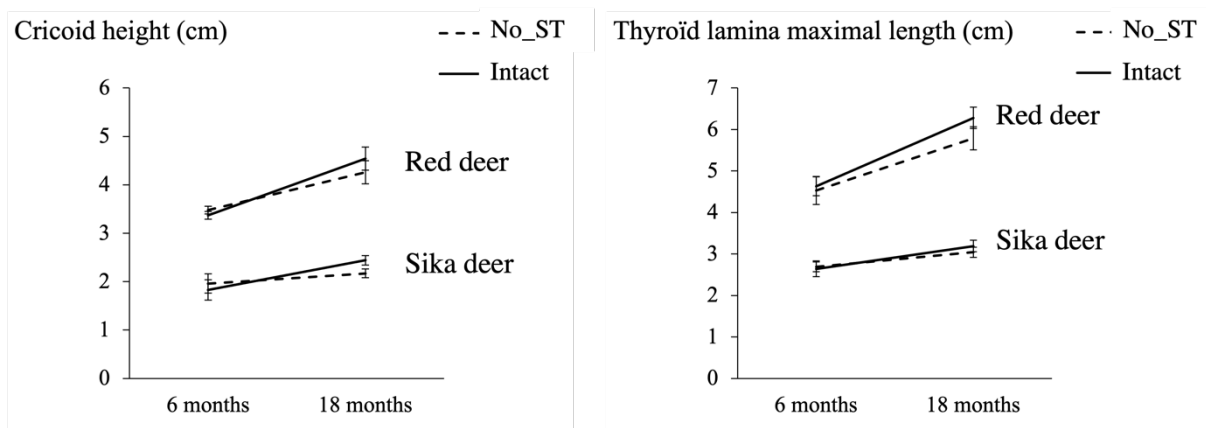


Figure 6. Mean \pm standard deviation (s.d.) of cricoid height (left) and maximal length of the thyroid lamina (right) in *Intact* and *No_ST* red deer and sika deer groups at 6 and 18 months.

Effect of supplementation in castrated red deer males

To compare the four red deer groups including *Phy_ST* and *High_ST* groups we conducted a LMM on the laryngeal descent while controlling for mandibula length. This analysis did not show an effect of group ($p = 0.35$), of age ($p = 0.94$) or of the interaction between age and group ($p = 0.99$). The castration and the supplementation in testosterone did not affect the descent of the larynx in red deer.

The PCA on the 9 laryngeal measurements in red deer indicates that the first axis of the PCA explains more than 89 % the total interindividual variability. All the variables are strongly correlated with this axis ($> 90\%$): a high value on the first PCA axis corresponds to high values of all 9 laryngeal measurements. LMM on this component shows only a significant effect of age (LMM; $F = 31.67$; $p < 0.001$) with a bigger larynx at the age of 18 months. LMM on each red deer laryngeal measurement showed similar results without group effect. Thus, neither the castration nor the supplementation in testosterone affected the general laryngeal growth (including vocal fold length).

Discussion

We tested the effect of testosterone on laryngeal development in two species of polygynous deer with contrasted vocal communication systems, the red deer and the sika deer. Our anatomical measurements at 6 and 18 months showed that at the age of 18 months, sika deer and red deer larynges are different, with a lower laryngeal position and greater laryngeal growth (specifically relatively longer vocal folds) in red deer. This is consistent with post-mortem observations made in adults of both species (Frey et al., 2021). Surprisingly, we only found a limited effect of castration on two laryngeal measurements (greater cricoid height and maximal length of thyroid lamina in intact males) in both species. Even more surprisingly, in red deer, there were no differences in laryngeal development between the four experimental groups tested (i.e., *Intact*, *No_ST*, *Phy_ST*, *High_ST*), indicating no effect of testosterone between 6 and 18 months on laryngeal development. Below we discuss these observations and their implications for our understanding of interspecific differences in vocal signals, and putatively androgen-mediated individual differences in laryngeal growth and associated acoustic traits (f_0 and formant spacing).

Influence of testosterone on the larynx growth

We found an effect of testosterone on two specific measurements: the cricoid height and the maximal length of the thyroid lamina. This result is consistent with a previous study in sheep (Beckford et al., 1985). However, contrary to our prediction we did not find any effect of testosterone on vocal fold length. Enlarged cricoid height and the maximal length of the thyroid lamina are not always correlated with the vocal fold length (e.g., Beckford et al., 1985; Frey & Riede, 2003). Thus, testosterone can influence the growth of laryngeal external parts without affecting the vocal fold length. Other anatomical parameters that we did not measure in this study can also be affected by testosterone such as the angle and the ossification of the thyroid cartilage as it was found in sheep and humans (Beckford et al., 1985; Mupparapu & Vuppalapati, 2005). Moreover, other hormones can also be implicated in larynx growth. Indeed, estrogen has been found in the female laryngeal area (Brunings et al., 2013; Holt et al., 1986; J. M. Kim et al., 2020; Narbaitz et al., 1980) and thus may have an important role in the laryngeal sexual dimorphism found in red deer and sika deer (Frey et al., 2021). Sexual selection can explain a part of the duality between the laryngeal growth and the relatively small vocal folds in red deer. Indeed, contrary to sika deer, in which the vocal behavior does not play a strong role in mate choice (Wyman et al., 2014), red deer are subject to a strong sexual selection in which estrous red deer hinds are more attracted by high pitch roar (Reby et al., 2010) usually associated to small animals but they are also attracted by lower formant spacing which is positively correlated with body size (Charlton et al., 2007; Fitch, 1997).

Influence of testosterone Vs body size on the laryngeal descent

We confirmed previous observations that red deer males have a lower laryngeal position than sika deer males (Frey et al., 2021). Interestingly, this difference persists even after a correction for individual size (using the mandibula length as a proxy). This result suggests that the greater body size of red deer males alone cannot explain the difference in laryngeal position between species. While absent in sika males, the laryngeal descent, characteristic of red deer and fallow deer males (Frey et al., 2021), induces an elongation of the vocal tract and a decrease in formant spacing in vocalizations (Charlton et al., 2011; Fitch, 1997; Fitch & Reby, 2001). The ability to lower the larynx when animals are vocalizing has been shown to support the exaggeration of perceived size in these two species (fallow deer: McElligott et al., 2006; red deer: Reby et al., 2005). Accordingly, size related formant variation is used by red deer hinds to select large mates

(Charlton et al., 2007) and also mediate competition males (Reby et al., 2005). In contrast, sika deer hinds show no clear preference between rutting calls given by conspecific versus heterospecific males, suggesting that acoustic signals likely play a weak role in mate choice in this species (Wyman et al., 2014). Thus, this sexual selection on acoustic parameters could drive the exaggeration of the laryngeal descent in red deer males contrasted by the high laryngeal position in sika deer.

In adult red deer, Frey et al., (2021) found a sexual dimorphism in laryngeal position with a more descended larynx in males. We hypothesized that testosterone could regulate the descent of the larynx in red deer male and predicted that the larynx would not descend in castrated males. However, our results in red deer groups with different levels of testosterone supplementation showed that although intact and supplemented groups had a higher testosterone level than non-supplemented castrated groups at 18 months, there were no significant differences in laryngeal position between these groups. Moreover, when we focused on the four red deer groups, we only found an effect of age when not controlling for animal size (data not shown). This suggests that differences in laryngeal position between red deer and sika deer appear before 6 months age and that, after 6 months, laryngeal descent in red deer males follows general body size growth (at least until 18 months). Future analyses are already planned to test if our results persist after 18 months.

As red deer lower their larynx when producing calls, it will be interesting to assess in post-mortem investigations whether testosterone has an effect on the length and elasticity of the thyrohyoid ligament involved during laryngeal movements accompanying call production. Indeed, Frey et al., (2021) found a higher potential for elongation in red deer males than in females, contrasting with an absence of elasticity in both sika deer sexes. Furthermore, species with a higher dimorphism in the laryngeal descent and size such as fallow deer (*Dama dama*) would be an excellent model for further investigating the effect of testosterone on larynx size and position.

Acknowledgements

We thank the team of animal keepers from the zoopark la Haute-Touche in charge of the animals presented in this study. We thank the veterinary team from Beauval zoopark in charge of the CT-scan.

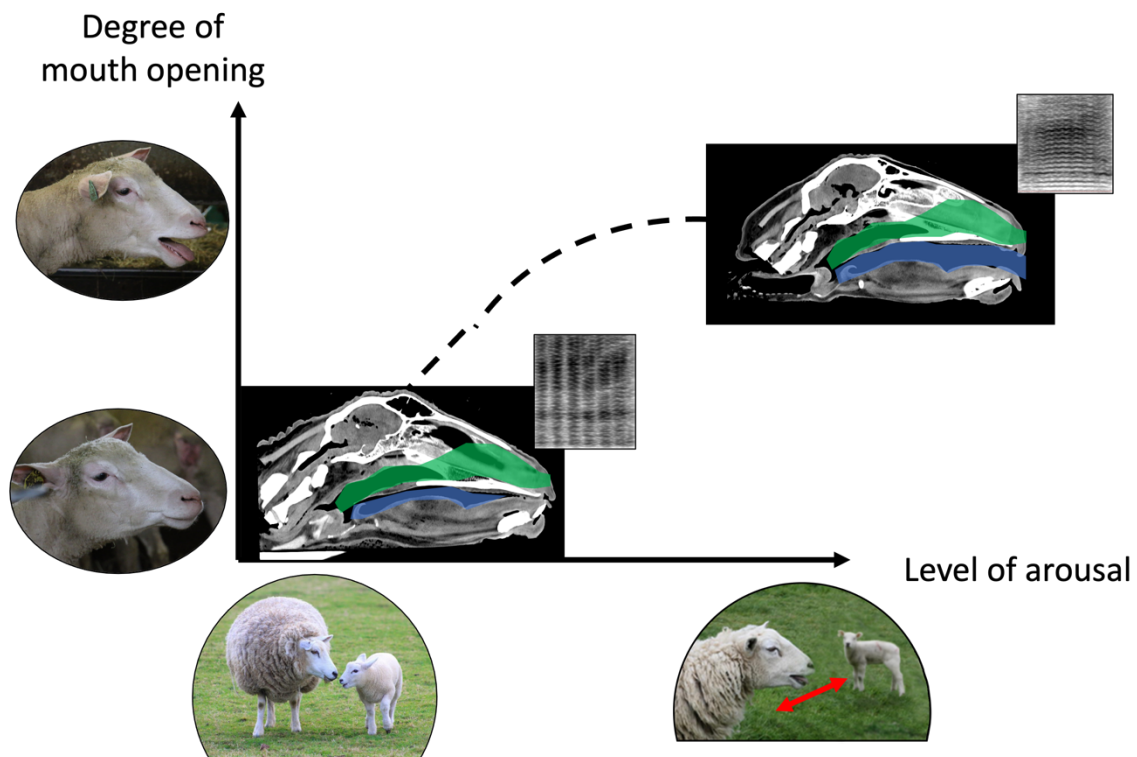
This work was carried out thanks to the financial support of the IDEXLYON Project of Lyon University within the framework of the *Programme Investissements d'Avenir* (ANR-16-IDEX-0005).

Chapter II: Mouth opening and expression of arousal in sheep



I explained in the introduction that the most common position of larynx in animal kingdom is the high position. This was assumed to restrict the vocal tract articulation as the pharyngeal part is reduced. However, basic articulation gestures such as mouth opening has been observed in several species including sheep. In this chapter, we combined behavioral analyses, acoustic analyses and post-mortem investigations to address the following questions:

- Does the degree of mouth opening encode arousal level in ewes?
- How is this information modulated by vocal tract modulation?



My results show that, even with a relatively simple vocal tract, ewes are able to modulate their vocal tract shape to encode their arousal level.

Article 2: Mouth opening as an indicator of the level of arousal in ewes

Lucie Barluet de Beauchesne¹, Nolwenn Risser¹, Léo Papet¹, Joël Gilbert^{2,†},
Matthieu Keller³, Frédéric Sèbe¹, Nicolas Mathevon^{1,4} and David Reby^{1,4}

¹ ENES Bioacoustics Research Laboratory, CRNL, CNRS, Inserm, University of Saint-Etienne, France

² Laboratoire d'Acoustique de l'Université du Mans, CNRS, 72085 le Mans, France

³ Laboratoire de Physiologie de la Reproduction & des Comportements, INRA/CNRS UMR 7247, Université de Tours, FRANCE

⁴ Institut Universitaire de France

[†]Deceased

This article was submitted to Journal of Experimental Biology in December 2022

Abstract

While the capacity for articulation is generally considered to be highly limited in non-human mammals, simple vocal tract movements are increasingly identified as contributing to the vocal diversity of a variety of species. Here we investigate the potential for such basic articulation in domestic sheep, where ewes give two main types of vocalizations when calling their lamb: “low” bleats which are produced with a closed mouth when the lamb is at close range and “high” bleats, produced with an open mouth when the lamb is away from its mother. First, we test whether mouth opening encodes information about the arousal state of the ewes. We show that ewes produce more open-mouth bleats and bleats with a higher degree of mouth opening as their level of arousal increases. We also demonstrate that this change in ewes’ mouth configuration affects the geometry of the vocal tract and thus formant frequencies. Finally, we show that bleat production in ewes simultaneously involves the oral and the nasal tracts in both open and closed mouth bleats. This articulation and the resulting phenotypic complexity allows the ewes to modulate the different information contained in their vocalization. We argue that documenting such articulatory gestures in species that lack speech contributes to understanding the selective forces that drive the diversification of articulation and thus phonetic complexity.

Introduction

Human speech relies on our unique ability to produce and control a rapid flow of phonetic sounds supported by complex articulatory maneuvers, such as tongue, mandible or lips movements (Lieberman, 1991). Understanding the evolutionary origins of these abilities is one of the key goals of research on vocal communication. Yet, decades of investigations have failed to identify animal equivalent of human speech articulation (Boë et al., 2019; Lieberman, 2012). Indeed, while some birds (e.g., Grey parrot (*Alex*): Pepperberg et al., 1991; see review: Beckers et al., 2004) or even mammals (e.g., harbor seal (*Phoca vitulina*) (Hoover): Ralls et al., 1985; elephant (*Elephas maximus*) (Koshik): Stoeger et al., 2012; beluga (*Delphinapterus leucas*) (Logosi): Eaton, 1980) manage to mimic speech sounds using various strategies, this vocal imitation usually involves a relatively narrow range of human speech sounds. For instance, Koshik the elephant (Stoeger et al., 2012) could only produce five Korean words using its trunk to modulate the resonant properties of its oral cavity, and Hoover the seal (Ralls et al., 1985) seemed to repeat relatively simple sentences such as “come over here”.

In non-human primates, there is some evidence of limited articulation abilities in rhesus monkeys, (*Macaca mulatta*: Hauser et al., 1993), ruffed lemurs (*Varecia variegata variegata*: Gamba & Giacoma, 2006) and Diana monkeys (*Cercopithecus diana*: Riede et al., 2005). Indeed, in these three species, vocalizers can shape their vocal tract to produce different call types. For instance, Diana monkeys are able to modulate their mouth closing, jaw movements and larynx movements to produce specific alarm calls depending on the predator species individuals detect (Riede et al., 2005). In rhesus monkeys, the opening of the mouth is correlates with changes in the dominant frequency but not with changes in the fundamental frequency, which suggest that rhesus monkeys use the change of vocal tract shape more than the modulation of vocal fold vibration rates to modulate the sound within a call (Hauser et al., 1993). Vocal tract modelling studies have shown that ruffed lemurs probably modify the configuration of their vocal tract to use either only the oral tract, only the nasal tract or both oral and nasal tracts depending on the call type produced (Gamba & Giacoma, 2006). Moreover, a study of living macaques combining magnetic resonance imaging (MRI) to determine their possible vocal tract geometries with acoustic modeling has demonstrated that primates vocal tracts should be able to perform the configurations required for producing a range of vocalizations comparable to human vowel and consonants (Fitch et al., 2016). Specifically, their vocal apparatus should allow them to produce at least the five clearly distinguishable vowels present in the words: “bit”, “bet”, “bat”, “but”, and “bought” as well as the stop

consonants: /p/, /b/, /k/ and /g/. Similar articulation abilities have been reported in Guinea baboon (*Papio papio*) (Boë et al., 2017): acoustic analysis of the first two formant patterns observed in their natural vocalizations showed that they can produce the distinct formant frequency patterns observed in five human vowels: /i/, /æ/, /ɑ/, /ɔ/, /u/, each requiring a different tongue position.

To further investigate the evolutionary origins of articulatory abilities it is useful to extend the comparative method to a wider range of non-primate species and to investigate the circumstances under which specific articulatory abilities evolve, by e.g., identifying such abilities in non-primate mammals, describing their mechanisms and understanding their function. This approach can benefit from a unified framework for the study of the mammal vocal signals: the source-filter theory of voice production. According to this theory, the vocal folds in the larynx (source) and of the vocal tract cavities (the filter) are involved independently in the vocal production: when the air flow coming from the lungs passes through the glottis (the air space between the vocal folds) the vocal folds vibrate and the speed of this vibration produce the fundamental frequency (hereafter f_0) (Behrman, 2021; Ohala, 1983; Titze, 1994). This sound is then propagated through the vocal tract, which potentially includes all the air cavities from the larynx to lips and /or nostrils. The vocal tract acts as a filter by amplifying frequencies called formants that will be radiated out of the mouth and/or nostrils (Fant, 1960). The center frequency of each formant is determined by the resonant characteristics of the vocal tract, and each modification of the vocal tract geometry (e.g. compression, elongation) will induce the modulation of specific formant frequencies (Schroeder, 1967).

One way to study basic forms of articulation is to observe the gestures that animals perform naturally when they vocalize. To do this, cineradiographic investigations allow the quantification of vocal tract modulations in vocalizing subjects. Cineradiography was first used in humans and consisted in performing multiple x-ray acquisition to reconstitute movements (Lieberman & Blumstein, 1988). Cineradiographic observations (x-ray video) of dogs (*Canis familiaris*), goats (*Capra hircus*), pigs (*Sus scrofa*) and cotton-top tamarins (*Sagunius Oedipus*) have been used to demonstrate that the larynx descends and the velum is raised to disengage the nasal cavity and, thus, favor of the oral cavity for the production of loud calls (Fitch, 2000). While this method allows articulation to be observed in real time when an animal vocalizes, animals must be habituated to the experimental setting, it cannot be applied to large animals and exposes the animals to the risks associated with radiation (Lin, 2010).

Alternatively, we can predict formants produced by alive animals based on the morphology of their vocal tract. Indeed, a simple vocal tract can usefully be approximated as a cylindrical tube closed at the glottis, open at the mouth/nostril, with the length of the vocal tract and the diameter corresponding to the average diameter of the cross-sectional areas along the vocal tract (Stevens, 2000). To add accuracy in the model, the vocal tract can also be modelled as a series of multiple tubes with diameter corresponding to each cross-sectional area along the vocal tract and the length corresponding to the distance between two cross-sectional area (Stevens, 2000). This method, called computational modeling, computer-modeling or acoustic modeling has been widely used in the literature on human speech to link each vowels (specific formant pattern) with specific vocal tract shape (Lieberman, 2012). The geometry of the vocal tract can be assessed using different scanning methods such as magnetic resonance imaging (MRI) or computer tomography scans (CT-scan) that allow accurate estimation of the cross-sectional areas along the vocal tract midline (e.g., MRI: Baer et al., 1991; Kim et al., 2009; CT-scan: Perrier et al., 1992; Sundberg et al., 2007). The first application of this method to a non-human mammals was performed on domesticated cat (*Felis catus*) where Carterette et al., (1979) demonstrated that an acoustic model simulating the acoustic properties of a simple cylindrical tube closed at one end could be used to predict formant frequencies in cat vocalizations. Since, acoustic modelling has been successfully applied to understand, for instance, the involvement of the nasal tract in ruffed lemur (*Varecia variegata variegata*: Gamba & Giacoma, 2006) and fallow deer (*Dama dama*: Reby et al., 2018) vocalizations and the involvement of the mammalian laryngeal air sacs on the formant production (Riede et al., 2008). These studies were focused on the role of complex anatomies on formant production, however, to our knowledge very few studies have focused on the effect of simple articulation such as the opening of the mouth that contributes to human speech articulation (Shipley, Carterette and Buchwald, 1991).

Another approach for studying vocal production is the manipulation of cadaveric specimens which has so far almost exclusively used excised larynges to study the production of source-related components of acoustic signals (for a review of this methodology see Garcia & Herbst, 2018). This method consists in positioning the larynx at the extremity of a tube blowing air while the vocal folds are adducted, either with prongs or strings and pulleys, to close the glottis and generate the oscillation of the vocal folds (Garcia & Herbst, 2018). By modulating the air pressure and/or the longitudinal tension of the vocal folds, this approach allows to study the

mechanical and acoustical factors that induce the modulation of the fundamental frequency (Garcia & Herbst, 2018).

In order to study the effects of the shape of the vocal tract and of its modulation, it is also possible to use whole head and neck of specimens. To the best of our knowledge, only one study (Ferrein, 1741) has previously used whole head of specimens to understand the mechanisms underlying vocal production in mammals. While this approach allowed Ferrein to conclude that the whole head was necessary to reproduce mammal vocalisation (Ferrein, 1741) he only focused on the sound produced in a single vocal tract configuration and did not investigate the effects of specific vocal tract shapes on formant modulation.

Here, in order to investigate basic articulation gestures and to determine their function, we apply the methods cited above (i.e., observation of natural vocalization, acoustic modelling based on CT-scan, manipulation of head specimens to do post-mortem recordings) to the vocal production in sheep. This model species has a relatively simple vocal repertoire consisting of three main types of vocalizations: snorts produced by short expulsion of air through nostrils in threatening context, and two types of bleats, the “low” (here after called *closed mouth bleats*) and “high” (here after called *open mouth bleats*) bleats, where the opening of the mouth has been hypothesized to follow the level of arousal of the caller (Dwyer et al., 1998). In the context of parturition, these bleats contribute to the individual recognition of the mother by the lamb (Sèbe et al., 2007, 2010). The closed mouth bleats are produced by the ewe within 24h after parturition and associated with physical contact between the mother and its lamb (Terrazas et al., 2002). While the open mouth bleats are mostly produced between 24h and 6 days after parturition and produce in context of separation between the mother and its lamb (Sèbe et al., 2007, 2010) and in a long distance individual recognition context (Sèbe et al., 2008, 2010). Although, open mouth and closed mouth bleats code for individual information, these two types of vocalization are known to be acoustically distinct by their fundamental frequency higher in open mouth bleats (mean [lower quartile – upper quartile]: 159.6 Hz [139.6 – 167.9]) than in closed mouth bleats (125.1 Hz [109.7 – 140.2]) (Sèbe et al., 2010). Here we focus on investigating the role of mouth opening in the modulation of formant patterns, in order to see if this mouth articulation gesture allows the ewes to modulate the different information contained in their vocalization such as arousal level.

First, we set out to confirm that the arousal level of ewes drives the degree of mouth opening during bleating and second, we investigate how this impacts the formant patterns and formant dynamics in the produced bleats. To investigate the relationship between arousal and vocal tract modulation, we perform separation experiments between mothers and lambs in which we analyze the correlation between arousal level and degree of mouth opening during bleat production by ewes. While physiological measurements, such as heart rate measurement or body-surface temperature, can be used to assess the arousal level of ewes, these methods have the inconvenient of being invasive and of requiring a long period of habituation (Reefmann, Bütikofer Kaszàs, et al., 2009). We thus decided to estimate the arousal level using established behavioral parameters of stress level in sheep (Poindron et al., 1994).

In order to study the impact of these basic articulatory gestures on vocal tract resonances, we compared formant frequencies in the bleats produced in three mouth configurations: closed mouth, slightly open and widely open. We predicted that open mouth bleats would have higher fundamental and formant frequencies than closed mouth bleats and that slightly open mouth bleats would have formant frequencies in between formants observed in closed mouth bleats and formant observed in widely open mouth bleats. To determine the mechanisms behind the different observed formant patterns, we investigated the involvement of the oral and nasal tracts in the production of open-mouth and closed-mouth bleats using three post-mortem methods: airflow through the vocal folds (called *post-mortem recordings*), acoustic modelling from CT-scan images and acoustic impedance measures obtained with an impedance sensor. These methods allow us to observe or predict formant frequencies based on controlled vocal tract geometries and to compare these values to that measured in the open and closed mouth bleats recorded from living ewes in our separation experiment. Based on previous findings in another ungulate species (Reby et al., 2018), we predicted that open mouth bleats would involve both oral and nasal tracts whereas closed mouth bleats should involve the nasal tract only.

Material and methods

Ethics

This protocol was approved by the welfare committee *Structure de Bien-Être Animal* (SBEA) of the Unité Experimentale de Physiologie Animale de l'Orfrasière (UEPAO; <https://doi.org/10.15454/1.5573896321728955E12>) of INRAE of Tours-Nouzilly (France).

Separation experiment

Subjects

The animals involved in this study were Ile-de-France ewes bred at the UEPAO (veterinary agreement G 37-175-2) of INRAE Val-de-Loire, Nouzilly. They were kept outdoor from May until October and indoor for the rest of the year.

In order to collect acoustic data from animals in contexts of varying arousal, we performed separation experiments between ewe and lambs in INRAE Val-de-Loire, Nouzilly, in March 2021 on 22 primiparous 2-years-old ewes with one or two lambs aged 3 ± 1 days old (mean \pm standard deviation).

Protocol of the experiment

The experiment was performed in the ewe and lamb barn at a sufficient distance to avoid interaction between the tested animals and their group (around 12 m, Morgan et al., 1975). In order to habituate them to the setup prior to the separation experiment, the mother and her lamb were placed together, for 5 min, in an enclosure of 200 x 60 x 95 cm (L x l x h, 'ewes' enclosure) containing food, before they were allowed to re-join their group in the pen. The test, conducted at least one hour after this habituation phase, consisted of three conditions (*baseline*, *separation* and *reunion*) repeated twice with two separation distances (1 m and 3 m). First, the baseline consisted of placing the mother and lambs together in the ewe's enclosure for 2 min allowing for visual, auditory and olfactory contact. Then, for the separation, the lambs were moved to another enclosure (120 x 120 x 80 cm; 'lambs' enclosure) located at fixed distances of 1 m or 3 m from the mother's enclosure, where they remained for 1 min (Fig.1). While remaining in visual, auditory and potentially olfactory (Morgan et al., 1975), physical contact was no longer possible during this period. Lambs were moved back to their mother's enclosure for 2 min, before mother and lambs were allowed to re-join their group in the pen.

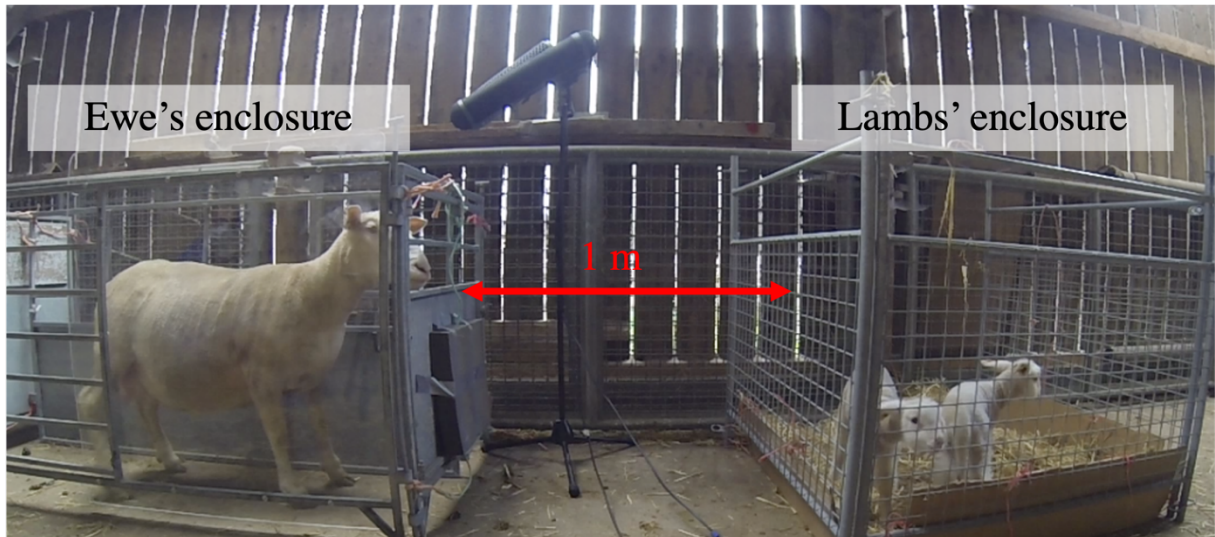


Figure 1. Experimental design used during the separation experiment. This photo represents the separation phase at 1 m with the mother in the ewes' enclosure and their lambs in the lambs' enclosure.

Video recordings were performed using a high-resolution camera (canon HD XA11, sampling frequency: 30 images/s) positioned perpendicular to the ewe's enclosure and at 1.6m from the Perspex screen. Two highly directive Sennheiser MKH70 microphones were placed at the top and the front of the ewe's enclosure and both connected to a Tascam DR-44WL recorder (Fig.1). As high bleats are produced with a higher amplitude higher than low bleats (Fig.2), we used different gain for each microphone to record both low and high bleats with a good signal/noise ratio and avoid saturation.

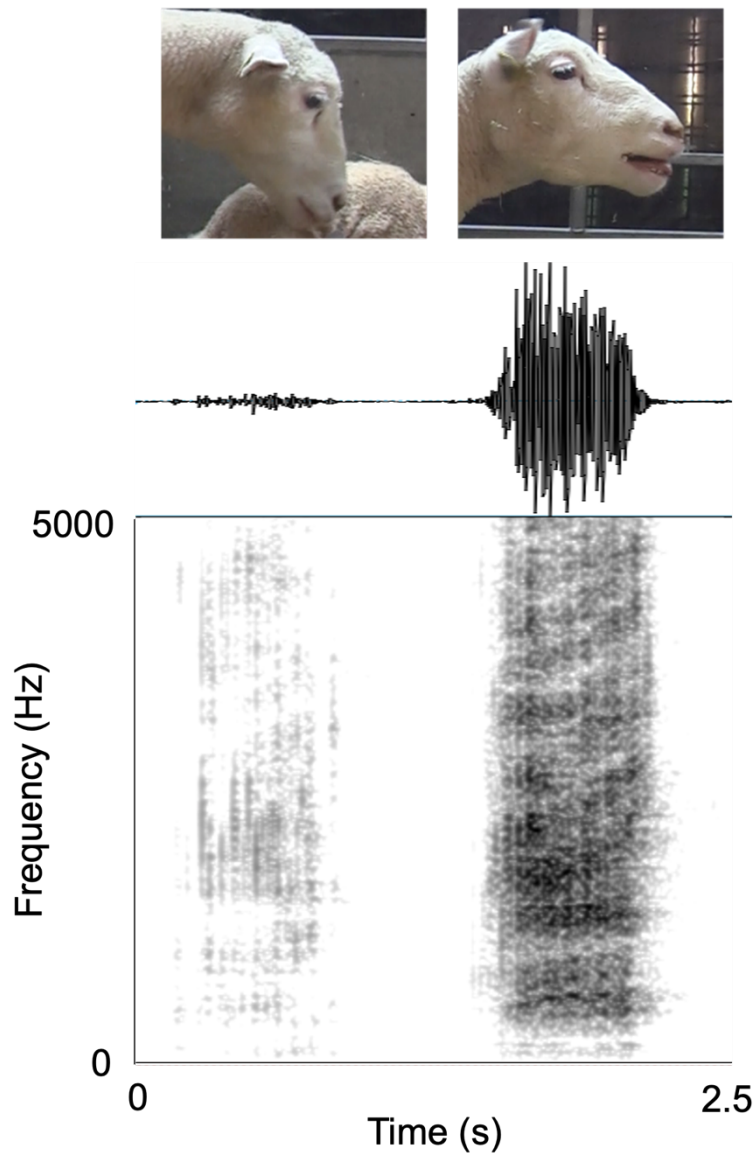


Figure 2. Example of one low bleat (left) and one high bleat (right) produced by an ewe during the habituation phase. Top: screen shot of the ewe producing the bleats. In the middle: oscillogram of each bleat. Bottom: spectrogram of corresponding bleat (window length: 0.05 s; time steps = 1000, frequency steps = 250, Gaussian window shape, dynamic range: 50 dB).

Evaluation of the arousal level

The temporary separation between the ewe and its lambs was expected to induce a mild stress compared to the baseline. This stress was characterized in previous studies by three components of the arousal level in sheep: the increased level of agitation, the head held high and the production of open mouth bleats (Poindron et al., 1994). Thus, to estimate afterward the arousal level of the ewes, we measure, for each condition of the separation experiment (baseline,

separation and reunion), the following behaviors, using Boris software (version 7.10.2; Friard & Gamba, 2016):

- *nCMB*: number of closed-mouth bleats per minute
- *nOMB*: number of open-mouth bleats per minute
- *tHup*: time with head held up divided by the condition duration
- *tMov*: time in movement (i.e., at least two steps forward or backward) divided by the condition duration

A PCA was performed on these four behavioral parameters and the axis that best represents data variance was chosen to represent the arousal scale (see *statistical analysis* for more details).

Selection of bleats

In order to investigate the hypothesis that arousal influences the degree of mouth opening, we selected open mouth bleats produced with the ewe's head was oriented perpendicular to the camera, allowing to accurately estimate the angle of the mouth. For the duration of each selected bleat, images were extracted from the video recording at the rate of 30 images/second. To calculate the angle of the mouth, three points were tracked using a semi-automated method on Matlab software (version 9.11.0): the nose, the eye and the lower jaw. The angle corresponding to maximal angle achieved during each bleat and the angle measured on closed mouth were retained. To assess the maximum mouth opening angle, the angle measured on closed mouth was subtracted from the maximal angle achieved during each bleat.

Extraction of the observed formants and of the fundamental frequency

The effect of mouth opening on formants was studied by comparing the first five formants extracted from the open mouth bleats and the closed mouth bleats recorded during the separation experiment. Open mouth bleats were separated into two categories according to the value of the maximal mouth opening angle: below the value of the median, bleats are categorized as *slightly open mouth bleats*, above the median, bleats are categorized as a *widely open mouth bleats*. The first five formants were measured from these sampled vocalizations using PRAAT (version 6.1.52). We used the *view spectral slice* function on which we applied a cepstral smoothing at 150 Hz and extracted the frequency peaks using both the function *To formant (peaks)* and a visual checking on the spectrogram (view range: 0-2500 Hz; window length: 0.05 s; dynamic range: 50 dB).

The mean of the fundamental frequency was extracted for each bleats using a custom-based in PRAAT. The script uses the function *To Pitch(cc)* with a detection range of 100-200 Hz, a time step: 0.01s and a silence threshold at 0.05.

Post-mortem investigation

Subjects

To investigate the mechanism underlying the production of closed and open mouth bleats, we conducted post-mortem experiments on two ewes (ewe 32128 and ewe 32142) that had not been involved in the separation experiment. These 6-year-old ewes were due to be sacrificed for reasons independent of our protocol (recurrent reproductive problems). Each head was obtained from the INRAE slaughterhouse (UEPAO, veterinary license G 37-175-2) by performing a section at the level of the first vertebra and were frozen at -20°C. The CT-scans were performed at the Pixanim platform of the PRC unit (https://www6.val-de-loire.inrae.fr/pixanim_eng/; INRAE Val-de-Loire) in May 2019 whereas post-mortem recordings and impedance measurements were performed at the ENES lab in June 2021 and May 2022.

Resonances predicted from CT-scan

Prior to scanning, each specimen was slowly thawed and flushed to remove debris and fluids into the oral and nasal cavities. During scanning, specimens were maintained in a plastic structure of 35 cm long, 28 cm large and with a “v” shape support to hold the head in order to avoid the collapsing of the trachea. The trachea was slightly pulled to disengage the larynx from the nasal cavity in order to simulate the natural laryngeal position of vocalizing bovids (Fitch, 2000). Both specimens were scanned in 2 configurations: first with a closed mouth (i.e., lips in contact) and second with an open mouth. Specimens were immediately refrozen at -20°C after scanning.

Following the methods described in Reby et al. (2018), we used the Osirix software (version 12.0) to extract the cross-sectional areas along the oral and nasal tract from scan. Contrary to Reby et al. (2018) and to reduce the time of processing, we used automatic detection with the function *grow region 2D/3D segmentation*, before manual adjusting with the function *ROI*. Moreover, the number of cross-sectional area was extracted every 5 mm instead of every 10 mm to increase the resolution (Fig.3).

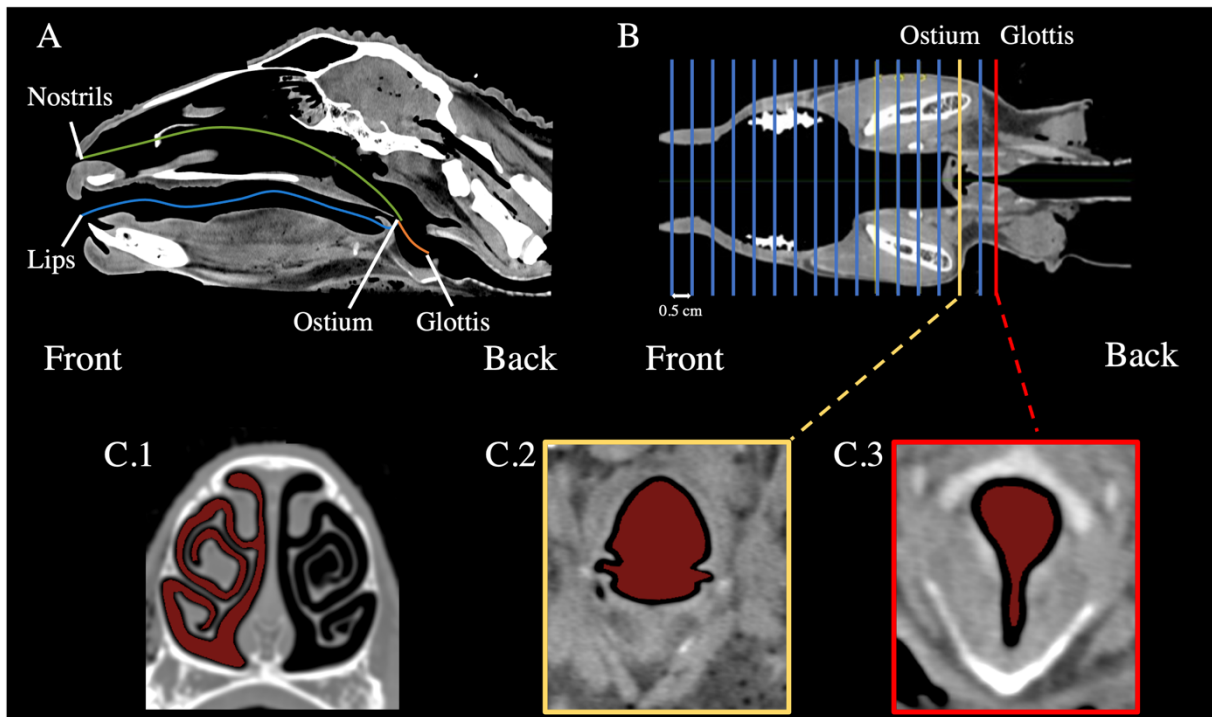


Figure 3. CT-scan of an ewe illustrating the method used to extract the cross-sectional areas. A) Median sagittal plane of the sheep head with the track of the oral cavity (blue line), the nasal cavities (green line) and the laryngeal tract (orange line) manually drawn. B) Position of the sections along the uncurved oral tract spaced by 0.5 cm (blue lines). C) Example of 3 transversal cross-sections: C.1. nostril, C.2. ostium, C.3. glottis.

As we were not able to open the mouth more widely than an angle of 6° for the ewe 32142 and 5° for the ewe 32128 (angle between nostril, eye and lower jaw) during the open mouth CT-scan because of the cadaveric rigidity, we estimated the cross-sectional areas along the oral and nasal tract as while simulating a mouth opening angle of 15° which corresponds to the average of the maximum mouth opening angles observed in natural bleats with the mouth widely open (mean \pm s.d.: $15.39 \pm 2.18^\circ$; $N = 10$ ewes; $n = 48$ bleats). We assumed that the degree of mouth opening should not affect the cross-sectional areas of the nasal cavities and the laryngeal tract, and thus for these cavities we used same cross-sectional areas as measured in the open mouth CT-scan. For the oral cavity, from the lips to the ostium, we assumed that the increase of the mouth opening angle would induce specific changes throughout the oral cavity and that these changes would be proportional to the degree of mouth opening. Thereby, we estimated the cross-sectional areas of the oral cavity using an extrapolation of the changes observed between closed mouth CT-scan and open mouth CT-scan, following the formula described in Fig.4.

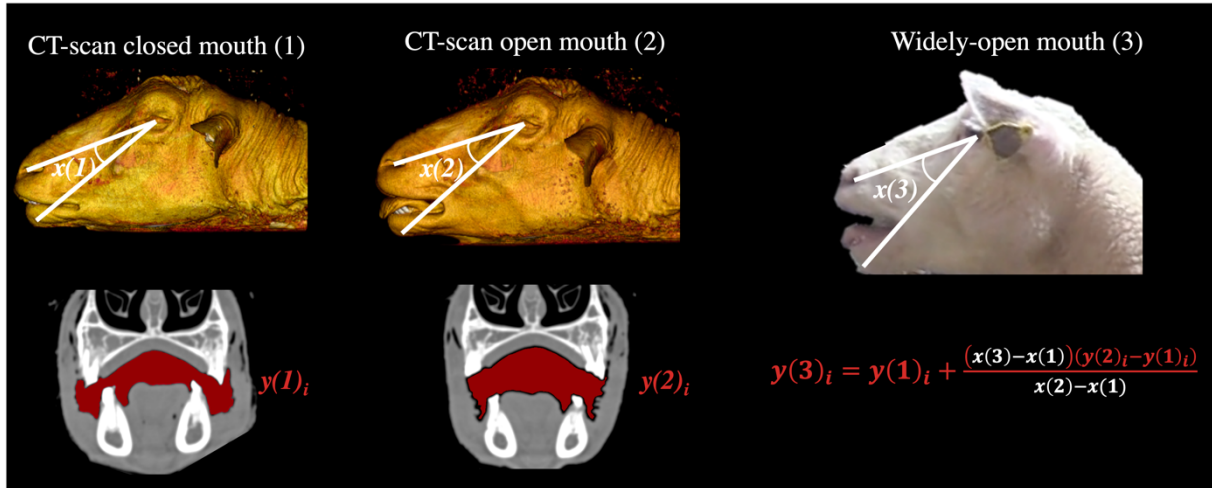


Figure 4. Method to extrapolate the cross-sectional areas to correspond to a widely open mouth angle. With x the angle of mouth and y the cross-sectional area at the i position and for the CT-scan with the mouth closed (1); the CT-scan with the mouth open (2) or the estimated cross-sectional area for the widely open mouth configuration (3).

For each specimen, three types of prediction were made based on three different configurations of the vocal tract:

- Only the oral tract (laryngeal tract and oral cavity) involved in production of the resonances, model called *oral tract only*.
- Only the nasal tract (laryngeal tract and nasal cavities) involved in the production of the resonances, model called *nasal tract only*.
- Both the oral and nasal tracts (laryngeal tract oral cavity and nasal cavities) involved in the production of the resonances, model called *both oral and nasal tracts*.

For each vocal tract configuration, the input impedance was estimated using the transfer matrix method (Chaigne & Kergomard, 2016) where the vocal tract is approximated as a series of cylindrical elements of fixed 0.5 cm thickness and variable cross-sectional areas. Based on the acoustic properties of tube nearly closed at one end (glottis) and open at the other end (lips and/or nostrils), the frequency peaks observed in the input impedance (hereafter called *predicted formants*) were used to predict center frequencies of the vocal tract resonances (Reby et al., 2018; Wolfe et al., 2009).

Post-mortem recordings

Each specimen was slowly thawed before experiment.

A metallic structure was built to maintain the head in a position suitable for the protocol (Fig.5). One experimenter was in charge of triggering vocal fold oscillation by applying manual pressure on the larynx. We monitored the fundamental frequency using real-time spectrograms produced by the Friture software (version 0.49) and adjusted the air pressure (minimum of 6 bar maximum of 8 kPa) to obtain a constant fundamental frequency in the range of the natural fundamental frequency of the species (i.e., [109.7 –167.9 Hz]; Sèbe et al., 2010). The second experimenter was in charge of positioning the mouth in the following configurations:

- oral open and nasal open
- oral closed and nasal open

The angle of mouth opening was extracted from videos recorded by a HC-VX980 Panasonic camera placed at 1 m from the side of the specimen. Sound was recorded at 1 m from the mouth of the specimen using a Behringer ECM8000 microphone. Formants were extracted for each configuration using PRAAT (smooth at 150 Hz, extraction of the frequency peaks and visual checking on the spectrogram).

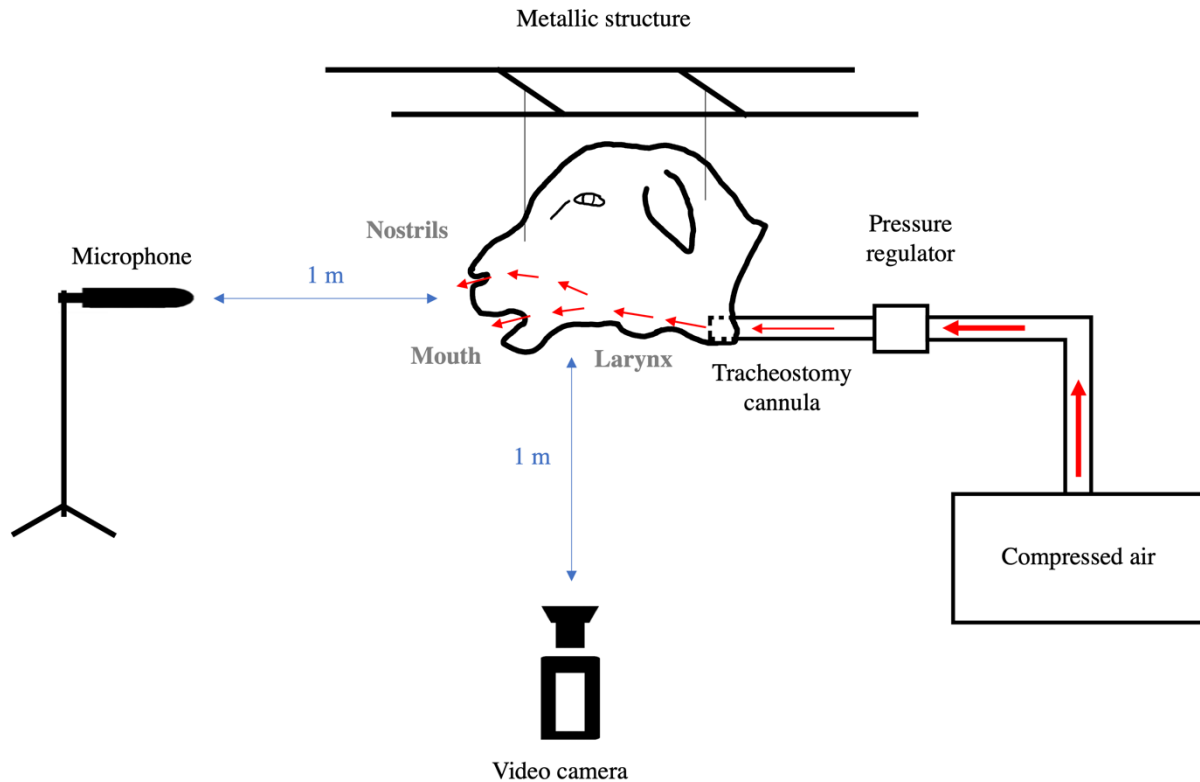


Figure 5. Schematic illustration of the experimental design of post-mortem recordings. The head is suspended using the metallic structure allowing the access to the larynx during the experiment. A compressor was used to generate an air flow (red arrow) into the trachea via a tracheostomy cannula. The air pressure was controlled via a pressure regulator.

Measures from impedance sensor

We used an impedance sensor, developed jointly by LAUM (Laboratoire d’acoustique de l’Université du Mans) and CTTM (Centre de Transfert de Technologie du Mans), to measure frequency resonances on two head specimens (ewe 32128 and 32142) slowly thawed before experiment. We 3D-printed a plastic adaptor (20 cm length and 1.5 cm diameter) to connect the sensor to the specimen glottis. The impedance sensor is a device constituted of a piezoelectric speaker as a sound source and two microphones measuring the sound pressure at the output and in the back cavity of sensor (Le Roux et al., 2012). The impedance sensor software (Capteur Z, version 2.7) was used to broadcast a swipe sound from 20 Hz to 6 kHz through the speaker of the impedance sensor, to record the pressure measurements and to calculate the input impedance based on the transfer function and taking into account the geometrical dimensions of the sensor and of the adaptor (Le Roux et al., 2012). The frequencies where the input impedance magnitude reaches peaks correspond to formants predicted by the impedance sensor.

Statistical analysis

All the statistical analyses were performed using R for Mac version 4.1.1. Assumptions for normality of residuals and homogeneity of variances were verified using qqplot and the distribution of the residuals.

To characterize the level of arousal for each ewe in the different conditions, a Principal Component Analysis (PCA) was performed on the four behavior variables: the number of closed mouth bleats produced per minute (*closed mouth bleats*); number of open mouth bleats produced per minute (*open mouth bleats*); percentage of time in movement (*time in movement*); percentage of time with head held high (*time head held up*). The arousal level was represented by the PCA axis that contributed for more than 40% of the variance.

The effect of arousal level on the maximum mouth opening angle in open mouth bleats was tested using a Linear Mixed-effect Model (LMM). The condition (baseline, separation, reunion), the distance (1 m, 3 m) and the arousal level were included in the model as fixed effect variables and the sheep identity as a random effect variable to account for repeated measures. This analysis was followed by a Tukey's HSD test, which allows to do adjusted pairwise comparisons (function *glht* in *multcomp* R package) to compare results of each condition.

To test whether the arousal level and/or the maximum mouth opening angle affects frequencies in bleats, a LMM was performed on the fundamental frequency and on each five first formants observed in natural bleats with three fixed effect variables: the arousal level, the maximum mouth opening angle and the interaction between the arousal level and the maximum mouth opening angle. The sheep identity was added as a random effect variable.

Results

Influence of arousal level on vocal production

The 22 ewes tested during the separation experiment produced open mouth bleats at least once across the experiment whereas only 17 ewes produced closed mouth bleats. The level of arousal was measured using a PCA on the four behavioral parameters (i.e., closed mouth bleats, open mouth bleats, time head held up and time in movement) for each condition of the separation experiment (i.e., baseline, separation, reunion) and distance (i.e., 1 m, 3 m) (Fig.6).

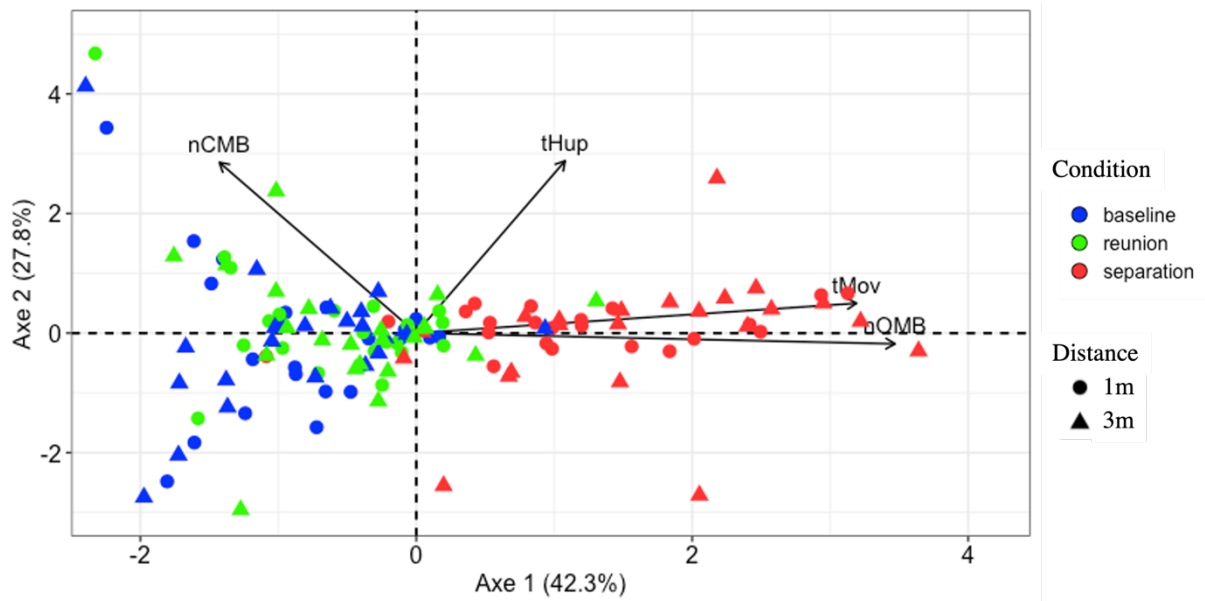


Figure 6. Results of the PCA on the behavioral response of the 22 ewes during each condition (baseline, separation, reunion) of the separation experiment. Experiments in which the separation between the mother and the lamb was at 1m are represented by a circle and experiments with a separation at 3m are represented by a triangular. Color corresponds to each condition baseline in blue, reunion in green and separation in red. *nCMB*: number of closed mouth bleats per minute; *nOMB* : number of open mouth bleats per minute; *tHup* : time with head held up divided by the condition duration; *tMov* : time in movement (i.e., at least two steps forward or backward) divided by the condition duration. Axis 1 explains 42.3% of the variability and mainly represented by *tMov* and *nOMB*. Axis 2 explains 27.8% of the variability and mainly represented by *nCMB* and *tHup*.

The axe 1 (x-axis) explains most of the behavioral data variability (42.3%). This axis is mainly represented by two types of arousal related behaviors: the time in movement (*tMov*; contribution at 40%) and the number of open mouth bleats (*nOMB*; contribution at 47%) both positively correlated with the first axis (Fig. 6). Furthermore, the LMM on the values of this axis indicates no significant effect of the interaction between the distance and the condition (LMM: $F(2, 110) = 1.37$; $p = 0.27$), nor of the distance ($F(1, 110) = 3.09$; $p = 0.08$) but a significant effect of the condition ($F(2, 110) = 140.52$; $p < 0.001$). The x-axis values are significantly higher during the separation than during the baseline (Tukey's HSD: $t = 15.09$; $p < 0.001$) and the reunion ($t = 12.96$; $p < 0.001$). The x-axis values are significantly higher during the reunion than during the baseline ($t = 2.13$; $p < 0.05$). Thus, the x-axis values reflect well the arousal level of the ewe induced by the separation with its lambs.

To test if the maximal mouth opening angle achieved during bleats was correlated with the emitter's arousal level, for each bleat for which we were able to analyze the mouth angle during open mouth bleats (N= 11 ewes, n= 45 bleats), we used the corresponding arousal value (x-axis of the PCA). The LMM testing the effect of the arousal level and of the experimental parameters (condition, distance) on the maximum mouth opening angle indicates no effects of distance (LMM: $F(1, 36) = 0.78$; $p = 0.38$), condition (LMM: $F(2, 40) = 0.005$; $p = 0.995$) or arousal level (LMM: $F(1, 43) = 3.62$; $p=0.06$). However, independently from condition and distance, there was a significant positive correlation between arousal level and maximum mouth opening angle (LMM: $F(1, 41) = 8.67$; $p<0.01$) (Fig. 7).

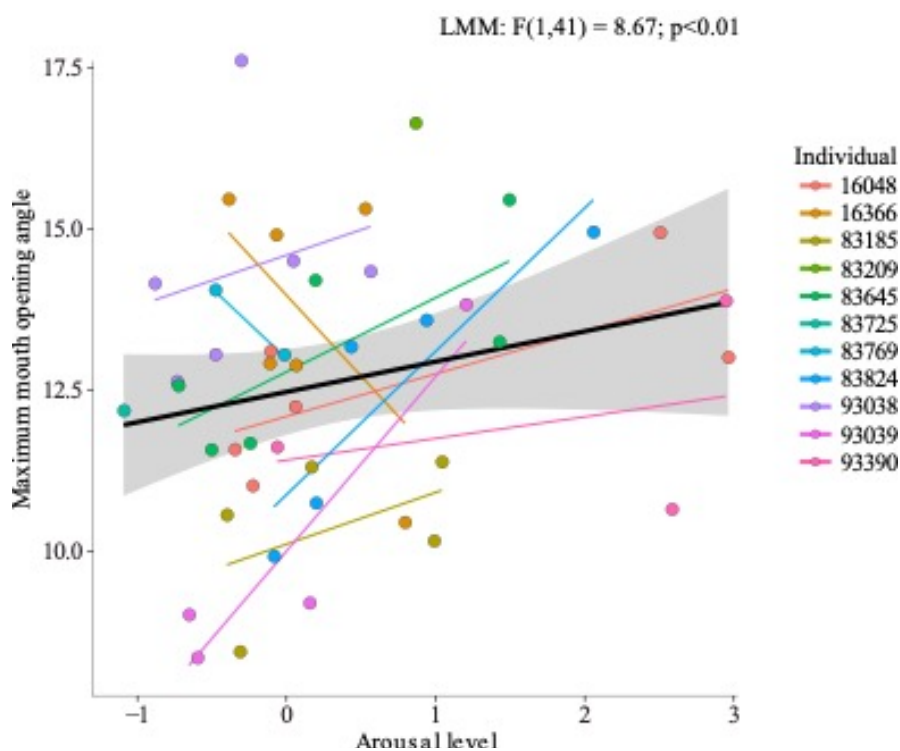


Figure 7. Correlation between the maximum mouth opening angle in the open mouth bleats and the arousal level (PCA x-axis) (N=11 ewes). The regression line is represented by the black line \pm the 95% confidence interval. Each open mouth bleat is represented by dot and line represent the regression for each individual represented by a specific color.

This result confirms that, in contrast with the other conditions (i.e., baseline, reunion), the separation induces a higher level of arousal characterized by a more movement activity and the production of more open mouth bleats. It also shows that ewes open mouth bleats are given with more widely opened mouths when they are more aroused, independently of the distance that separates them from their lambs.

Effect of mouth opening on the fundamental frequency and on the formants in natural bleats

In order to investigate the effect of mouth opening on formants, we measured the first five formants in 29 closed mouth bleats (N = 8 ewes) and 96 open mouth bleats (N = 11 ewes) produced during the separation experiment.

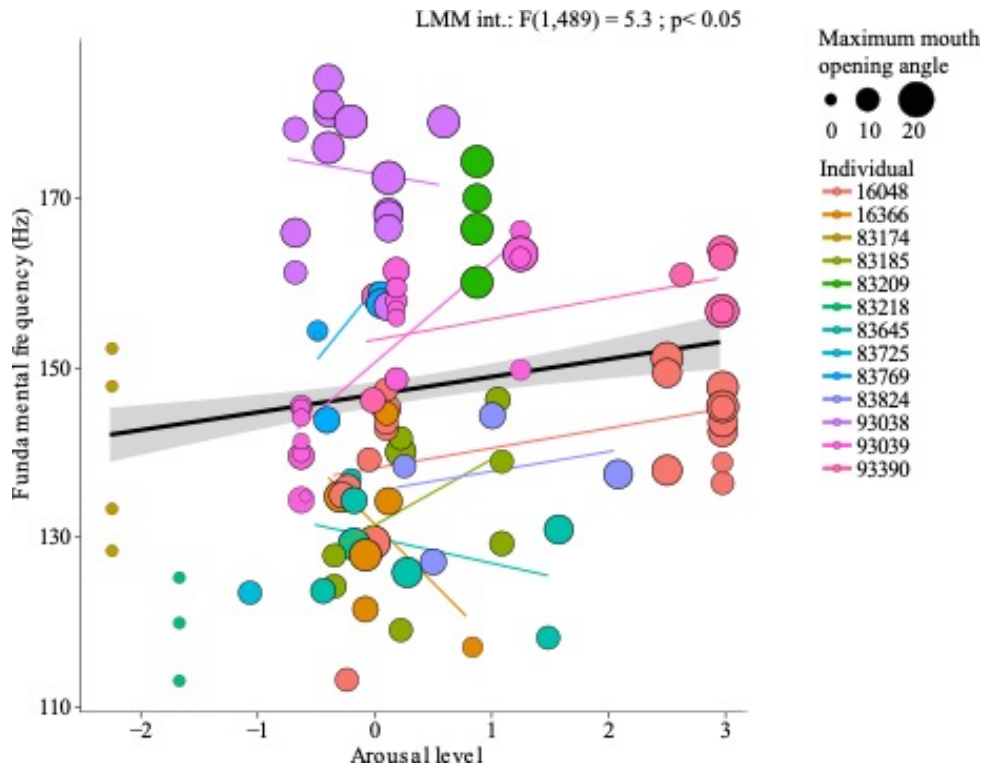


Figure 8. Correlation between fundamental frequency and the arousal level (PCA x-axis). The regression line is represented by the black line \pm the 95% confidence interval. Each open mouth bleat is represented by dot. The maximum mouth opening angle is represented by the dot size. Regression for each individual is represented by specific color line. The linear mixte model (LMM) was significant for the interaction between the arousal level and the maximum mouth opening angle $F(1,489)=5.3$; $p<0.05$.

Closed mouth bleats have a fundamental frequency at $126.2 \text{ Hz} \pm 11.5 \text{ Hz}$ (mean \pm s.d.) whereas open mouth bleats have a fundamental frequency at ...Hz \pm ... Hz. The fundamental frequency is positively correlated with the arousal level (LMM: $F(1,487)=22.27$; $p<0.001$), with the maximum mouth opening angle (LMM: $F(1,511)=18.26$; $p<0.001$) and the interaction between the arousal level and the maximum mouth opening angle (LMM int.: $F(1,489)=5.3$; $p<0.05$) (Fig.8).

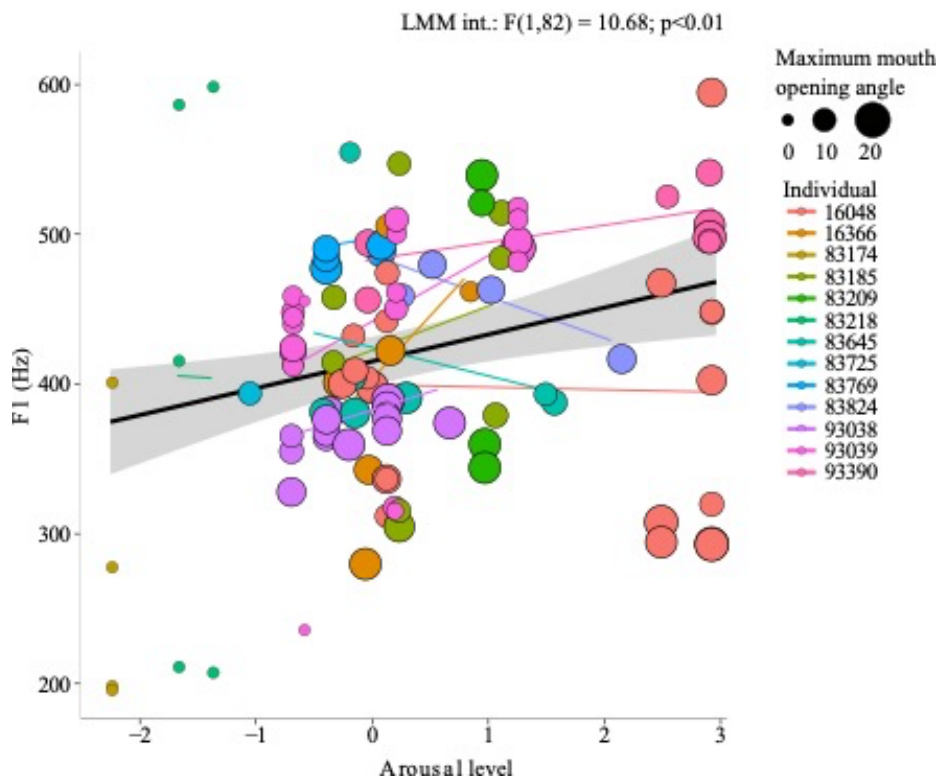


Figure 9. Correlation between the first formant frequency (F1) and the arousal level. The regression line is represented by the black line \pm the 95% confidence interval. Each open mouth bleat is represented by dot. The maximum mouth opening angle is represented by the dot size. Regression for each individual is represented by specific color line. The linear mixed model (LMM) was significant for the interaction between the arousal level and the maximum mouth opening angle $F(1,82)=10.68; p<0.01$.

Concerning the effect on the five first formants, we only found a significant effect of the arousal level (LMM: $F(1,41)=16.6; p<0.001$) and the interaction between the arousal level and the maximum open mouth angle (LMM int.: $F(1,42)=13.36; p<0.001$) on F1 (Fig.9).

The more the ewe is aroused the more it produces bleats with high f_0 and F1 frequencies.

Vocal tract geometry involved in the production of closed and open mouth bleats

To determine the mechanisms behind the different observed formant patterns, we tested the involvement of the oral and nasal tracts in production of open and closed-mouth bleats using three post-mortem methods: post-mortem recordings, acoustic modelling from CT-scan images and impedance measures obtained with an impedance sensor.

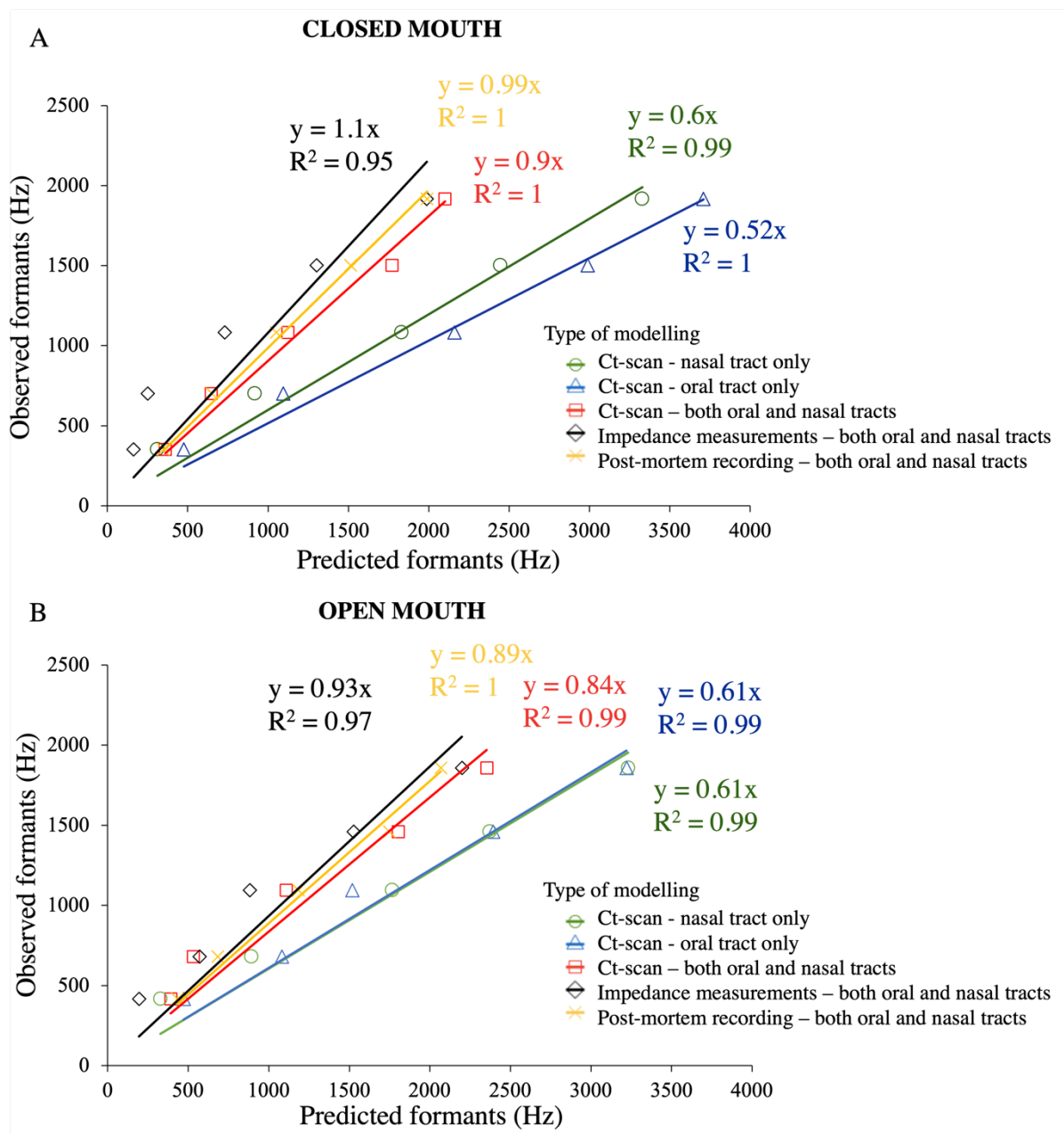


Figure 10. Correlations between formants observed in vocalization of ewes (y-axis) and resonances predicted by the 5 types of modelling (x-axis) in closed mouth (A) and open mouth (B) configurations. Observed formants are the average frequencies of the first five formants extracted from 29 bleats produced with closed mouth (N = 8 ewes) (A) and from 48 open mouth bleats produced with a mouth angle above 12.43° which corresponds to the median of angles observed in open mouth bleats (B). Predicted formants are the average frequencies of the first five peaks predicted by the 5 types of modelling based on the vocal tract of 2 ewes. The intercept was forced to pass by zero. The regression slope informs the fit of the formant spacing to that of the predicted resonances, the more the slope tends to 1.0 the more the overall formant spacing of the observed resonances matches that of the predicted resonances. The R² values provide the fit of formants frequencies pattern to the frequencies of the

predicted resonances, the more R^2 tends towards 1.0 the more the frequency pattern of the predicted formants matches that of the observed formants.

All the models reliably predicted the distribution of the first 5 formants' frequencies observed in natural ewes' vocalizations ($R^2 > 0.9$). The regression slopes indicate that the formants predicted from the impedance measurements, from the CT-scan geometry, or measured in the post-mortem experiment that include both oral and nasal tracts most reliably predicted formant spacing observed in both open and closed mouth bleats (Fig. 10). In contrast, models considering only the oral tract (slope for open mouth configuration = 0.61; slope for closed mouth = 0.52) or only the nasal tract (slope for open mouth configuration = 0.61; slope for closed mouth = 0.6) all underestimated the number of formants (i.e., predicted less formants than observed in the recorded live vocalizations). This strongly suggests that both oral and nasal tracts are involved to produce the formant spacing and pattern observed in both open and closed mouth bleats. Our CT scans show that, even when the mouth is closed at the level of the lips, the oral cavity remains hollow and our modelling confirms that this cavity contributes to vocal tract filtering.

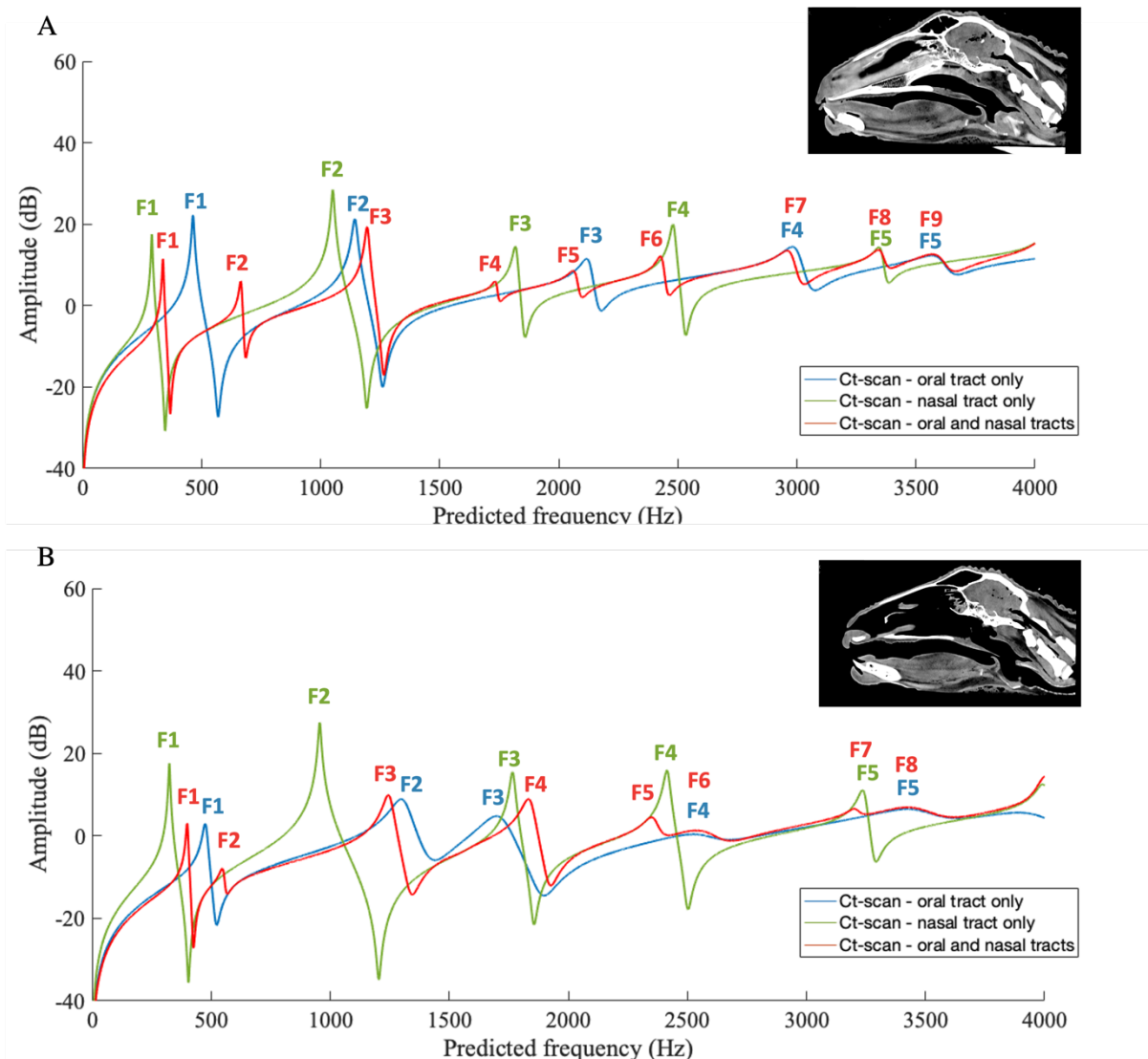


Figure 11. Acoustic impedance calculated from CT-scan of the ewe 32142, where frequency peaks correspond to predicted formants. A) prediction from closed mouth CT-scan. B) prediction from open mouth CT-scan.

To assess the respective and joint contribution of the oral and nasal tracts to the formant frequency patterns observed in the closed and open mouth bleats, we compared the formant predicted by the three CT-scan models (oral only, nasal only, oral & nasal) established for both closed and open mouth (Fig.11). In both closed and open mouth, the model considering only the oral tract predicts formant frequencies higher than formant frequencies predicted by the model considering only the nasal tract (Fig.11).

In both closed and open mouth, when combining both oral and nasal tracts: F1 is shifted to an intermediate position between those predicted by single tracts, a new formant appears around 600hz and F3 remains at the same frequency as F2 predicted by the oral tract only.

When the mouth is open, all of the three models predict a formant around 1780 Hz corresponding to F4 predicted by both oral and nasal tracts and the F5 seems to be affiliated to the nasal tract. When the mouth is closed, F4 predicted by both oral and nasal tracts is closer to the F3 predicted by the nasal tract than formants predicted by the oral tract, thus F4 seems to be affiliated to the nasal tract and, in the same way, F5 seems to be affiliated to the oral tract.

In summary, open and closed mouth bleats appear to involve both oral and the nasal tracts. In open mouth bleat the oral tract play the role of an open vocal tract whereas it plays the role of a closed cavity in closed mouth bleats.

Discussion

We investigated the effect of variation in mouth opening on the acoustic structure of bleats produced by female sheep in the context of separation from their lamb. We sought simple articulatory correlates of arousal and assessed the effect of these vocal gestures on formant frequency patterns. In a separation test we found that ewes produced more open mouth bleats in high arousal contexts and that these bleats were also characterized by a higher degree of mouth opening. The analysis of formant frequencies in the natural bleats produced during this experiment showed that the arousal level was correlated with f_0 and F1. To understand the mechanisms underlying the modulation of F1, we applied a combination of post-mortem techniques to predict formant frequencies in different vocal geometries: airflow through the vocal folds, vocal tract modelling from CT-scan, and acoustic impedance sensor measurements. These three methods allowed us to conclude that both oral and nasal tracts are simultaneously involved in the production of open and closed mouth bleats. Below we discuss the methodological and theoretical implications of these observations.

The open and closed mouth bleats have an important function in mother-young communication in sheep. Terrazas et al., (2002) highlighted the role of closed mouth bleats at an early stage of lamb development and associated to close proximity between ewe and lamb. Closed mouth bleats are produced mostly in the 24h following birth, yet in our behavioral experiment, lambs were days-old, which explains why we didn't observe the production of closed mouth bleats in

each tested ewe (N = 17 ewes under a total of 22). When the production of closed mouth bleats was observed it was only in context of close proximity with the lamb (i.e., baseline and reunion conditions) which confirms the observation made by Terrazas et al., (2002). Moreover, the production of open mouth bleats was positively correlated with the percentage of time in movement and was significantly higher in the separation condition than in the other conditions. The arousal level did have a positive correlation with the degree of mouth opening, suggesting that the arousal level of the ewe can drive the ewe mouth articulation during mother – young communication. Indeed, there is a direct link between the emotion arousal level and the somatic and autonomic nervous systems, which are responsible of the tension and movements of voice production muscles in humans (Porges & Lewis, 2010). Hence, the modification of the emotion arousal level induce the modification of the vocal tract resulting in particular changes of voice parameters (Porges & Lewis, 2010). The tension on the laryngeal muscles increases the tension on the vocal fold and which can explain the higher the fundamental frequency while the arousal level increases. We also found that F1 is positively corelated with the arousal level of the ewes. Formants are independent from the source, thus this correlation cannot be explain by the increase of the vocal fold tension (Taylor & Reby, 2010). However, formants depends on the vocal tract anatomy and the modification of the degree of mouth opening can explain the modulation of F1. In human, the constriction the anterior oral cavity lowers F1 which is highly affected by mouth articulation (e.g. mandible, lips, tongue) (Behrman, 2021; Fant, 1960; Titze et al., 2011). Similarly, Gamba et al., (2011) found that the fundamental frequency and F1 were higher in indris (*Indri indri*) call when the mouth was completely open than in indris call with the mouth barely open or in intermediate position.

We used a range of post-mortem production or modelling methods to investigate how the modification of the vocal tract configurations affects formant frequencies patterns in ewes' bleats. We found that both oral and nasal tracts were involved in the production of formants in open mouth bleats. This result is consistent with observations that the production of open mouth groans by fallow deer males also involves both oral and nasal tracts (Reby et al., 2018). We hypothesized that during closed mouth bleats only the nasal tract is implicated in the production of formants and thus that the vocal tract is similar to a simple tube. Interestingly, our investigations suggest that in closed mouth bleats the oral cavity remains hollow and is involved in the vocal tract (as an attached cavity). To our knowledge the adjunction of a hollow cavity in mammalian voice production has only previously been described in species that possess supralaryngeal air sacs (de Boer, 2008; Riede et al., 2008). Supralaryngeal air sacs are large

cavities connected to the vocal tract as it was shown in reindeer (Frey, Gebler, et al., 2007) and primates (Hayama, 1970). For some species this supralaryngeal air sacs has a role of resonator adding formants in lower frequencies (de Boer, 2008). This lower formant frequencies are known to be implicated in the exaggeration of male size in several species (Koala: [Charlton et al., 2011](#); Columbus monkey: [Harris et al., 2006](#); Howling monkey: [Schön, 1971](#)) and play a role in sexual selection by attracting females or to avoid agonistic interaction with male competitor (Owings & Morton, 1998). In sheep sexual selection was reduced at the minimum due to the domestication (Polak & Frynta, 2009). Another hypothesis is that the use of oral cavity allows to preserve the information about the mother identity between closed and open mouth calls. Indeed, ewes produce closed mouth calls mainly in the first few hours postpartum which are progressively replaced by open mouth calls (Dwyer et al., 1998; Sèbe et al., 2007), thus the preservation of formants frequencies may facilitate the learning of the acoustic signature by the lamb once it has learned it in the closed mouth bleats. This finding can be compared with acoustic gradation observed across bonobo's vocal repertoire which the reliable of the individual signature decreases with the increase of the arousal level (Keenan et al., 2020). We showed that, in ewes, mouth articulation gestures allow the acoustic encoding of more complex information than the individual signature (Sèbe et al., 2007, 2010), by adding dynamic information such as arousal, which can be modulated based on the individual's current internal state (Briefer, 2020; Pisanski et al., 2016; Raine et al., 2017; Taylor et al., 2016).

Our study thus reveals that the production of bleats in ewes involves relatively complex vocal tract configurations and that its acoustic structure cannot be produced by a static vocal tract consisting of a simple tube closed at one end and opened at the other. As such, we recommend that future studies of vocal tract production in terrestrial mammals use post-mortem methods presented in this article in order to study the involvement of supralaryngeal cavities in vocal production and their dynamic. Combined with preliminary observations of vocal gestures in live animals, these methods allow to systematically study vocal production in reasonably realistic anatomical context.

To conclude, our postmortem approach has allowed us to demonstrate how basic articulatory gestures involving relatively complex geometries affect formant patterns in the vocal signals of domestic sheep. We suggest that this method could also be used to investigate a wider range of articulation gestures, e.g. involving lips, tongue or laryngeal movements. We also suggest that acknowledging and investigating basic articulation abilities in a wider range of non-primate

species may help better understanding the evolutionary origins of more complex articulatory maneuvers in our own lineage.

Supplementary

Table S1. The five first formants frequencies (F1, F2, F3, F4, F5) extracted from natural bleats, predicted from CT-scan, post-mortem recordings, and the impedance sensor.

Method	Condition	Mouth configuration	F1 (mean ± s.d.)	F2 (mean ± s.d.)	F3 (mean ± s.d.)	F4 (mean ± s.d.)	F5 (mean ± s.d.)
Natural bleats	“High” bleats (N=11 ewes, n= 96 bleats)	Open	429.6 ± 71 Hz	675.9 ± 50.5 Hz	1076.8 ± 112.6 Hz	1442.6 ± 89.6 Hz	1816.8 ± 152.1 Hz
	“Low” bleats (N=8 ewes, n= 29 bleats)	Closed	351.2 ± 117.7 Hz	699.4 ± 139.2 Hz	1080.8 ± 175.7 Hz	1500.1 ± 192.4 Hz	1915.1 ± 225.8 Hz
CT-scan	Oral tract only (N=2 ewes)	Open	469.3 ± 1.8 Hz	1056.9 ± 275 Hz	1509.9 ± 246.2 Hz	2394.6 ± 190.9 Hz	3205 ± 262.8 Hz
		Close	476.4 ± 17 Hz	1096.4 ± 67.3 Hz	2160.2 ± 61.1 Hz	2990 ± 12.7 Hz	3709.2 ± 202.8 Hz
	Nasal tract only (N=2 ewes)	Open	325.45 ± 3 Hz	879.5 ± 74.2 Hz	1757.3 ± 10 Hz	2360.7 ± 47.1 Hz	3224.3 ± 0.1 Hz
		Closed	312.7 ± 29.6 Hz	917.9 ± 188.8 Hz	1832.5 ± 20.8 Hz	2446 ± 47.5 Hz	3327.9 ± 23.1 Hz
	Both oral and nasal tracts (N=2 ewes)	Open	387.5 ± 9.8 Hz	533.8 ± 26.9 Hz	1091.8 ± 172 Hz	1792.8 ± 28 Hz	2350.1 ± 10 Hz
		Closed	362.1 ± 33.5 Hz	646.5 ± 26.2 Hz	1124.7 ± 100 Hz	1773 ± 57.1 Hz	2100.3 ± 56.4 Hz
Post-mortem recordings	Both oral and nasal tracts (N=2 ewes)	Open	392.3 ± 6.4 Hz	651.5 ± 53.7 Hz	1208.5 ± 53.4 Hz	1705.4 ± 8.3 Hz	2096.5 ± 228.7 Hz
		Closed	348.6 ± 41.9 Hz	651.9 ± 54.7 Hz	1052.3 ± 61 Hz	1515.5 ± 45.3 Hz	1991 ± 96 Hz
Impedance sensor	Both oral and nasal tracts (N=2 ewes)	Open	194.2 ± 16.1 Hz	569.3 ± 28.1 Hz	880.3 ± 138.2 Hz	1526.4 ± 242.7 Hz	2198.2 ± 570.5 Hz
		Close	163.3 ± 6.6 Hz	253.2 ± 20.1 Hz	731.7 ± 140.4 Hz	1303.4 ± 21.8 Hz	1988.1 ± 162.2 Hz

Acknowledgements

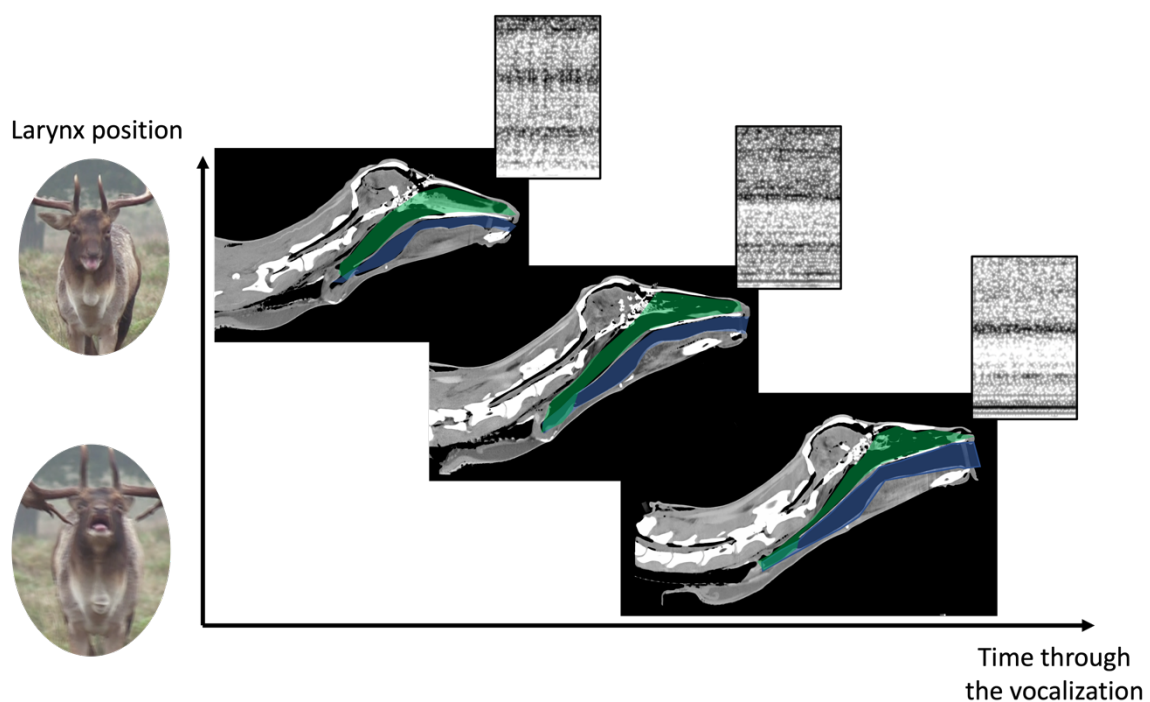
We thank François Lecompte for performing the CT-scans at INRAE Nouzilly, Chantal Porte and Maxime Meunier for their help during the separation experiment and the UEPAO staff for handling the animals. We thank Nicolas Boyer for his help building the CT-scan and post-mortem recordings setups.

This work was carried out thanks to the financial support of the IDEXLYON Project of Lyon University within the framework of the *Programme Investissements d'Avenir* (ANR-16-IDEX-0005).

Chapter III: Laryngeal descent in fallow deer



I showed in chapter II that ewes, a species with a high resting larynx, are able of basic articulatory gestures (i.e., mouth opening). In the introduction, I discussed how laryngeal anatomy has diverged between species in response to variable selection pressures. As a result of strong sexual selection, several species have developed a larynx in low resting position, a position that, in humans, has previously been associated with speech human ability. Fallow deer is one such species where males present the particularity of lowering their larynx further down the sternum when they produce their groans. In this chapter we investigate this laryngeal dynamic using post-mortem approach to highlight the anatomical changes involved in the formant modulation throughout vocalization.



My results show that, as fallow deer males extend their vocal tract during vocalization, both the oral and nasal vocal tracts are involved and their combined extension contributes to lowering formant frequencies.

Article 3: Modelling the laryngeal descent using post-mortem approaches

Lucie Barluet de Beauchesne¹, Léo Papet¹, Joël Gilbert^{2,+}, Yann Locatelli³,
Matthieu Keller⁴, Nicolas Mathevon^{1,5} and David Reby^{1,5}

¹ ENES Bioacoustics Research Laboratory, CRNL, CNRS, Inserm, University of Saint-Etienne, France

² Laboratoire d'Acoustique de l'Université du Mans, CNRS, 72085 le Mans, France

³ Réserve de la Haute Touche, Muséum National d'Histoire Naturelle, Obterre, France

⁴ Laboratoire de Physiologie de la Reproduction & des Comportements, INRA/CNRS UMR 7247, Université de Tours, FRANCE

⁵ Institut universitaire de France

[†] Deceased

Abstract

While articulation is generally considered as a hallmark of human speech production, there is growing evidence that animals can also change the shape and length of their supralaryngeal vocal tract to modulate the resonance frequencies in their calls (i.e., formants). This ability may ultimately increase the potential of encoding information in their vocalizations. During the breeding season, fallow deer produce groans to attract females and repulse males. While groaning was correlated to a descent movement of the larynx down to the sternum, the casual link between the shape of the vocal tract and the formant patterns remains unknown. In order to understand this mechanism, we used an innovative approach by combining two post-mortem methods: i) scanning fallow deer bucks head and neck specimens positioned in various vocal production postures and measuring the cross-sectional area of the vocal tracts to predict formant frequencies, and ii) using compressed air to generate sound in the glottis and measuring the resulting formant frequencies, across the same range of postures. Our results demonstrated that the descent of the larynx is responsible for the lowering of formant frequencies and that the formant frequency pattern is influenced by both the oral and nasal cavities of the vocal tract. Moreover, our study showed that post-mortem methods are reliable to investigate formant dynamic in mammal vocalization.

Introduction

Articulation has been widely studied in human from mechanistic, cognitive and functional aspects using a wide variety of methods (Dang & Honda, 2002; de Boer & Fitch, 2010; Fitch & Giedd, 1999; Schroeder, 1967; Takemoto et al., 2006). Articulators organs such as the lips, the tongue and the mandible induce constrictions/expansion of the vocal tract in several points, which influence formant frequencies composing the vowels (Fant, 1980; Titze, 1994). For instance, the tongue positioned high inside the mouth creates a constriction of the oral tract which increases the first formant and contributes to the production of the vowels /i/ and /u/. In contrast, a low position of the tongue reduces the constriction of the oral tract, resulting in lowering the first formant while producing the vowels /a/ or /æ/ (Schwartz, 1968).

Understanding the evolutionary origins of human articulatory abilities requires the application of the comparative method. While long underestimated, it is now recognized that the production of formants by the vocal tract plays an important role in animal communication (Bradbury & Vehrencamp, 1998). For instance, formant dispersion, defined as the averaged difference between successive formant frequencies, was shown to be an indicator of body size (Reby & McComb, 2003a; Riede & Fitch, 1999). Indeed, there is a negative correlation between the formant dispersion and the body size (Reby & McComb, 2003a). This acoustic cue is crucial for mate choice in estrus red deer hinds or in female koalas with both preferring calls containing lower formants, indicating larger males (Charlton et al., 2007, 2012). In the context of male-male competition, size-related formant dispersion is also used by males to assess the fighting ability of competitors, allowing them to avoid unnecessary fights with perceived stronger and more dominant males (sea lion: Charrier et al., 2011; giant panda: Charlton et al., 2010; red deer: Reby et al., 2005; dogs: Taylor et al., 2010). Formants are, therefore, under strong selective forces and some species may have evolved vocal apparatus that supports the exaggeration of the expression of body size in their calls (Charlton & Reby, 2016; Ryan, 1986). For instance, red deer, impalas and humans males have all developed descended larynges allowing them to exaggerate their perceived body size conveyed by their calls (Fitch & Reby, 2001; Frey et al., 2011, 2020).

Fallow deer belongs to the family *Cervidae* and subfamily *Cervinae* and both sexes have a descended larynx, albeit much more pronounced in males than in females (Fitch & Reby, 2001; Frey et al., 2021). During the rutting period, males further extend their vocal tract during groans by lowering their larynx toward sternum (Fitch & Reby, 2001). Reby et al., (2018) CT-scanned

the head and neck of two fallow deer bucks with larynx fully descended toward the sternum and predicted the resonance frequencies using three acoustic models considering either (1) only the oral tract, (2) only the nasal tract or (3) both oral and nasal tracts. The authors found that the best predicting model for groan formants was the one obtained using both oral and nasal tracts. This result indicated that, in addition to the laryngeal descent, formant pattern in groaning involves both oral and nasal tracts to participate to the size exaggeration conveyed by the groans. Interestingly, the groan starts when larynx is close to the resting position and then larynx extends over the groan, up to the end where the vocal tract is at its maximal extension (Fitch & Reby, 2001; McElligott et al., 2006). During the laryngeal descent formants are lowered and the formant spacing is negatively correlated with the position of the larynx during its descent (McElligott et al., 2006). The geometries responsible for the formants observed during the lowering larynx before reaching the maximal extension of the vocal tract remain unknown. Here, using post-mortem approaches, we aimed at investigating the involvement of oral and nasal tracts in the production of groans. We first scanned the head and neck of two fallow deer for which the larynx was pulled in three different positions (i.e., rest, intermediate and maximal). Then, across the same range of postures (i.e., rest, intermediate and maximal), we compared formants predicted by similar acoustic models as those used in Reby et al., (2018) with formants observed in sound recordings obtained using an innovative method, hereafter called *post-mortem recordings*. The post-mortem recordings consist of blowing air into the trachea of fallow deer head and applying pressure on both thyroid cartilage lamina to generate a sound source signal at the glottis. As this method enables to use same specimens for both acoustic model and post-mortem recordings, we predicted that we would have higher matches between predicted and observed formants than in Reby et al., (2018). Furthermore, we hypothesized that both oral and nasal tracts are involved in call production throughout the entirety of the laryngeal descent.

Material and methods

Dead specimens

Specimens were two fallow deer bucks H426 and H438 respectively aged of 4 and 2 years old, from the zoo parc La Haute Touche (France). The two animals died from injuries caused while fighting with conspecific rival males. The head and neck specimens were obtained by a section at the sternum and then frozen at -20°C.

CT-scan

CT-scans were performed at the Pixanim platform of the PRC unit (https://www6.val-de-loire.inrae.fr/pixanim_eng/; INRAE Val-de-Loire). Prior scanning, each specimen was slowly thawed and flushed to remove debris and fluids from the oral and nasal cavities. They were suspended for about 30 min to dry before scanning. Specimens were maintained in a structure to avoid the trachea collapsing, and the mouth was maintained opened with a polyethylene wedge (Fig.1).



Figure 1. 3D CT-scan image representing the structure used for scanning the specimens.

To simulate the dynamic of the laryngeal descent over the fallow deer buck groaning, CT-scans were performed for three different positions of the larynx: resting position (*rest position*), intermediate extended position (*int position*) and maximum extended position (*max position*). The *rest position* was obtained without pulling on the trachea. The *max position* was achieved by maximally pulling the trachea toward the sternum and fastening it to the back of the structure hosting the head. The *int position* was achieved by releasing the pressure on the trachea until the larynx was in intermediate position between the max and the rest position.

Specimens were refrozen at -20°C after scanning and transported to the ENES laboratory to perform the post-mortem sound recordings.

Prediction of vocal tract resonances

Using the software Osirix (version 12.0) cross-sectional areas were semi-automatically extracted every 1 cm in the oral and nasal tracts (Fig.2) from CT-scan following the methods described in Reby et al. (2018).

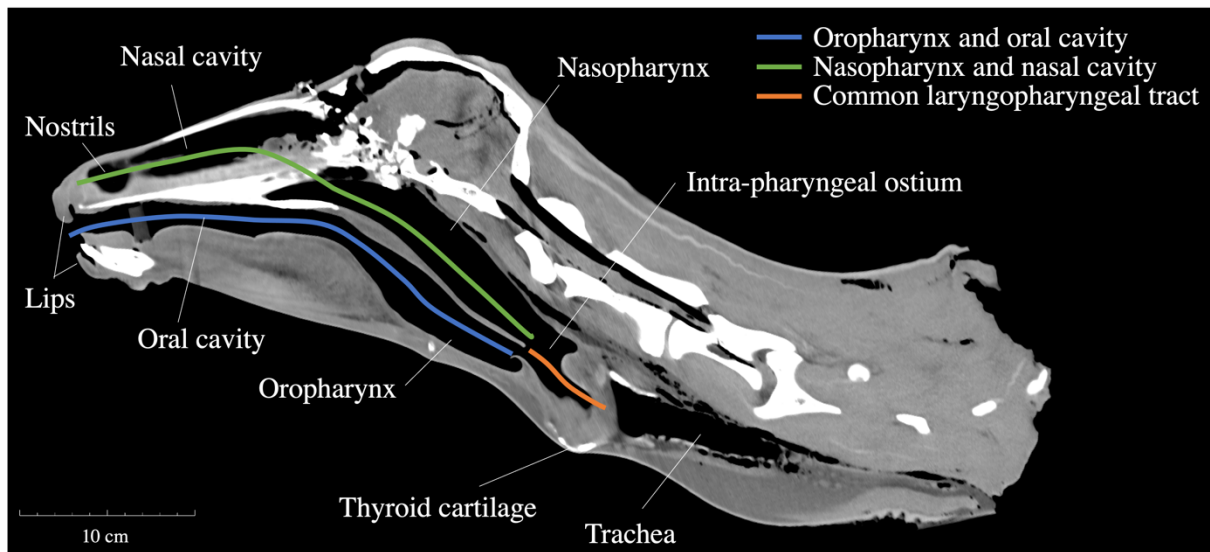


Figure 2. Mid sagittal slice of a CT-scan of male H438 with the larynx in rest position, illustrating the manual drawing of the dorsoventral transverse midline through the oral vocal tract and the nasal vocal tract. The oral vocal tract corresponds to the oropharynx and oral cavity (blue) and the common laryngopharyngeal tract (orange). The nasal vocal tract corresponds to the nasopharynx and nasal cavity (green) and the common laryngopharyngeal tract (orange).

The transfer matrix method (Chaigne & Kergomard, 2016) was used to predict the frequency of vocal tract resonances (Reby et al., 2018). For each specimen, three types of acoustic models were made based on three different configurations of the vocal tract (see Fig. 2):

- Only the oral tract (laryngopharyngeal tract, oropharynx, and oral cavity; Fig.2) involved in the production of the resonances, model called *oral tract only*.
- Only the nasal tract (laryngopharyngeal tract, nasopharynx, and nasal cavities; Fig.2) involved in the production of the resonances, model called *nasal tract only*.

- Both the oral and nasal tracts (common laryngopharyngeal tract, oropharynx, oral cavity, nasopharynx, and nasal cavities; Fig.2) involve in the production of the resonances, model called *combined oral and nasal tracts*.

Post-mortem recordings

Each specimen was slowly thawed and drained before the experiment.

The head of each specimen was maintained in a position suitable for the protocol using a metallic structure (Fig.3). One experimenter was in charge to establish the vocal fold oscillation by pressure on the thyroid cartilage of the larynx. Using a pressure regulator, the air pressure was maintained between 4 and 6 kPa to induce the vibration of the vocal folds. Three experiments were performed one for each laryngeal position (*rest, int, max*). The intermediate and max positions of the larynx were performed using cable ties attached to the trachea and to the metallic structure. The degree of mouth opening was extracted from the video recorded with a HC-VX980 Panasonic camera placed at 1 m from the side of the specimen.

Sound was recorded at 1 m from the mouth using a Behringer ECM8000 microphone. The first five formants were measured using PRAAT (version 6.1.52). We used the *view spectral slice* function on which we applied a cepstral smoothing at 150 Hz and extracted the frequency peaks using both the function *To formant (peaks)* and a visual checking on the spectrogram (view range: 0-2000Hz; window length: 0.05s; dynamic range: 50dB).

The estimated vocal tract length (*eVTL*) was calculated using the following equation : $eVTL=c/2 \times DF$, where c is the speed of sound of 35000 cm.s^{-1} and DF is the average of the formant spacing in vocalization (Reby & McComb, 2003a).

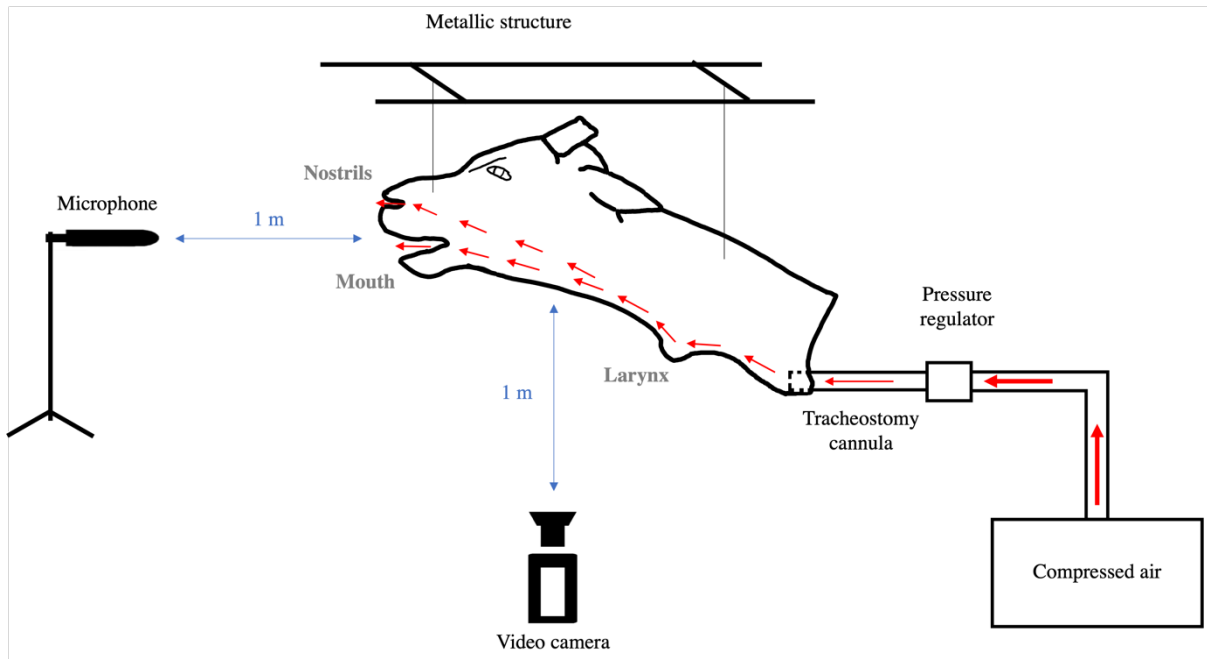


Figure 3. Schematic illustration of the experimental design of post-mortem recordings. The head is suspended using the metallic structure allowing the access to the larynx during the experiment. A compressor was used to generate an air flow (red arrow) into the trachea. The air pressure was manually controlled by a pressure regulator.

Results

Formants predicted from CT-scan

Based on the mandibula length, male H438 (mandibula length: 18.5 cm) was found to be slightly smaller than male H426 (mandibula length: 19 cm).

In the resting position, the larynx was located at the level of the cervical 2 for both males. The oral tract length measured from the glottis to the lips was of 34.9 cm for male H426 and of 32.8 cm for male H438. The nasal tract length from glottis to nostrils was 35.5 cm for male H426 and 34.7 cm for male H438.

To obtain the intermediate position of the larynx, the vocal tract length was extended by 2 cm for the male H426 and by 5.8 cm for the male H438 from the rest position. For the max position, the vocal tract was extended by 5 cm for the male H426 (panel 1 of Fig.4) and by 8.6 cm for the male H438 from the rest position (panel 1 of Fig.5).

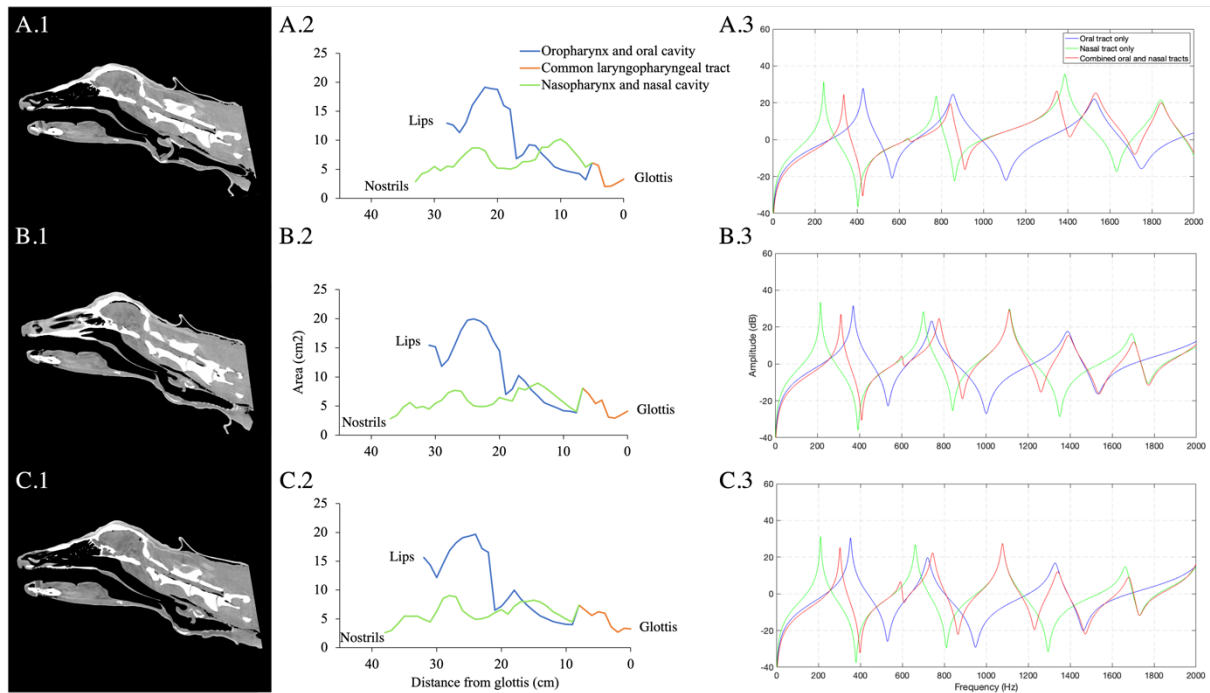


Figure 4. CT-scan (1), cross-sectional area (2) and predicted resonances (3) of the male H426 in three laryngeal positions: rest position (A), intermediate position (B) and max position (C). The cross-sectional area for the oral cavity were corrected for the angle of mouth. Predicted resonances were calculated for three vocal tract modelling: oral tract only (blue), nasal tract only (green) and combined oral and nasal tracts (red).

During the CT-scan, the elongation of the vocal tract from the rest to the max position extended the nasal tract (from nostrils to the ostium) by 3.1 cm for the male H426 and by 3.4 cm for the male H438; the oral tract by 2.2 cm for the male H426 and by 4.8 cm for the male H438 (panel 2 of Fig.5).

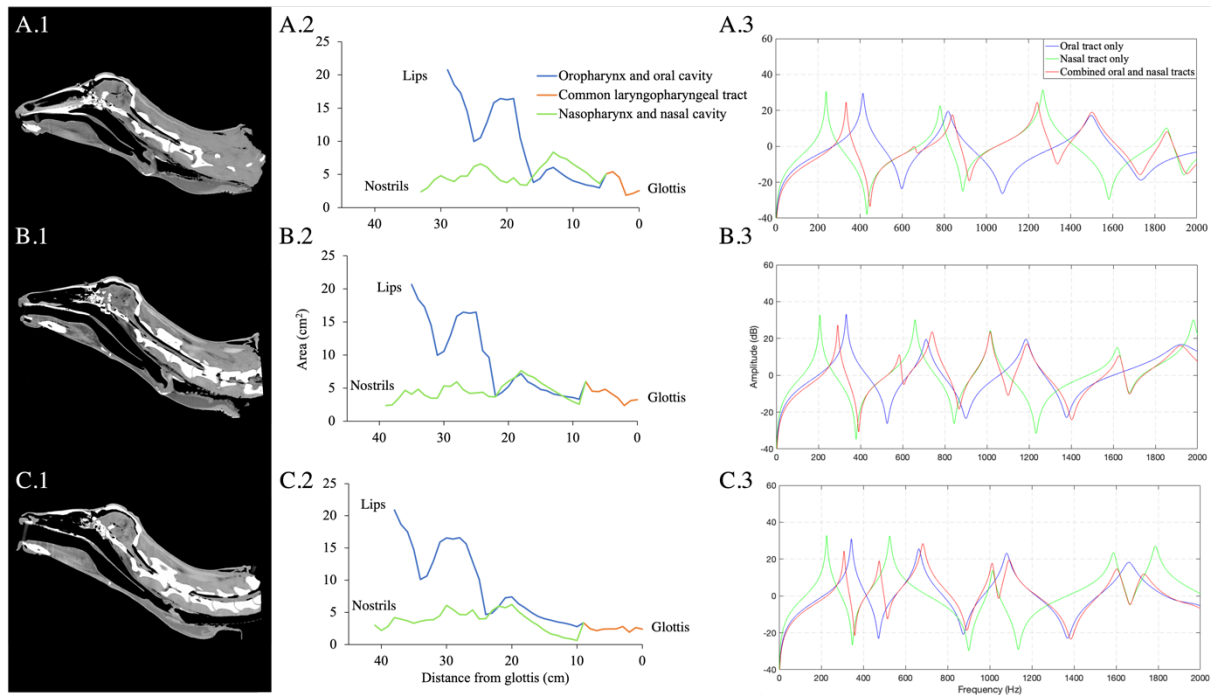


Figure 5. CT-scan (1), cross-sectional area (2) and predicted resonances (3) of the male H438 in three larynx configurations: rest position (A), intermediate position (B) and max position (C). The cross-sectional area for the oral cavity was corrected for the angle of mouth. Predicted resonances were calculated for three vocal tract modelling: oral tract only (blue), nasal tract only (green) and combined oral and nasal tracts (red).

Comparison of predicted frequencies between the three laryngeal positions (i.e., *rest*, *int*, *max*) showed, for both males, that the predicted frequencies tended to be lower with an elongation of the vocal tract (panel 3 of Fig.4; panel 3 of Fig.5). This tendency was true for all of the three acoustic models (i.e., *oral tract only*, *nasal tract only*, *combine oral and nasal tracts*; Fig.4; Fig.5). The first formant (F1) predicted by the acoustic model combining the oral and nasal tracts had an intermediate frequency between the F1 frequency values predicted by the two other models. F2 predicted by the acoustic model combining the oral and nasal tracts had a value between F1 of the *oral tract only* model and F2 of the *nasal tract only* model. F3 of the combined oral and nasal tracts model was close to F2 predicted by *oral tract only*. F4 predicted by *combined oral and nasal tracts* was close to F3 predicted by *nasal tract only*. F5 predicted by *combined oral and nasal tracts* was close to F3 predicted by *oral tract only*.

Formants observed in post-mortem recordings

During the post-mortem recordings, the degree of mouth opening of the male H426 was maintained at $20.5 \pm 0.2^\circ$ (mean \pm s.d.) and at $15.3 \pm 0.3^\circ$ for the male H438 through the three laryngeal positions (rest, int, max).

For both males (H426 and H438), between the rest position and the max position, there were stronger changes in the third (-103 Hz for H426 and -110 Hz for H438), the fourth (-112 Hz for H426 and -123 for H438) and fifth formant (-541 Hz for H426 and -528 for H438) than in the first (-29 Hz for H426 and -4 for H438) and the second formant (-59 Hz for H426 and -8 for H438) (**Erreur ! Source du renvoi introuvable.**).

Formant spacing observed in post-mortem recordings at the rest position, provided an estimate of the vocal tract length of 45 cm for the male H426 and 46 cm for the male H438 (**Erreur ! Source du renvoi introuvable.**). The elongation of the vocal tract until the intermediate position increased this estimation by 10 cm for the male H426 and by 11 cm for the male H438. At the max position the estimated vocal tract length was elongated by 23 cm for the male H426 and by 24 cm for the male H438.

Table 2. The five first formants (F1, F2, F3, F4, F5) observed in the post-mortem recordings of males H426 and H438 in the three laryngeal positions (rest, int, max). DF corresponds to the mean spacing between formants. Estimate VTL corresponds to the vocal tract length estimated using the mean spacing.

Laryngeal position	Male	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	F5 (Hz)	DF (Hz)	Estimated VTL (cm)
Rest	H426	329	575	760	1225	1870	385	45
	H438	334	558	777	1218	1863	382	46
Int	H426	329	461	658	1141	1600	318	55
	H438	340	598	758	1212	1565	306	57
Max	H426	300	516	657	1113	1329	257	68
	H438	330	550	667	1095	1335	251	70

Comparison between formants extracted from postmortem recordings and formants extracted from natural groans (data from (McElligott et al., 2006) indicated a good correlation for both individuals, in the scaling (slope > 0.8), and in frequency match ($R^2 > 0.9$), for the rest and the maximal extended positions (Fig.6).

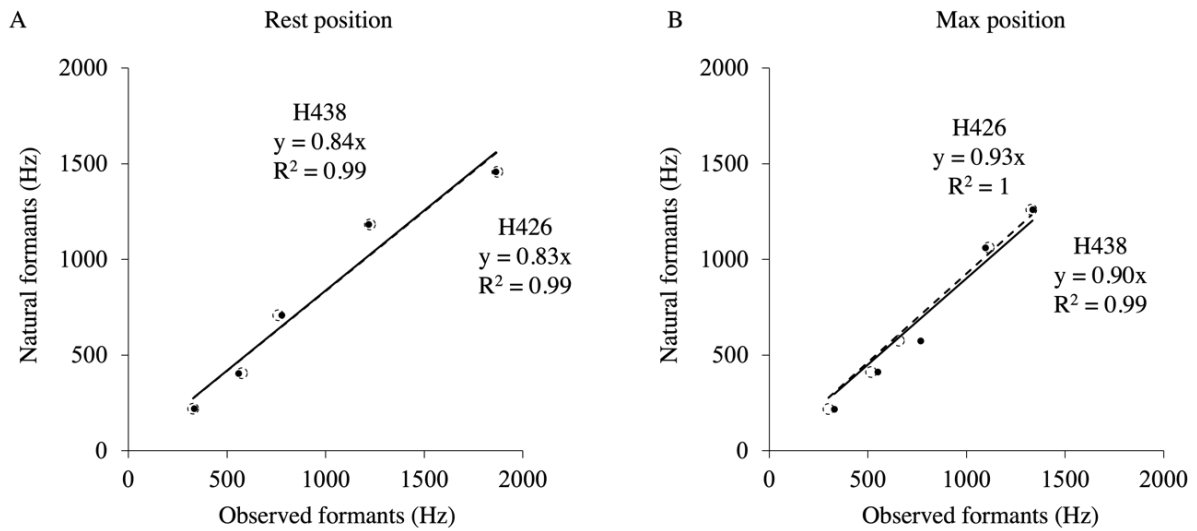


Figure 6. Comparison of formants observed in postmortem recordings (x-axis) and natural formants observed in fallow deer groan (n=37 groans, N= 9 males). Observed formants in rest position are compared to natural formants extracted at minimum VTL during the groan (A). Observed formants in max position are compared to natural formants extracted at maximum VTL during the groan (B). The intercept was forced to pass by zero. The regression slopes inform the fit of the formant spacing to the predicted resonances, more the slope tends to 1.0 more the formant spacing match with the spacing of the predicted resonances. The R^2 values provide the fit of formants frequencies to the frequencies of the predicted resonances, more R^2 tends to 1.0 more the frequencies match. (Data of natural formants from McElligott et al., 2006).

Correlation between observed and predicted formants

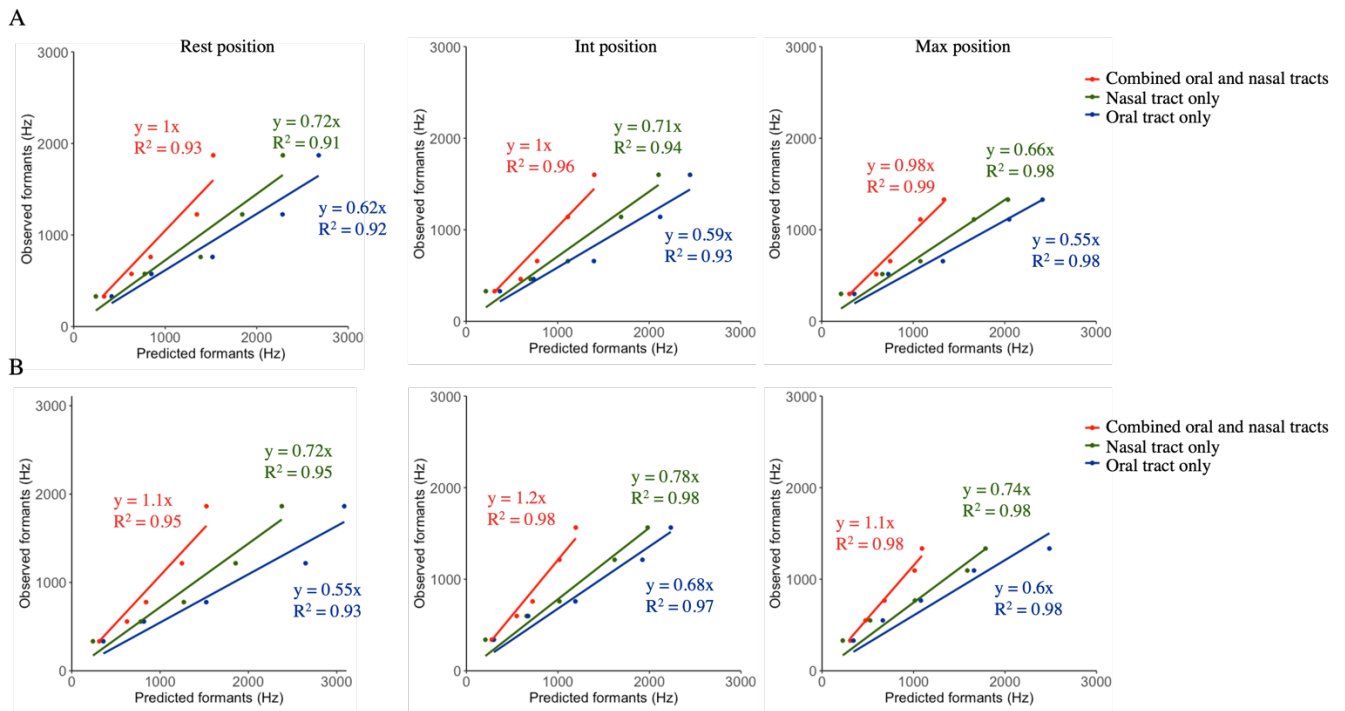


Figure 7. Correlation between the first five formants observed in post-mortem recordings and the five first formants from the three acoustic models (oral tract only, nasal tract only, combined oral and nasal tracts) for male H426 (A) and male H438 (B). Observed and predicted formants were extracted in three laryngeal positions: rest, int and max. The intercept was forced to pass by zero. The regression slopes inform the fit of the formant spacing to the predicted resonances, more the slope tends to 1.0 more the formant spacing match with the spacing of the predicted resonances. The R^2 values provide the fit of formants frequencies to the frequencies of the predicted resonances, more R^2 tends to 1.0 more the frequencies match.

For the three laryngeal positions, we obtained good correlations between formants extracted from post-mortem recordings and formants predicted from the three acoustic models (i.e., *oral tract only*, *nasal tract only*, *combined oral and nasal tracts*) ($R^2 > 0.9$; Fig.7).

In all of the three laryngeal positions, the acoustic model simulating the vocal tract as combination of two tubes (i.e., combined oral and nasal tracts) predicted formants with a better scale (the slope of the regression equal to 1) than the acoustic model simulating the vocal tract as a single tube (i.e., oral tract only, nasal tract only) (Fig.7). Both acoustic models simulating the vocal tract as a single tube underestimates the formant spacing observed in post-mortem recording of male H426 and H438 in all of the three laryngeal positions (Fig.7).

When the larynx was at a resting position, the formant spacing predicted by the combined oral and nasal tracts model matched perfectly to formant spacing observed in post-mortem recording

of the male H426 whereas this model underestimated by 10% the formant spacing observed in post-mortem recording of the male H438. At an intermediate position of the larynx, the formant spacing predicted by the combined oral and nasal tracts model matched perfectly to formant spacing observed in post-mortem recording of the male H426 whereas it was underestimated by 20% for the male H438. When the larynx was at its max extended position, the combined oral and nasal tracts overestimated by 2% and underestimated by 10% the formant spacing observed in post-mortem recording of the male H426 and male H438 respectively (Fig.7).

Discussion

Using post-mortem approaches, we investigated the effect of vocal tract extension and laryngeal descent on formant frequency patterns and scaling in fallow deer groans. The vocal tract of two males were artificially extended to maintain the larynx in three different positions simulating natural laryngeal descent during groaning: resting, intermediate and maximal positions. For each of these three positions, we measured formants observed in post-mortem recordings and compared them to the resonance frequencies predicted by the corresponding vocal tract geometries. Our results suggest that male fallow deer use both oral and nasal tracts during the laryngeal descent when vocalizing. We discuss our results in regard to previous findings and the pertinence of our methodology.

We used same acoustic models in our study and observed similar pattern of predicted formants as found by Reby et al., (2018), except for the second formant predicted by the combined oral and nasal tracts model which appeared to be less associated to the nasal tract in our study. This difference can be explained by the fact that we used two smaller, young adult, fallow deer males which could affect the vocal tract anatomy and thus modify the predicted formant pattern. Another explanation could be that, contrary to Reby et al., (2018), we were able to stretch the neck of specimens to reproduce the natural posture of vocalizing fallow deer. This position could modify the anatomy of oral and nasal tracts by decreasing the difference of length between these two tracts.

We compared acoustic model prediction to formant observed in post-mortem recordings consisting in blowing air into the trachea while the vocal folds vibrate under manual pressure on the larynx. This method was first used by Ferrein (1741) to understand the mechanisms underlying the vocal production. Using heads and excised larynges of several species including human, he demonstrated the role of the vocal folds in the sound production (Ferrein, 1741).

Later, several studies used excised larynges to understand the mechanisms influencing the modulation of the fundamental frequency (Alipour et al., 2007; Alipour & Jaiswal, 2008; Garcia & Herbst, 2018; Herbst et al., 2013; Maytag et al., 2013; Titze et al., 2010). To our knowledge, this is the first study using postmortem method on the whole head to understand the link between formant and vocal tract configuration rather than just larynx. This method allowed us to use same specimens that we used to predict formants with the previously developed acoustic models. As a result, we could control for the animal size which explain the better match between predicted and observed formants ($R^2 > 0.98$) than previously found by Reby et al., (2018) ($R^2 > 0.94$) for the maximal extended position. When the fallow deer vocal tract was maximally extended, Reby et al., (2018) have found that the acoustic model combining the oral and nasal tracts predicted better the formant observed in natural fallow deer groans than configurations including only one tract. Our results confirm this finding and add a dynamic aspect to this study as we showed that this configuration combining both oral and nasal tracts was the best for predicting formants in all three laryngeal positions studied. In addition to the low laryngeal position, the involvement of both oral and nasal tracts may enable the fallow deer to increase the exaggeration of the perceived body size. Exaggeration of body size through formant frequencies has also been found in other mammals and bird species that possess extension of their vocal tract or other resonators organs (Charlton & Reby, 2016; Fitch, 1999). It is well known that the exaggeration of perceived body size represent an advantage for fallow deer to protect their harem against rival males (McElligott & Hayden, 1999; Pitcher et al., 2015).

The artificial modulation of the vocal tract allowed us to record formants similar to those observed in fallow deer natural groans, in frequencies and dynamic (Fitch & Reby, 2001; McElligott et al., 2006). As such, we observed the decrease of the formant spacing while the larynx is lowered, similar to formant pattern in natural groans (Fitch & Reby, 2001). To study the dynamic of formants, cineradiographic and 3D cine-MRI methods have been largely used in human (Badin et al., 1995; Jallon & Berthommier, 2006; Y.-C. Kim et al., 2009; Lindau, 1979; Takemoto et al., 2006) and on several other species such as dogs, goats, and pigs (Fitch, 2000). This method presents the advantage to take into account muscles movements during call production which can be a strong limitation in post-mortem studies. Moreover, post-mortem changes and frozen storage could have changed the elasticity of muscles and ligaments involved in the laryngeal descent and the head movements (Bendall, 1960; Fitzgerald, 1975; Takahashi & Saito, 1979). Thus, using cadaveric samples can lower the possibility of movements when performing the experiment and can modify the interior structures of the oral and nasal tract

inducing errors when measuring the cross-sectional areas. However, studying alive animals has also currently insurmountable limitations such as the management of alive wild animals which can implicate a long habituation phase and the constraint of the animal size.

We conclude that, even if it has some limitations, post-mortem investigations are reliable methods to study vocal tract articulation and allows to get over the difficulties encountered when studying alive animals. We showed strong evidence that fallow deer use both oral and nasal tracts all along with the laryngeal descent and we need further investigation for studying the function of the formant dynamic in their groans.

Acknowledgements

We thank François Lecompte in charge of the CT-scan at INRAE, Nicolas Boyer for his help to build the CT-scan and post-mortem recordings setups and the animal keepers from the zoopark la Haute Touche (France).

This work was carried out thanks to the financial support of the IDEXLYON Project of Lyon University within the framework of the *Programme Investissements d'Avenir* (ANR-16-IDEX-0005).

Annex to chapter III: Playback experiment to test the information encoded by the laryngeal descent of fallow deer

Introduction

In most mammal species it is known that vocalization play a crucial role to transmit information about the motivation of the caller (Morton, 1977; Owings & Morton, 1998). In rutting period, Fallow deer bucks produce vocalization while they lower their larynx toward sternum. This vocal tract modulation allows them to exaggerate the perceived body size and represent an advantage for fallow deer to protect their harem against rival males (McElligott & Hayden, 1999; Pitcher et al., 2015). While the maximal extended position was associated to specific context (Charlton & Reby, 2011), the function of the formant dynamic in their vocalizations has not been yet studied. Throughout playback experiment simulating three different speed of laryngeal descent, we hypothesis that rivals males would react more to vocalization simulating a higher descent speed than slow laryngeal descent.

Material and methods

Playback experiment

Playback experiments were performed, during the rutting period, in October 2019 at Petworth park, West Sussex (U.K.), in semi-captive fallow deer population (enclosure of 2.8 km²).

Stimuli were made based on natural groans of two different males previously recorded the season before or during the same season. Vocalizations of both males used to do the stimuli were recorded using a Sennheiser MKH416P48 microphone connected to a Canon XA11 HD camera to identify afterward the focal animal recorded. All fallow deer males used in this study were identified visually using the unique antler shape and coat marking of each male.

Using the software Praat (version 6.1.52), the pulses of the groans were extracted and reassembled to create three types of vocalizations: simulating a constant maximum vocal tract length (constant minimum formant spacing), simulating a slow laryngeal descent (progressive

descended slope of the formant frequencies), simulating a fast laryngeal descent (drastic descended slope to quickly arriving to the minimum formant spacing) (Figure 1). One playback is composed by three synthesized vocalizations randomly ordered and repeated for a total duration of 1 min. Between each playback, we let a minimum of 10 min without sound diffusion.

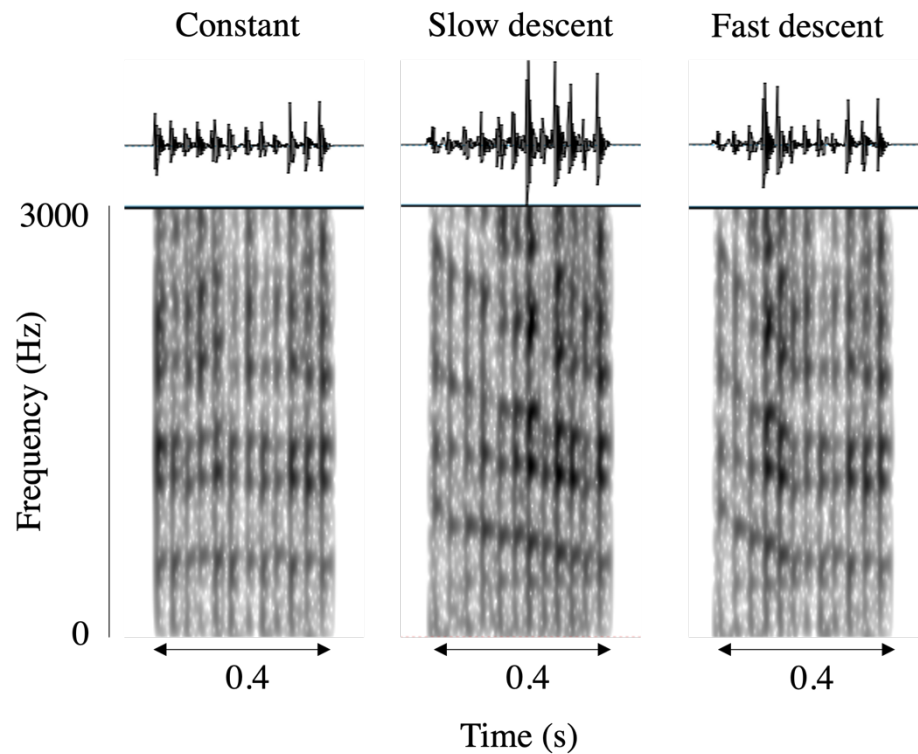


Figure 1. Example of synthesized vocalization used for each playback type: constant, slow descent and fast descent. (Window length: 0.02 s; dynamic range: 70 dB).

Playbacks were broadcasted using a speaker (Behringer EUROPORT MPA40BT) and a player (zoom H4n) at 60 m from the focal male. To reduce the variability in the behavioral response, we controlled for the context before starting the playback experiment: the focal male is active and at less than 3 m of its harem (presence of at least two females at less than 5 m from the focal male) before starting the experiment (Figure 2).



Figure 2. Context in which playback was broadcasted. The focal fallow deer buck is active and inside his harem constituted several females before playback experiment.

Measure of the behavioral response

The behavior of the focal male was filmed, using a Canon XA11 HD camera, for a minimum of 1 min before the playback, during the playback and a minimum of 1 min after the end of playback. The video recordings were analyzed using the software Boris (version 7.12.2).

We measured the following behavior from the start of playback:

- The latency before starting to be immobile
- The latency before turning the head toward the sound source
- The latency before emitting the first call
- The latency before the first contact with female

We measured the following behaviors before and during the playback:

- Number of calls
- Number of physical contacts with female
- Number of agonistic behaviors toward other males
- The duration of immobility
- The duration of head toward the sound source

To measure the behavioral change between before the playback and during the playback, we calculated for each behavior measured before and during the playback a score of reaction consisting to substrate the value before the playback to the value during the playback.

Statistical analysis

All behavioral data were included in a Principal Component Analysis (PCA) to reduce the number of response variable and Linear Mixed Models were performed on the coordinate of the PCA. The stimulus type (constant, slow descent, fast descent), the playback version (from male 1, from male 2) and the interaction between the stimulus type and the playback version, were added in the model as fixed effect variables and the identity of the focal male as a random effect variable.

Results

We tested 17 fallow deer bucks for a total of 40 playback experiments (16 with the slow descent stimulus, 17 with the fast descent stimulus and 7 with the constant stimulus). We were able to test four males with the three types of stimuli and nine with the slow descent and fast descent stimuli.

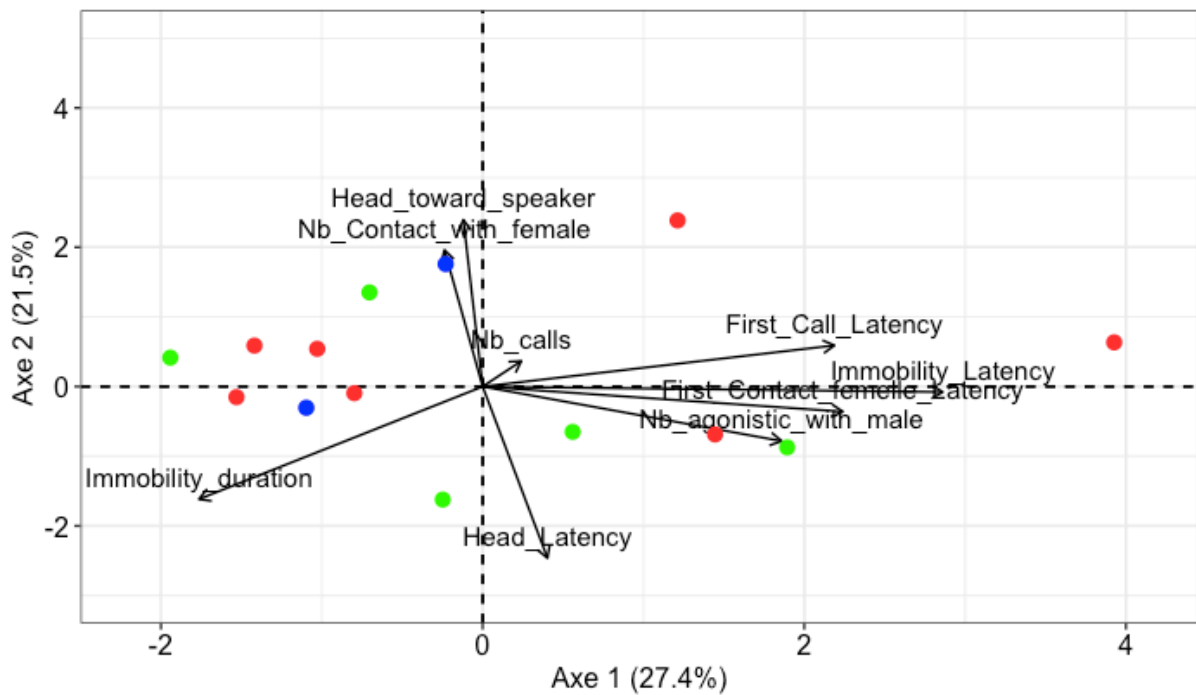


Figure 3. Result of the two first axes of the PCA including the behavioral data for the constant stimulus (blue), the slow stimulus (red) and fast stimulus (green).

In order to characterize the response of the focal fallow deer bucks in response to playbacks, we performed a Principal Component Analysis (PCA) on all the nine behavioral parameters measured. The first axis (axis 1) of the PCA explicates 27.4% of the total variability and the second axis (axis 2) 21.5% (Figure 3). The first axis is most represented by the latency before starting to be immobile (32.96%), the latency before the first contact with female (20.2%) and before the first emitting call (19.3%). The second axis is most represented by the latency before turning the head toward the speaker (31.1%), the duration of head turns in direction of the sound source (29.4%) and the number of contacts with female (19.7%).

To test if the speed of the decrease of formant spacing influence the behavior of the focal male tested, we performed Linear Mixte Model (LMM) on the two first axis of the PCA. The LMM on the first axis shows no significant differences between stimulus type ($F = 0.74$; $p = 0.5$), nor between playback version ($F = 1.33$; $p = 0.27$), nor between the interaction between stimulus type and playback version ($F = 0.4$; $p = 0.54$). The LMM on the second axis shows no significant differences of the stimulus type ($F = 3.1$; $p = 0.08$), nor of the playback version ($F = 1.4$; $p = 0.26$), nor of the interaction between the stimulus type and the playback version ($F = 0.03$; $p = 0.87$).

Conclusion

We did not find any effect of the formant dynamic on the fallow deer bucks behaviours. The hypothesis that the speed of laryngeal descent could code for motivation level in male interactions is therefore not supported by our playback experiments. Thus, our results suggest that the maximal extended position is sufficient for males to detect the aggressive motivation of competitors (Charlton & Reby, 2011). Further investigations are now needed to see if this dynamic is more directed toward females than males.

Discussion and perspectives

The aim of the work reported in this thesis was to improve our understanding of the origins of the vocal diversity in terrestrial mammals, and to identify examples of articulatory gestures in the natural vocal communication system of non-primate species.

In the first chapter, I describe the laryngeal development and test the effect of testosterone on this development in two closely related species: the sika deer and the red deer. These two species are interesting contrasting models as they present differences in their larynx position and in their level of sexual dimorphism of their vocal apparatus. Specifically, while sika deer possess a larynx characterized by a relatively high resting position like most terrestrial mammals, red deer possess a larynx in a relatively lower resting position (as also described in e.g. fallow deer or humans). In addition, red deer are characterized by a higher sexual dimorphism in their vocal apparatus than sika deer. The position of the larynx is lower in red deer males compared to females whereas there is no such difference between sexes in sika deer. While I expected a stronger effect of testosterone level on the laryngeal development of red deer, I found that post-natal testosterone had a marginal effect on the laryngeal development at both species, at 18-month-old individuals.

After studying the development of the vocal apparatus in species that possess either a larynx in high or in low resting position, I investigated the presence of two examples of articulatory gestures (i.e., the mouth opening and the laryngeal descent) in two ungulate species that present two different types of vocal apparatus (i.e., characterized by low or high resting position of the larynx). Specifically, I focused on domestic sheep that open their mouth when producing vocalizations with their high-resting, immobile larynx (chapter II), and on fallow deer that present a relatively low resting but mobile larynx (chapter III). Using a multi-disciplinary approach, combining behavioral measurements, acoustic modelling, and innovative postmortem experiments, I showed that vocal tract movements play an important role in the modulation of vocalizations in these two species. I also demonstrated that the modulation of the vocal tract geometry (i.e., degree of mouth opening) encodes ewes arousal level in a context of mother-lamb separation (see chapter II) and that the speed of the laryngeal descent hypothesized as a possible cue to motivation, does not actually play a role in male-male communication in fallow deer (see annex to chapter III).

Therefore, in my thesis, I provide evidence of similarities in vocal anatomy and basic articulation gestures in species phylogenetically distant from humans. In this discussion, I highlight the possible selective forces driving the evolution of vocal anatomy and vocal

articulation in the studied ungulate species. Then, I argue how the fundamental frequency can modulate (i.e., attenuate or reinforce) the information encoded by formants, while also discussing my methodologies and their broader implication for our understanding of vocal communication in mammals.

Selective forces driving the evolution of the vocal apparatus

The low resting position of the larynx

In my thesis, I studied three species from the *cervinae* sub-family the sika deer (chapter I), the red deer (chapter I) and the fallow deer (chapter III). Like most mammal species sika deer possess a larynx in high resting position whereas red deer and fallow deer possess a larynx in low resting position (Frey et al., 2021). This dimorphism reported in closely related species have also been described in primates where, to the best of my knowledge, only human possess a larynx in low resting position (Fitch & Reby, 2001), thus suggesting that the low resting position of the larynx is an evolutionary convergence.

In primate species, the low resting position of the larynx, as only observed in humans initially led to the hypothesis that the laryngeal position evolved with the emergence of articulated speech as it allows higher range of articulation gestures (Lieberman et al., 1969). However, this theory has been later refuted by several studies showing that non-human primate species are also able of articulation gestures even with a larynx positioned relatively high in their throat (Boë et al., 2017; Fitch et al., 2016). This is also supported by my thesis findings as I demonstrated that basic articulation gestures, such as the mouth opening in sheep (chapter II) and the laryngeal descent in fallow deer (chapter III), are sufficient to significantly modulate formant frequencies.

To explain evolutionary convergences towards a low larynx position in terrestrial mammals, another theory suggested by Fitch & Reby, (2001) proposes that the development of such vocal apparatus may function for the communication of size exaggeration by male vocalisers. This theory is indeed supported in deer species, where sexual selection has favored the evolution towards relatively large males, and in humans, where sex differences in vocal tract length

appear after the puberty and do not impact the phonetic abilities whereas the lowering of males' larynx increases their perceived size (Fitch & Giedd, 1999).

Mouth opening in sheep

In the second chapter of my thesis, I studied the mechanisms of vocal production of the low and high bleats in ewes. I found that, both oral and nasal tracts are engaged into the vocal production of both types of calls. Due to the difference in the mouth configuration (i.e., open in high bleats and closed in low bleats), formant F1 was significantly higher in high bleats than in low bleats. This finding is consistent with studies conducted on humans, where it has been shown that the frequency of F1 depends on the anterior oral cavity area (Fant, 1960). When the mouth opens, the pressure at the anterior oral cavity decreases and thus the frequency of F1 increases. More specifically, mouth opening occurs in syllables comprising bilabial consonant (i.e., produced by the articulation of both lips) followed by a vowel. This articulation gesture induces modification in the vocal tract from a position (closed mouth) to another one (open mouth) inducing a progressive change in formant frequencies. For example, when producing the nasalized consonant /m/ followed by the vowel /ae/, first, the velum lowers to engage the nasal tract in addition of the oral tract, then the velum gradually raises until blocking the nasal tract to involve the oral tract only to produce the vowel /ae/.

In contrast to humans, sheep possess a relatively simple vocal repertoire that are mostly represented by mouth opening when vocalizing. This relatively simple vocal repertoire and the relatively low level of vocal sexual dimorphism can be a consequence of domestication which is known to affect selection pressures on species (e.g., lower predation pressures in domesticated versus wild species) (Zohary et al., 1998). Indeed, sheep were domesticated 8-10000 years ago for their wool, pelts, meat and milk (Nowak et al., 2008). Domestication has led to artificial selection made by humans which has induced morphological, behavioral and physiological differentiations of domestic sheep from wild ancestors (Zohary et al., 1998). For example, protection against predators and human-controlled reproduction have lowered sexual dimorphism in the overall body size of sheep (Polak & Frynta, 2009). Hence, in contrast to humans and fallow deer, the lack of sexual selection on their vocal behavior may have led domesticated sheep to evolve only basic vocal apparatus (i.e., high laryngeal resting position) and articulatory gestures such as the mouth opening essential for the mother-lamb communication (Sèbe et al., 2010).

Laryngeal descent in fallow deer

While most mammal species possess a larynx in high position as we saw in sheep, wild animals that are facing different selective pressures may develop specific anatomical structures in their vocal apparatus (Ryan, 1986). In fallow deer, the vocal behavior of males is likely to be under strong sexual pressures linked to a relatively strong competition for female access (Clutton-Brock et al., 1988). This has probably led to the development of low larynx position in vocalizers to exaggerate the communication of vocal cues to size (Fitch & Reby, 2001). In addition to their larynx characterized by a low resting position, fallow deer males are able to perform laryngeal movement consisting in a retraction of their larynx toward the sternum (Fitch & Reby, 2001; McElligott et al., 2006). I showed in chapter III that the laryngeal descent induces the elongation of the vocal tract and affects formant frequencies by decreasing their spacing. This articulation can be compared to the laryngeal descent that occurs in humans when producing two consecutive vowels such as /i/ and /u/ (Hoole & Kroos, 1998). Indeed, the vowel /i/ is produced when the larynx is relatively higher in the throat compared to the vowel /u/ inducing similar formants modulation (Hillenbrand et al., 1995).

While formants convey cues to the behavioral context in fallow deer buck's call (Charlton & Reby, 2011), the communicative function of the formant modulation remains unknown. In the annex of the chapter III, I investigated the function of laryngeal descent speed in a context of male-male competition for females during the rutting period and I hypothesized that the speed of formant modulation when lowering the larynx would influence the behavioral response in rivals. My playback experiments do not show an influence of the descent speed on the behavior of fallow deer bucks suggesting that the modulation of formants likely plays a rudimental role in male-male competition. Further investigations are needed to explore the response of female to this modulation.

Effect of the source on the filter

According to the source-filter theory, the source (i.e., larynx) and the filter (i.e., vocal tract) are independent and encode information including static (i.e., identity, size, sex) and dynamic (i.e., motivations, emotions) information (Taylor et al., 2016). My thesis mainly focuses on the filter

to identify precursors of human articulation (see chapter II on the mouth opening and chapter III on the laryngeal descent). I found that the filter plays a crucial role for the modulation of formants between call types (low and high bleats in sheep) and within a single call type (groans in fallow deer). While the formant modulation allows to produce a higher variety of sounds to encode multiple information, such as the arousal level via the degree of mouth opening in sheep, this information can be reinforced or attenuated by the fundamental frequency (Ryalls & Lieberman, 1982).

My study (chapter II) showed that ewe's vocalization can encode for the arousal level through the degree of mouth opening in context of mother-young separation. Moreover, I found that the first formant and the f_0 are higher when the mouth is opened than when the mouth is closed. Hence, both source- and filter-related parameters convey information about the emitter level of arousal. The reinforcement of this information highlights the necessity to transmit clearly this information. It would be interesting to do playback experiments on lambs to investigate whether they are able perceive the modification of the acoustic parameters (i.e., raise F1 and F5) while the arousal level of the ewe changes.

In fallow deer buck groan, the f_0 is relatively low (28.2 ± 0.3 Hz; Vannoni & McElligott, 2007). As such, the spectral density of these vocalizations is relatively high and formants strongly excited and perceptually salient (Kewley-Port et al., 1996; Ryalls & Lieberman, 1982). It is now well established that both f_0 and formants play a crucial role in vocal communication in fallow deer males. Indeed, while f_0 and formants encode information about the male quality and body size, these two vocal information are also complementary to inform receivers about the caller's dominance (Briefer et al., 2010; Vannoni & McElligott, 2008).

In the first chapter of my thesis, I discuss the possible contradictory information between the source and the filter in red deer species. Indeed, if females red deer prefer males with high fundamental frequency, linked to small size in a lot of species, they are also attracted by low formant spacing, which is linked to big size (Charlton & Reby, 2016). Thus, this species may face a tradeoff between the production of high f_0 and low formant frequencies. This could explain the development of relatively short vocal folds length compared to their body size (Frey et al., 2021) contrasted by the very low laryngeal position in red deer males (chapter I).

Hence, it is important to take into account information encoded by both f_0 and formants when studying vocal production and the communicative function of vocalizations.

New tools for studying the mammal vocal production

The development of new methodologies has allowed to improve our knowledge about the vocal production and its evolution across species. Indeed, for long time, vocal tract movements were considered as a uniquely human ability and thus studies were focused on human vocal production.

Through this multi-disciplinary thesis, I used various methods to study the vocal production from the laryngeal development (chapter I), the mechanism of articulatory gestures (the mouth opening in sheep chapter II and the laryngeal descent in fallow deer chapter III) and the function of articulatory gestures (annex of the chapter III). Here, I discuss the pertinence of our methods to study animal vocal production.

Behavioral and acoustic recordings

In order to link the vocal activity of ewes with a specific context, in the chapter II of my thesis, I recorded behavior and vocalizations produced by ewes in a specific context (i.e., separation). This method is commonly used in ethology and bioacoustics to associate the production of vocalization and specific behavior (Gerhardt, 1998). In my study, I controlled for the context with an isolation of the focal animal from the group to be able to record only the focal ewe. This isolation can affect the natural behavior of the animal as it changes its biological condition (i.e., live in group for sheep). As such, the isolation of the mother from its group might affect the stress level of the ewe during the experiment and thus induces a variability in the ewe's arousal level during the basal condition (i.e., the phase just before the separation between mother and lamb). Physiological measurements such as electrocardiogram would be great tools to confirm the stress level of the tested ewes, however, it also needs a long period of habituation to the physiological measurement equipment such as leather belt (Reefmann, Wechsler, et al., 2009).

Behavioral and acoustic recordings can also be done in the field which present the advantage to study the animal in their natural environment. However, recordings in the field can be challenging because of the weak signal-to-noise ratio, the absence of the animals and the accessibility to approach sufficiently the animals.

To assess to the function of vocalization playback experiment is broadly used in the literature (McGregor, 1992). This method consists of replaying recordings of natural or synthesized signals to animals and to record their response. In the annex of chapter III, I did playback experiments on fallow deer bucks to assess the information perceived by other males. This method requires ensuring the nature of the acoustic parameters tested (McGregor, 1992). For example, to test the function of the dynamic of formant in fallow deer groan, I synthesized vocalizations to normalize the speed of the formant lowering, the fundamental frequency and the minimum frequency of formants. The context in which the playback is broadcasted is also important as the behavioral response depends on the motivation of the focal animals. In my experiment, I tried to control for the context by considering the season (i.e., the rutting period), the behavior of the tested male before the exposure (i.e., mobile) and the presence of females. However, other factors can play in the animal response such as the distance with other males and the place from where the playback is played (e.g., territory of another male), therefore it is important to replicate the experiments.

CT-scan to assess the anatomy of vocal apparatus

In all of the three chapter of my thesis I used computed tomography scanner (CT-scan) either on anesthetized animals to assess the anatomy of the larynx size and position in sika and red deer (chapter I) or on dead specimens to assess the vocal tract anatomy in order to do acoustic modeling (chapter II and III). To assess the vocal apparatus anatomy several methods can be employed such as radiograph (e.g., domestic dog (*Canis familiaris*): Riede & Fitch, 1999), magnetic resonance imaging (MRI) (e.g., human: Badin et al., 1995) or CT-scan (e.g., fallow deer: Reby et al., 2018). The use of CT-scan presents the advantage to give an accurate measurement contrary to radiograph and is faster than acquisition from MRI.

CT-scan on dead specimens present the advantage to change the configuration of their vocal tract to be close to the natural configuration that animals performed while vocalize. Hence, this method allowed me to show that while the mouth is closed in sheep the oral cavity can remain hollow (chapter II) and that the extension of the vocal tract of fallow deer decrease of the differential length between the oral and nasal tract (chapter III). Based on my CT-scan acquisitions I was able to make acoustic modelling that predicted accurately the formant frequencies observed in natural vocalizations. However, this method has certain limitations such as the problem of the cadaveric rigidity, which made it impossible for me to open widely

the mouth of sheep (chapter II), and the fact that it does not take into account natural muscle contraction/relaxation that occurs in natural vocal behavior (Fitch, 2000).

An alternative to this method is to do cineradiography on alive animals to take account for the internal changes when animal vocalize. This method consists of several radiography while the animal vocalizes and was applied on several species (Badin et al., 1995; Fitch, 2000; Ghazanfar et al., 2012) to see in real time movement of the vocal tract such has the velum or laryngeal movements (Fitch, 2000). The major inconvenient of this method is that it exposes the subject to high level of radiation (Reynolds, 1934) and requires habituation and training of the animals.

Post-mortem recordings

During this thesis I developed an innovative method, called the *post-mortem recordings*, to study the influence of animal articulatory gestures on the modulation of formants. This method consists in flowing air into the back of the trachea of a head specimen positioned on a metal structure while an experimenter is in charge of triggering vocal fold oscillation by applying manual pressure on the larynx. This method successfully predicted the formant frequencies naturally observed in vocalization of ewes (chapter II) and fallow deer bucks (chapter III). In future research this *post-mortem recordings* method could be applied to a wider range of articulation gestures such as tongue or lips movements which are known to played an important role in human vocalizations (Titze, 1994).

Like CT-scan on post-mortem specimens, this approach has the inconvenient to not consider the natural role of muscles involved in the vocal production. Thus, to understand the overall changes involved by vocal production it will be interesting to compare our results to cineradiography on vocalizing animals.

Impedance sensor

In the second chapter of my thesis, I used an impedance sensor, developed by LAUM (Laboratoire d'acoustique de l'Université du Mans) and CTTM (Centre de Transfert de Technologie du Mans). The impedance sensor is a device including a piezoelectric speaker as a sound source and two microphones measuring the sound pressure (Le Roux et al., 2012). This tool is commonly used to predict resonant frequencies in wind instruments and to assess human

tracheal areas from measurements at the lips (Dalmont, 2001). To my knowledge, this is the first time that such a tool is used at the glottis to predict formant frequencies. One limitation is the need of an adaptor between the device and the glottis which should be adjusted to each head specimen. However, this method presents a major advantage as it allows to assess very quickly and accurately the formant frequencies observed in natural ewe's vocalizations (chapter II). Future experiments could apply this tool in other species such as fallow deer with different laryngeal position.

Moreover, it would be useful to test this tool on 3D printed model of vocal tract to assess formant frequencies without manipulation of the head specimen.

Conclusion

The fascination for the human articulation ability started from the Greek Antiquity, therefore, the mechanisms of the human vocal production have been well described. However, the implication of the androgens on sexual dimorphism in this vocal apparatus remains unknown. Using the comparative method on deer species, my findings provide the first investigation of the effect of androgens on the development of the sexually dimorphic vocal apparatuses. Moreover, while the evolution of complex articulation, characterizing the human speech, is well accepted to be unique in mammal species, I demonstrated the importance of basic articulation gestures in natural vocal communication of non-primate mammal species (i.e., mouth opening in sheep and laryngeal descent in fallow deer). Thus, using robust methods and the development of innovative methodologies, this multi-disciplinary thesis provides a useful background for our understanding of the evolutionary origins of human vocal anatomy and vocal production. It also opens the way to further investigations into the evolution of the vocal apparatus in non-human mammals.

References

- Alipour, F., & Jaiswal, S. (2008). Phonatory characteristics of excised pig, sheep, and cow larynges. *The Journal of the Acoustical Society of America*, *123*(6), 4572-4581.
- Alipour, F., Jaiswal, S., & Finnegan, E. (2007). Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models. *Annals of Otology, Rhinology & Laryngology*, *116*(2), 135-144.
- Aufdemorte, T. B., Sheridan, P. J., & Holt, G. R. (1983). Autoradiographic evidence of sex steroid receptors in laryngeal tissues of the baboon (*Papio cynocephalus*). *The Laryngoscope*, *93*(12), 1607-1611.
- Bachorowski, J.-A., & Owren, M. J. (2008). Vocal expressions of emotion. *Handbook of emotions*, *3*, 196-210.
- Badin, P., Gabioud, B., Beutemps, D., Lallouache, T., Bailly, G., Maeda, S., Zerling, J.-P., & Brock, G. (1995). Cineradiography of VCV sequences : Articulatory-acoustic data for a speech production model. *International Congress on Acoustics*, 349-352.
- Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging : Vowels. *The Journal of the Acoustical Society of America*, *90*(2), 799-828. <https://doi.org/10.1121/1.401949>
- Bauer, H. R. (1987). Frequency code : Orofacial correlates of fundamental frequency. *Phonetica*, *44*(3), 173-191.
- Beckers, G. J. L., Nelson, B. S., & Suthers, R. A. (2004). Vocal-Tract Filtering by Lingual Articulation in a Parrot. *Current Biology*, *14*(17), 1592-1597. <https://doi.org/10.1016/j.cub.2004.08.057>
- Beckford, N. S., Schaid, D., Rood, S. R., & Schanbacher, B. (1985). Androgen Stimulation and Laryngeal Development. *Annals of Otology, Rhinology & Laryngology*, *94*(6), 634-640. <https://doi.org/10.1177/000348948509400622>
- Behrman, A. (2021). *Speech and voice science*. Plural publishing.
- Bendall, J. R. (1960). Postmortem changes in muscle. *The structure and function of muscle*, *2*(Part 1), 243-309.

- Blant, M., & Gaillard, J. M. (2004). Use of biometric body variables as indicators of roe deer (*Capreolus capreolus*) population density changes. *Game and Wildlife Science*, 21(1), 21-40.
- Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., Becker, Y., Rey, A., & Fagot, J. (2017). Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech Precursors. *PLOS ONE*, 12(1), e0169321. <https://doi.org/10.1371/journal.pone.0169321>
- Boë, L.-J., Sawallis, T. R., Fagot, J., Badin, P., Barbier, G., Captier, G., Ménard, L., Heim, J.-L., & Schwartz, J.-L. (2019). Which way to the dawn of speech? : Reanalyzing half a century of debates and data in light of speech science. *Science Advances*, 5(12), eaaw3916.
- Boyd, S. K., Wissing, K. D., Heinsz, J. E., & Prins, G. S. (1999). *Androgen Receptors and Sexual Dimorphisms in the Larynx of the Bullfrog*. 10.
- Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*.
- Briefer, E. F. (2020). Coding for ‘dynamic’ information : Vocal expression of emotional arousal and valence in non-human animals. In *Coding strategies in vertebrate acoustic communication* (p. 137-162). Springer.
- Briefer, E. F., Tettamanti, F., & McElligott, A. G. (2015). Emotions in goats : Mapping physiological, behavioural and vocal profiles. *Animal Behaviour*, 99, 131-143. <https://doi.org/10.1016/j.anbehav.2014.11.002>
- Briefer, E., Vannoni, E., & McElligott, A. G. (2010). Quality prevails over identity in the sexually selected vocalisations of an ageing mammal. *BMC biology*, 8(1), 1-15.
- Brinkmann, A. O., Blok, L. J., de Ruiter, P. E., Doesburg, P., Steketee, K., Berrevoets, C. A., & Trapman, J. (1999). Mechanisms of androgen receptor activation and function. *The Journal of Steroid Biochemistry and Molecular Biology*, 69(1), 307-313. [https://doi.org/10.1016/S0960-0760\(99\)00049-7](https://doi.org/10.1016/S0960-0760(99)00049-7)
- Brunings, J. W., Schepens, J. J., Peutz-Kootstra, C. J., & Kross, K. W. (2013). The expression of estrogen and progesterone receptors in the human larynx. *Journal of Voice*, 27(3), 376-380.
- Carterette, E. C., Shipley, C., & Buchwald, J. S. (1979). Linear prediction theory of vocalization in cat and kitten. *Frontiers in speech communication research*, 245-257.

- Chaigne, A., & Kergomard, J. (2016). *Acoustics of musical instruments*. Springer.
- Charlton, B. D., Ellis, W. A. H., Brumm, J., Nilsson, K., & Fitch, W. T. (2012). Female koalas prefer bellows in which lower formants indicate larger males. *Animal Behaviour*, *84*(6), 1565-1571. <https://doi.org/10.1016/j.anbehav.2012.09.034>
- Charlton, B. D., Ellis, W. A. H., McKinnon, A. J., Cowin, G. J., Brumm, J., Nilsson, K., & Fitch, W. T. (2011). Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows : Honesty in an exaggerated trait. *Journal of Experimental Biology*, *214*(20), 3414-3422. <https://doi.org/10.1242/jeb.061358>
- Charlton, B. D., & Reby, D. (2011). Context-Related Acoustic Variation in Male Fallow Deer (*Dama dama*) Groans. *PLOS ONE*, *6*(6), e21066. <https://doi.org/10.1371/journal.pone.0021066>
- Charlton, B. D., & Reby, D. (2016). The evolution of acoustic size exaggeration in terrestrial mammals. *Nature Communications*, *7*(1), Article 1. <https://doi.org/10.1038/ncomms12739>
- Charlton, B. D., Reby, D., & McComb, K. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*(4), 382-385. <https://doi.org/10.1098/rsbl.2007.0244>
- Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2010). Giant pandas perceive and attend to formant frequency variation in male bleats. *Animal Behaviour*, *79*(6), 1221-1227. <https://doi.org/10.1016/j.anbehav.2010.02.018>
- Charrier, I., Ahonen, H., & Harcourt, R. G. (2011). What makes an Australian sea lion (*Neophoca cinerea*) male's bark threatening? *Journal of Comparative Psychology*, *125*(4), 385.
- Charuta, A., Dzierżęcka, M., & Wysocki, J. (2009). Evaluation of sexual dimorphism in horses on the basis of the morphology and morphometry of the larynx. *Bulletin of the Veterinary Institute in Pulawy*, *53*(3), 477-486.
- Clutton-Brock, T. H., Green, D., Hiraiwa-Hasegawa, M., & Albon, S. D. (1988). Passing the buck : Resource defence, lek breeding and mate choice in fallow deer. *Behavioral Ecology and Sociobiology*, *23*(5), 281-296. <https://doi.org/10.1007/BF00300575>
- Clutton-Brock, T. H., & Vincent, A. C. (1991). Sexual selection and the potential reproductive rates of males and females. *Nature*, *351*(6321), 58-60.
- Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among

men. *Personality and Individual Differences*, 27(4), 801-804. [https://doi.org/10.1016/S0191-8869\(98\)00272-4](https://doi.org/10.1016/S0191-8869(98)00272-4)

Dalmont, J.-P. (2001). Acoustic impedance measurement, part I : A review. *Journal of Sound and Vibration*, 243(3), 427-439. <https://doi.org/10.1006/jsvi.2000.3428>

Dang, J., & Honda, K. (2002). Estimation of vocal tract shapes from speech sounds with a physiological articulatory model. *Journal of Phonetics*, 30(3), 511-532. <https://doi.org/10.1006/jpho.2002.0167>

de Boer, B. (2008). The acoustic role of supralaryngeal air sacs. *The Journal of the Acoustical Society of America*, 123(5), 3779-3779. <https://doi.org/10.1121/1.2935420>

de Boer, B., & Fitch, W. T. (2010). Computer Models of Vocal Tract Evolution : An Overview and Critique. *Adaptive Behavior*, 18(1), 36-47. <https://doi.org/10.1177/1059712309350972>

Dwyer, C. M., McLean, K. A., Deans, L. A., Chirnside, J., Calvert, S. K., & Lawrence, A. B. (1998). Vocalisations between mother and young in sheep : Effects of breed and maternal experience. *Applied Animal Behaviour Science*, 58(1), 105-119. [https://doi.org/10.1016/S0168-1591\(97\)00113-5](https://doi.org/10.1016/S0168-1591(97)00113-5)

Eaton, R. L. (1980). A beluga whale imitates human speech. *Carnivore*, 2(NOV), 22-23.

Evans, S., Neave, N., & Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males : An evolutionary explanation for a deep male voice. *Biological Psychology*, 72(2), 160-163. <https://doi.org/10.1016/j.biopsycho.2005.09.003>

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4-5), 783-788.

Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton.

Fant, G. (1980). The Relations between Area Functions and the Acoustic Signal. *Phonetica*, 37(1-2), 55-86. <https://doi.org/10.1159/000259983>

Feighny, J. A., Williamson, K. E., & Clarke, J. A. (2006). North American elk bugle vocalizations : Male and female bugle call structure and context. *Journal of Mammalogy*, 87(6), 1072-1077.

Ferrein, A. (1741). De la formation de la voix de l'homme. *Hist Acad Roy Sci*, 3, 409-432.

Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2002). Acoustic features of male baboon loud calls : Influences of context, age, and individuality. *The Journal of the Acoustical Society of America*, *111*(3), 1465-1474.

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. *Behavioral ecology and sociobiology*, *56*(2), 140-148.

Fischer, L. M., Catz, D., & Kelley, D. B. (1995). Androgen-directed development of the *Xenopus laevis* larynx : Control of androgen receptor expression and tissue differentiation. *Developmental biology*, *170*(1), 115-126.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Acoustical Society of America*, *10*.

Fitch, W. T. (1999). Acoustic exaggeration of size in birds via tracheal elongation : Comparative and theoretical analyses. *Journal of Zoology*, *248*(1), 31-48. <https://doi.org/10.1111/j.1469-7998.1999.tb01020.x>

Fitch, W. T. (2000). The Phonetic Potential of Nonhuman Vocal Tracts : Comparative Cineradiographic Observations of Vocalizing Animals. *Phonetica*, *57*(2-4), 205-218. <https://doi.org/10.1159/000028474>

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*(12), e1600723. <https://doi.org/10.1126/sciadv.1600723>

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract : A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 1511-1522. <https://doi.org/10.1121/1.427148>

Fitch, W. T., & Hauser, M. D. (2003). Unpacking “honesty” : Vertebrate vocal production and the evolution of acoustic signals. In *Acoustic communication* (p. 65-137). Springer.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1477), 1669-1675. <https://doi.org/10.1098/rspb.2001.1704>

Fitzgerald, E. R. (1975). Dynamic mechanical measurements during the life to death transition in animal tissues. *Biorheology*, *12*(6), 397-408. <https://doi.org/10.3233/bir-1975-12611>

- Geist, V., & Bayer, M. (1988). Sexual dimorphism in the *Cervidae* and its relation to habitat. *Journal of Zoology*, 214(1), 45-53. <https://doi.org/10.1111/j.1469-7998.1988.tb04985.x>
- Gerhardt, H. C. (1998). Acoustic Signals of Animals: Recording, Field Measurements, Analysis and Description. In S. L. Hopp, M. J. Owren, & C. S. Evans (Éds.), *Animal Acoustic Communication: Sound Analysis and Research Methods* (p. 1-25). Springer. https://doi.org/10.1007/978-3-642-76220-8_1
- Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., & Fitch, W. T. (2012). Cineradiography of Monkey Lip-Smacking Reveals Putative Precursors of Speech Dynamics. *Current Biology*, 22(13), 1176-1182. <https://doi.org/10.1016/j.cub.2012.04.055>
- Gilbert, J., Dalmont, J.-P., Potier, R., & Reby, D. (2014). Is Nonlinear Propagation Responsible for the Brassiness of Elephant Trumpet Calls? *Acta Acustica United with Acustica*, 100(4), 734-738. <https://doi.org/10.3813/AAA.918752>
- Greenwood, P. J., & Adams, J. (1987). Sexual Selection, Size Dimorphism and a Fallacy. *Oikos*, 48(1), 106-108. <https://doi.org/10.2307/3565694>
- Harris, T. R., Fitch, W. T., Goldstein, L. M., & Fashing, P. J. (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. <https://doi.org/10.1111/j.1439-0310.2006.01247.x>
- Hattori, S., Yamamoto, K., & Fujimura, O. (1958). Nasalization of vowels in relation to nasals. *The Journal of the Acoustical Society of America*, 30(4), 267-274.
- Hauser, M. D. (2000). The sound and the fury : Primate vocalizations as reflections of emotion and thought. *The origins of music*, 77-102.
- Hauser, M. D., Evans, C. S., & Marler, P. (1993). The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations. *Animal behaviour*, 45(3), 423-433.
- Hayama, S. (1970). The saccus laryngis in primates. *Journal of the Anthropological Society of Nippon*, 78(4), 274-298.
- Herbst, C. T., Švec, J. G., Lohscheller, J., Frey, R., Gumpenberger, M., Stoeger, A. S., & Fitch, W. T. (2013). Complex vibratory patterns in an elephant larynx. *Journal of Experimental Biology*, 216(21), 4054-4064.

- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97(5), 3099-3111.
- Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia phoniatrica et logopaedica*, 26(2), 89-94.
- Holt, G. R., Aufdemorte, T. B., & Sheridan, P. J. (1986). Estrogen receptor in the larynx of the aged baboon (*Papio cynocephalus*). *Annals of Otology, Rhinology & Laryngology*, 95(6), 608-617.
- Hoole, P., & Kroos, C. (1998). Control of larynx height in vowel production. *5th International Conference on Spoken Language Processing (ICSLP 1998)*, paper 1097-0. <https://doi.org/10.21437/ICSLP.1998-360>
- Hunt, J., Breuker, C. J., Sadowski, J. A., & Moore, A. J. (2009). Male–male competition, female mate choice and their interaction : Determining total sexual selection. *Journal of Evolutionary Biology*, 22(1), 13-26. <https://doi.org/10.1111/j.1420-9101.2008.01633.x>
- IPA Chart with Sounds – International Phonetic Alphabet Sounds*. (s. d.). Consulté 2 décembre 2022, à l'adresse <https://www.internationalphoneticalphabet.org/ipa-sounds/ipa-chart-with-sounds/>
- Jallon, J. F., & Berthommier, F. (2006). Semi-automatic extraction of vocal tract movements from cineradiographic data. *INTERSPEECH 2006-ICSLP*, 569-572.
- Joliveau, E., Smith, J., & Wolfe, J. (2004). Tuning of vocal tract resonance by sopranos. *Nature*, 427(6970), 116-116. <https://doi.org/10.1038/427116a>
- Kahane, J. C. (1982). Growth of the human prepubertal and pubertal larynx. *Journal of Speech, Language, and Hearing Research*, 25(3), 446-455.
- Keenan, S., Mathevon, N., Stevens, J. M. G., Nicolè, F., Zuberbühler, K., Guéry, J.-P., & Levréro, F. (2020). The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Animal Behaviour*, 169, 9-21. <https://doi.org/10.1016/j.anbehav.2020.08.024>
- Kempster, G. B., Larson, C. R., & Kistler, M. K. (1988). Effects of electrical stimulation of cricothyroid and thyroarytenoid muscles on voice fundamental frequency. *Journal of Voice*, 2(3), 221-229.

- Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of speech and hearing disorders*, 52(4), 367-387.
- Kewley-Port, D., Li, X., Zheng, Y., & Neel, A. T. (1996). Fundamental frequency effects on thresholds for vowel formant discrimination. *The Journal of the Acoustical Society of America*, 100(4), 2462-2470.
- Kim, J. M., Shin, S.-C., Park, G.-C., Lee, J.-C., Jeon, Y. K., Ahn, S. J., Thibeault, S., & Lee, B.-J. (2020). Effect of sex hormones on extracellular matrix of *lamina propria* in rat vocal fold. *The Laryngoscope*, 130(3), 732-740. <https://doi.org/10.1002/lary.28086>
- Kim, Y.-C., Narayanan, S. S., & Nayak, K. S. (2009). Accelerated 3D MRI of vocal tract shaping using compressed sensing and parallel imaging. *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 389-392.
- Le Roux, J.-C., Pachebat, M., & Dalmont, J.-P. (2012). A new impedance sensor for industrial applications. In S. F. d'Acoustique (Éd.), *Acoustics 2012*. <https://hal.archives-ouvertes.fr/hal-00810602>
- Lieberman, P. (1991). *Uniquely Human: The Evolution of Speech, Thought, and Selfless Behavior*. Harvard University Press.
- Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics*, 40(4), 608-622. <https://doi.org/10.1016/j.wocn.2012.04.001>
- Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press.
- Lieberman, P., Klatt, D., & Wilson, W. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, 164(3884), 1185-1187.
- Lin, E. C. (2010). Radiation risk from medical imaging. *Mayo Clinic Proceedings*, 85(12), 1142-1146.
- Lindau, M. (1979). The feature expanded. *Journal of Phonetics*, 7(2), 163-176.
- Maytag, A. L., Robitaille, M. J., Rieves, A. L., Madsen, J., Smith, B. L., & Jiang, J. J. (2013). Use of the rabbit larynx in an excised larynx setup. *Journal of voice: official journal of the Voice Foundation*, 27(1), 24-28. <https://doi.org/10.1016/j.jvoice.2012.08.004>

- McElligott, A. G., Birrer, M., & Vannoni, E. (2006). Retraction of the mobile descended larynx during groaning enables fallow bucks (*Dama dama*) to lower their formant frequencies. *Journal of Zoology*, 270(2), 340-345. <https://doi.org/10.1111/j.1469-7998.2006.00144.x>
- McElligott, A. G., & Hayden, T. J. (1999). Context-related vocalization rates of fallow bucks, *Dama dama*. *Animal Behaviour*, 58(5), 1095-1104. <https://doi.org/10.1006/anbe.1999.1237>
- McGregor, P. K. (Éd.). (1992). *Playback and studies of animal communication*. Plenum Press.
- Minami, M., & Kawamichi, T. (1992). Vocal repertoires and classification of the sika deer *Cervus nippon*. *Journal of the Mammalogical Society of Japan*, 17(2), 71-94.
- Moore, C., Shalet, S., Manickam, K., Willard, T., Maheshwari, H., & Baumann, G. (2005). Voice abnormality in adults with congenital and adult-acquired growth hormone deficiency. *The Journal of Clinical Endocrinology & Metabolism*, 90(7), 4128-4132.
- Morgan, P. D., Boundy, C. A. P., Arnold, G. W., & Lindsay, D. R. (1975). The roles played by the senses of the ewe in the location and recognition of lambs. *Applied Animal Ethology*, 1(2), 139-150. [https://doi.org/10.1016/0304-3762\(75\)90083-8](https://doi.org/10.1016/0304-3762(75)90083-8)
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*.
- Mupparapu, M., & Vuppapapati, A. (2005). Ossification of laryngeal cartilages on lateral cephalometric radiographs. *The Angle Orthodontist*, 75(2), 196-201.
- Narbaitz, R., Stumpf, W. E., & Sar, M. (1980). Estrogen target cells in the larynx : Autoradiographic studies with 3H-diethylstilbestrol in fetal mice. *Hormone Research in Paediatrics*, 12(2), 113-117.
- Nowak, R., Porter, R. H., Blache, D., & Dwyer, C. M. (2008). Behaviour and the Welfare of the Sheep. In C. M. Dwyer (Éd.), *The Welfare of Sheep* (p. 81-134). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8553-6_3
- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech* (p. 189-216). Springer.
- Owens, I. P. F., & Hartley, I. R. (1998). Sexual dimorphism in birds : Why are there so many different forms of dimorphism? *Proceedings of the Royal Society of London. Series B:*

Biological Sciences, 265(1394), 397-407. <https://doi.org/10.1098/rspb.1998.0308>

Owings, D. H., & Morton, E. S. (1998). *Animal Vocal Communication : A New Approach*. Cambridge University Press.

Pepperberg, I. M., Brese, K. J., & Harris, B. J. (1991). Solitary sound play during acquisition of English vocalizations by an African Grey parrot (*Psittacus erithacus*) : Possible parallels with children's monologue speech. *Applied Psycholinguistics*, 12(2), 151-178. <https://doi.org/10.1017/S0142716400009127>

Perrier, P., Boë, L.-J., & Sock, R. (1992). Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast : Modeling the transition with two sets of coefficients. *Journal of Speech, Language, and Hearing Research*, 35(1), 53-67.

Pfefferle, D., & Fischer, J. (2006). Sounds and size : Identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*. *Animal Behaviour*, 72(1), 43-51. <https://doi.org/10.1016/j.anbehav.2005.08.021>

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice Modulation : A Window into the Origins of Human Vocal Control? *Trends in Cognitive Sciences*, 20(4), 304-318. <https://doi.org/10.1016/j.tics.2016.01.002>

Pitcher, B. J., Briefer, E. F., & McElligott, A. G. (2015). Intrasexual selection drives sensitivity to pitch, formants and duration in the competitive calls of fallow bucks. *BMC Evolutionary Biology*, 15(1), 149. <https://doi.org/10.1186/s12862-015-0429-7>

Poindron, P., Caba, M., Gomora, P., Krehbiel, D., & Beyer, C. (1994). *Reaction a la separation sociale et a la privation du jeune chez la brebis : Consequences pour l'entretien des ongles domestiques a la parturition*. <https://hal.inrae.fr/hal-02848701>

Polak, J., & Frynta, D. (2009). Sexual size dimorphism in domestic goats, sheep, and their wild relatives. *Biological Journal of the Linnean Society*, 98(4), 872-883.

Porges, S. W., & Lewis, G. F. (2010). The polyvagal hypothesis : Common mechanisms mediating autonomic regulation, vocalizations and listening. In *Handbook of Behavioral Neuroscience* (Vol. 19, p. 255-264). Elsevier.

Raine, J., Pisanski, K., & Reby, D. (2017). Tennis grunts communicate acoustic cues to sex and contest outcome. *Animal Behaviour*, 130, 47-55.

- Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, 63(5), 1050-1056.
- Reby, D., Charlton, B. D., Locatelli, Y., & McComb, K. (2010). Oestrous red deer hinds prefer male roars with higher fundamental frequencies. *Proceedings of the Royal Society B: Biological Sciences*, 277(1695), 2747-2753. <https://doi.org/10.1098/rspb.2010.0467>
- Reby, D., & McComb, K. (2003a). Anatomical constraints generate honesty : Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, 65, 519-530. <https://doi.org/10.1006/anbe.2003.2078>
- Reby, D., & McComb, K. (2003b). Vocal communication and reproduction in deer. *Advances in the Study of Behavior*, 33, 231-264.
- Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, 272(1566), 941-947. <https://doi.org/10.1098/rspb.2004.2954>
- Reby, D., Wyman, M. T., Frey, R., Charlton, B. D., Dalmont, J. P., & Gilbert, J. (2018). Vocal tract modelling in fallow deer : Are male groans nasalized? *Journal of Experimental Biology*, 221(17). <https://doi.org/10.1242/jeb.179416>
- Reefmann, N., Bütikofer Kaszàs, F., Wechsler, B., & Gygax, L. (2009). Physiological expression of emotional reactions in sheep. *Physiology & Behavior*, 98(1-2), 235-241. <https://doi.org/10.1016/j.physbeh.2009.05.017>
- Reefmann, N., Wechsler, B., & Gygax, L. (2009). Behavioural and physiological assessment of positive and negative emotion in sheep. *Animal Behaviour*, 78(3), 651-659. <https://doi.org/10.1016/j.anbehav.2009.06.015>
- Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F 0) and formant profiles of human vowels and vowel-like baboon grunts : The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117(2), 944-955.
- Reynolds, R. J. (1934). Cineradiography. *The British Journal of Radiology*, 7(79), 415-424.
- Riede, T., Bronson, E., Hatzikirou, H., & Zuberbühler, K. (2005). Vocal production mechanisms in a non-human primate : Morphological data and a model. *Journal of Human*

Evolution, 48(1), 85-96.

Riede, T., & Fitch, W. T. (1999). Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). *Journal of Experimental Biology*, 202, 2859-2867.

Riede, T., Tokuda, I. T., Munger, J. B., & Thomson, S. L. (2008). Mammalian laryngeal air sacs add variability to the vocal tract impedance : Physical and computational modeling. *The Journal of the Acoustical Society of America*, 124(1), 634-647. <https://doi.org/10.1121/1.2924125>

Ryalls, J. H., & Lieberman, P. (1982). Fundamental frequency and vowel perception. *The Journal of the Acoustical Society of America*, 72(5), 1631-1634. <https://doi.org/10.1121/1.388499>

Ryan, M. J. (1986). Factors Influencing the Evolution of Acoustic Communication : Biological Constraints. *Brain, Behavior and Evolution*, 28(1-3), 70-82. <https://doi.org/10.1159/000118693>

Saez, S., & Sakai, F. (1976). Androgen receptors in human pharyngo-laryngeal mucosa and pharyngo-laryngeal epithelioma. *Journal of Steroid Biochemistry*, 7(11-12), 919-921.

Sanders, I., & Mu, L. (2013). A three-dimensional atlas of human tongue muscles. *The Anatomical Record*, 296(7), 1102-1114.

Scherer, K. R. (1986). Voice, stress, and emotion. In *Dynamics of stress* (p. 157-179). Springer.

Scheumann, M., Roser, A.-E., Konerding, W., Bleich, E., Hedrich, H.-J., & Zimmermann, E. (2012). Vocal correlates of sender-identity and arousal in the isolation calls of domestic kitten (*Felis silvestris catus*). *Frontiers in zoology*, 9(1), 1-14.

Schön, M. A. (1971). The anatomy of the resonating mechanism in howling monkeys. *Folia Primatologica*, 15(1-2), 117-132.

Schroeder, M. R. (1967). Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements. *The Journal of the Acoustical Society of America*, 41(4B), 1002-1010. <https://doi.org/10.1121/1.1910429>

Schwartz, M. F. (1968). The acoustics of normal and nasal vowel production. *The Cleft palate journal*, 5(2), 125-140.

Sèbe, F., Aubin, T., Boué, A., & Poindron, P. (2008). Mother–young vocal communication and

acoustic recognition promote preferential nursing in sheep. *Journal of Experimental Biology*, 211(22), 3554-3562. <https://doi.org/10.1242/jeb.016055>

Sèbe, F., Duboscq, J., Aubin, T., Ligout, S., & Poindron, P. (2010). Early vocal recognition of mother by lambs : Contribution of low-and high-frequency vocalizations. *Animal behaviour*, 79(5), 1055-1066.

Sèbe, F., Nowak, R., Poindron, P., & Aubin, T. (2007). Establishment of vocal communication and discrimination between ewes and their lamb in the first two days after parturition. *Developmental Psychobiology*, 49(4), 375-386. <https://doi.org/10.1002/dev.20218>

Shiple, C., Carterette, E. C., & Buchwald, J. S. (1991). The effects of articulation on the acoustical structure of feline vocalizations. *The Journal of the Acoustical Society of America*, 89(2), 902-909. <https://doi.org/10.1121/1.1894652>

Smith, J. M., & Price, G. R. (1973). The Logic of Animal Conflict. *Nature*, 246(5427), Article 5427. <https://doi.org/10.1038/246015a0>

Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). MIT press.

Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., Herbst, C. T., Kwon, S., & Fitch, W. T. (2012). An Asian elephant imitates human speech. *Current Biology*, 22(22), 2144-2148.

Stojanović, S., Belic, B., Stojanović, S., & Belic, B. (2013). Laryngeal Manifestations of Rheumatoid Arthritis. In *Innovative Rheumatology*. IntechOpen. <https://doi.org/10.5772/51730>

Sundberg, J., Birch, P., Gümöes, B., Stavvad, H., Prytz, S., & Karle, A. (2007). Experimental Findings on the Nasal Tract Resonator in Singing. *Journal of Voice*, 21(2), 127-137. <https://doi.org/10.1016/j.jvoice.2005.11.005>

Sundberg, J., & Rossing, T. D. (1990). *The science of singing voice*. Acoustical Society of America.

Takahashi, K., & Saito, H. (1979). Post-mortem changes in skeletal muscle connectin. *The Journal of Biochemistry*, 85(6), 1539-1542.

Takemoto, H., Honda, K., Masaki, S., Shimada, Y., & Fujimoto, I. (2006). Measurement of temporal changes in vocal tract area function from 3D cine-MRI data. *The Journal of the Acoustical Society of America*, 119(2), 1037-1049.

- Taylor, A. M., Charlton, B. D., & Reby, D. (2016). Vocal Production by Terrestrial Mammals : Source, Filter, and Function. In R. A. Suthers, W. T. Fitch, R. R. Fay, & A. N. Popper (Éds.), *Vertebrate Sound Production and Acoustic Communication* (p. 229-259). Springer International Publishing. https://doi.org/10.1007/978-3-319-27721-9_8
- Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research : Advances in vocal communication research. *Journal of Zoology*, 280(3), 221-236. <https://doi.org/10.1111/j.1469-7998.2009.00661.x>
- Taylor, A. M., Reby, D., & McComb, K. (2009). Context-Related Variation in the Vocal Growling Behaviour of the Domestic Dog (*Canis familiaris*). *Ethology*, 115(10), 905-915. <https://doi.org/10.1111/j.1439-0310.2009.01681.x>
- Taylor, A. M., Reby, D., & McComb, K. (2010). Size communication in domestic dog, *Canis familiaris*, growls. *Animal Behaviour*, 79(1), 205-210. <https://doi.org/10.1016/j.anbehav.2009.10.030>
- Terrazas, A., Nowak, R., Serafin, N., Ferreira, G., Lévy, F., & Poindron, P. (2002). Twenty-four-hour-old lambs rely more on maternal behavior than on the learning of individual characteristics to discriminate between their own and an alien mother. *Developmental Psychobiology*, 40(4), 408-418. <https://doi.org/10.1002/dev.10041>
- Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*, 85(2), 901-906.
- Titze, I. R. (1994). *Principles of voice production*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey,.
- Titze, I. R., Fitch, W. T., Hunter, E. J., Alipour, F., Montequin, D., Armstrong, D. L., McGee, J., & Walsh, E. J. (2010). Vocal power and pressure–flow relationships in excised tiger larynges. *Journal of Experimental Biology*, 213(22), 3866-3873.
- Titze, I. R., Luschei, E. S., & Hirano, M. (1989). Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice*, 3(3), 213-224.
- Titze, I. R., Worley, A. S., & Story, B. H. (2011). *Source-Vocal Tract Interaction in Female Operatic Singing and Theater Belting*. 13.
- Tuohimaa, P. T., Kallio, S., Heinijoki, J., Aitasalo, K., Virolainen, E., Karma, P., & Tuohimaa,

P. J. (1981). Androgen receptors in laryngeal carcinoma. *Acta Oto-laryngologica*, 91(1-6), 149-154.

Van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of speech and hearing research*, 1(3), 227-244.

Vannoni, E., & McElligott, A. G. (2007). Individual Acoustic Variation in Fallow Deer (*Dama dama*) Common and Harsh Groans : A Source-Filter Theory Perspective. *Ethology*, 113(3), 223-234. <https://doi.org/10.1111/j.1439-0310.2006.01323.x>

Vannoni, E., & McElligott, A. G. (2008). Low Frequency Groans Indicate Larger and More Dominant Fallow Deer (*Dama dama*) Males. *PLOS ONE*, 3(9), e3113. <https://doi.org/10.1371/journal.pone.0003113>

Weissenhuber, G. E., Forstenpointner, G., Peters, G., Kübber-Heiss, A., & 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. <https://doi.org/10.1046/j.1469-7580.2002.00088.x>

Wolfe, J., Garnier, M., & Smith, J. (2009). Vocal tract resonances in speech, singing, and playing musical instruments. *HFSP journal*, 3(1), 6-23.

Wyman, M. T., Locatelli, Y., Charlton, B. D., & Reby, D. (2014). No preference in female sika deer for conspecific over heterospecific male sexual calls in a mate choice context. *Journal of Zoology*, 293(2), 92-99. <https://doi.org/10.1111/jzo.12123>

Zohary, D., Tchernov, E., & Horwitz, L. K. (1998). The role of unconscious selection in the domestication of sheep and goats. *Journal of Zoology*, 245(2), 129-135.