



**HAL**  
open science

## Comparative Anatomy of the Baboon and Human Vocal Tracts: Renewal of Methods, Data, and Hypotheses

Frédéric Berthommier, Louis-Jean Boë, Adrien Meguerditchian, Thomas R. Sawallis, Guillaume Captier

### ► To cite this version:

Frédéric Berthommier, Louis-Jean Boë, Adrien Meguerditchian, Thomas R. Sawallis, Guillaume Captier. Comparative Anatomy of the Baboon and Human Vocal Tracts: Renewal of Methods, Data, and Hypotheses. Louis-Jean Boë; Joël Fagot; Pascal Perrier; Jean-Luc Schwartz. Origins of human language: Continuities and discontinuities with nonhuman primates, pp.101-135, 2017, Speech Production and Perception, 978-3-631-73726-2. 10.3726/b12405 . hal-01649499v2

**HAL Id: hal-01649499**

**<https://hal.science/hal-01649499v2>**

Submitted on 5 Mar 2019

**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.

Frédéric Berthommier<sup>1</sup>, Louis-Jean Boë<sup>1</sup>,  
Adrien Meguerditchian<sup>2</sup>, Thomas R. Sawallis<sup>3</sup>,  
Guillaume Captier<sup>4</sup>

<sup>1</sup>*GIPSA-Lab, CNRS and Grenoble Alpes University,  
Saint-Martin-d'Hères, France*

<sup>2</sup>*Cognitive Psychology Laboratory, CNRS and  
Aix-Marseille University, Marseille, France*

<sup>3</sup>*New College, The University of Alabama, Tuscaloosa,  
Alabama, United States of America*

<sup>4</sup>*Anatomy Laboratory, Montpellier University,  
Montpellier, France*

## Comparative Anatomy of the Baboon and Human Vocal Tracts: Renewal of Methods, Data, and Hypotheses

**Abstract:** This chapter focuses on the emergence of speech during human evolution, revisiting exaptation hypotheses (Fitch, 2010; MacNeilage, 1998) with new data from comparison with baboons. Speech necessarily evolved to be compatible with aerodigestive anatomy, reusing its functions of suction, chewing and swallowing. The tongue is involved with every feeding gesture, and also has a central position for speech. We analyze the evolution of the tongue position taking into account the distinction between the morphogenetic fields of HOX and non-HOX genes involved in the development of the pharyngeal arches and the cephalic structures, anatomical and neurological components, and functional support for breathing and swallowing. The hyoid bone is the locus of insertion of the tongue muscles as well as a precise marker of the glottis position. It is not fixed because it partly depends on the development of the facial area controlled by non-HOX genes. In contrast, the vertebral column has stable dimensions because it is controlled by HOX genes. After a detailed presentation of a baboon head dissection, we present a new method for mapping hyoid bone position relative to the vertebral axis, applied to MRI images. This is compared to a set of radiographs of 3–7.5 year human children. We observe that the hyoid bone is 1 vertebra lower in human infants than in adult baboons. The normalized oral cavity length is shorter, in agreement with prognathism reduction as controlled by non-HOX genes. Using the cervical vertebrae and their axis as a reference allows the conclusion that there is indeed laryngeal descent from baboons to humans and that it is accompanied by compensatory facial shortening.

This preserves the vocal tract length as well as the relationship between the tongue and the oropharyngeal cavity, which is important for swallowing and other feeding gestures.

**Keywords:** exaptation, baboons, vocal tract anatomy, HOX genes, laryngeal descent

## 1. Introduction

### 1.1 Why link speech emergence and primate vocalizations?

The existence of speech as a characteristic of the human species raises a series of questions that, for the most part, have remained open and unanswered for several centuries. What are the anatomical and cognitive prerequisites for vocal communication? When, where, and how did this type of communication arise? By what steps has this evolution taken place? Did gestural communication originate earlier? Or did gestures and vocalizations arise simultaneously?

Researchers have at their disposal human fossils which, though rarely complete, do allow us, to some extent, to trace the anatomical evolution of the head and neck, and thus the architecture of the vocal tract. Obviously, there are no recordings of their sound productions.

Already by the second third of the 19<sup>th</sup> century, Youatt (1835) had trouble understanding why the chimpanzees lacked the power to speak while they were able to shout loudly. We can understand why the anatomy of the vocal organs of chimpanzees has since then aroused great interest (Vrolik, 1841), but what explains that with very similar organs, these primates cannot use them in the same way humans do? More generally, for insights into the evolution of the cerebral cortex and cognition in human ancestors, researchers have long studied the comparative anatomy of the chimpanzee brain (Clark et al., 1936; Falk, 2014; Walker and Fulton, 1936).

Since we share common ancestors with both apes and monkeys we hypothesize that the current vocalizations of these primates provide us with an underexploited window for exploring the nature of speech, and can inform us about the stages of its emergence. Indeed, we assume that the system of speech communication was gradually established over the course of the millions of years of evolution that separate us from our common ancestors. Animal communication has evolved on several levels:

anatomical, cognitive, ethological, all under the constraining influence of the environment.

On the other hand, the other descendants of these common ancestors would not have followed the same evolution. We can therefore assume that their vocalizations have changed little. The vocalizations of present-day monkeys would thus be *relics* (Pisanski et al., 2016) of earlier vocal tract abilities and, we could say metaphorically, *fossil traces* of the communication of our common ancestors.

Monkey and ape vocalizations depend on the sex, status, and age of the analyzed individual. Among primates, baboons produce a repertoire of fourteen vocalizations identified and associated with situations ethologically described (including behavior and communication) (Hall and DeVore, 1965; Zuberbühler, 2012). There are several acoustic analyzes of baboon vocalizations (e.g. Andrew, 1976; Fischer et al., 2002; Owren et al., 1997; Rendall et al., 2005) and more recently it has been shown that they can produce five differentiated vocalizations corresponding to five different ethological situations (Boë et al., 2017).

## 1.2 The exaptation hypothesis

This chapter does not focus on the acoustic analysis of baboons vocalizations but rather on the anatomical aspects that enable and condition the production of these vocalizations, that is to say on the anatomy of the larynx and on the vocal tract and its position with respect to the larynx and to the cervical vertebrae.

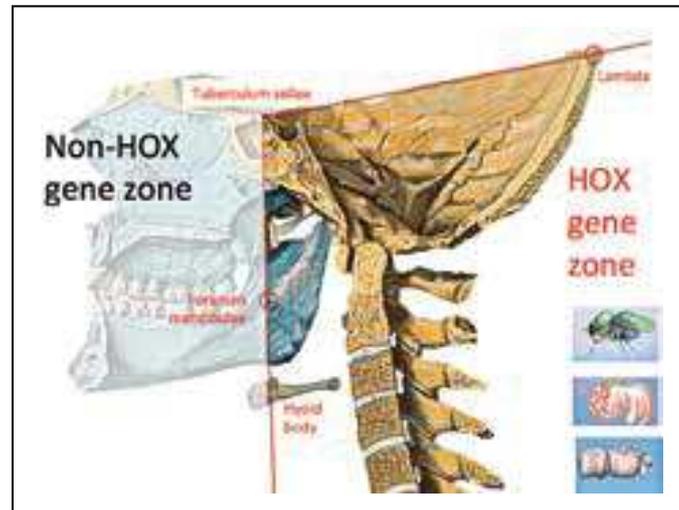
Indeed, the vocalizations of mammals, and thus of human and non-human primates, are all produced by the same process. The sound generated by the vibration of the vocal folds (the source) is acoustically filtered by the resonance characteristics of the vocal tract (the filter), that extends from the glottis (the gap between the vocal folds) to the lips which radiate the filtered signal: this is the source-filter theory (Fant, 1960). By controlling the action of the vocal folds, by modifying the vocal tract shape through control of the articulators (tongue, mandible, lips), or by engaging the nasal passages (through lowering of the velum) it is thus possible for humans to articulate sufficiently differentiated vowels and consonants and to combine them in syllables and syllable sequences.

The question, then, is whether anatomical reasons explain why primates would not be able to produce differentiated vocalizations. The question arises all the more so since for almost 50 years a widespread and longstanding theory (Lieberman et al., 1969; Lieberman, 1975, 1984, 1998, 2007, 2015) has claimed that nonhuman primates, including pre-modern hominids, were incapable of producing systems of vowel-like sounds involving control of their vocal tract, due to their high larynx position and the resulting articulatory anatomy.

The comparative study of the anatomy of the upper aero-digestive tract of *Papio papio* and *Papio anubis* baboons and humans reveals similarities and differences. The first difference is the transition to the upright posture, which caused the centering of the foramen magnum, and triggered reductions in prognathism and the weight of the face. There is therefore a modification of the aerodigestive crossroads at the level of the epiglottis which ends up in a lower position and which is no longer in contact with the soft palate in humans. The second difference is the less flexed skull base, which has the biomechanical consequence of modifying the position of the hyoid bone (Reidenberg and Laitman, 1991). (Note that skull base flexion is measured as the angle of the orbital plane with that of the foramen magnum; the increased flexion in humans indicates a ventral displacement of the foramen magnum to accommodate upright posture.)

Hence, the observed human-baboon similarities are constrained by the primary functions, but the critical point is to consider these *similarities* as the *true prerequisites* to the production of speech. In contrast, many authors assumed that the differences are the markers of limits on the speech production ability. In this context, we will describe the evolution of the laryngeal elements (vocal folds, cartilages, thyro-hyoid membrane), the hyoid bone and the oral elements (tongue, palate). We will use a common reference frame for vertebrates based on the axis of the cervical vertebral column, with the apex of the odontoid as its origin. This landmark is under the control of the HOX development genes, as are the hyo-branchial apparatus and the larynx (Figure 1).

*Figure 1: HOX and non-HOX zones delimited by anatomical landmarks on the skeleton. Insets: top, Drosophila, which is an important model for understanding body generation, and below, mouse and human embryos, which present HOX genes.*



After the first 15 to 20 days following fertilization in vertebrates (Couly et al., 1993; Couly and Bennaceur, 1998; Couly et al., 2002), the HOX genes are responsible for embryo development and determine its anterior-posterior and dorsoventral organization, and thus the placement of the base of the skull, the head, and the body. Consequently, these genes are involved in the growth of relevant bones, which form the framework in which the vocal tract is situated. This system is highly conserved in vertebrates (McGinnis et al., 1984) and it can be assumed that this regulation maintains a suitable morphology for swallowing and protecting the airways.

Conversely, located in non-HOX areas, the oral part of the vocal tract is considered variable. This oral part, derived from the first pharyngeal arch, is under the control of a variety of genes and must be negative HOX for normal development (Chai and Maxson, 2006; Kuratani et al., 1997).

Speech would have evolved through possibilities and constraints external to speech: “speech from nonspeech” (MacNeilage and Davis, 2000a, 2000b), hence the interest in finely analyzing the anatomical structures of the vocal apparatus of non-hominin primates, because they are likely to enlighten us regarding the path followed during the emergence of speech. Thus, gestures of the tongue, the mandible, and the lips were compared across feeding and speech production (Green and Wang, 2003; Hiiemae, 2000; Hiiemae et al., 2002; Hiiemae and Palmer, 2003; Serrurier et al., 2012). Part of the control

might also have been exapted (for discussion, see Ballard et al., 2003; Bunton, 2008; Folkins et al., 1995; Martin, 1991; Ziegler, 2003).

The vocal tract's original and still primary function is digestive. It is divided into two main parts that evolved with their own constraints and their own regulatory genes. The anterior part is dedicated to feeding, with suction and chewing as well as swallowing, and the posterior part is mainly related to swallowing. This chapter revisits the hypothesis of exaptation (Gould and Vrba, 1982) of speech from tongue anatomy as well as from these functions in several ways. First, speech gestures may be derived from feeding gestures. For example, suction and lip rounding are related. Second, they can reuse the existing anatomy. For example, the ability of the tongue for swallowing, which guides food from the anterior to the posterior, is related to its musculature. For speech, this permits constrictions inside the vocal tract at well-controlled positions. Third, the skill at chewing a variety of foods has an impact on the agility of tongue, as well as on the development of oral somatosensory perception and feedback necessary for speech. We now continue with an anatomical description of these anterior and posterior components of the vocal tract, followed by a quantitative analysis of their evolution from baboon to human.

### 1.3 The central position of the hyoid bone

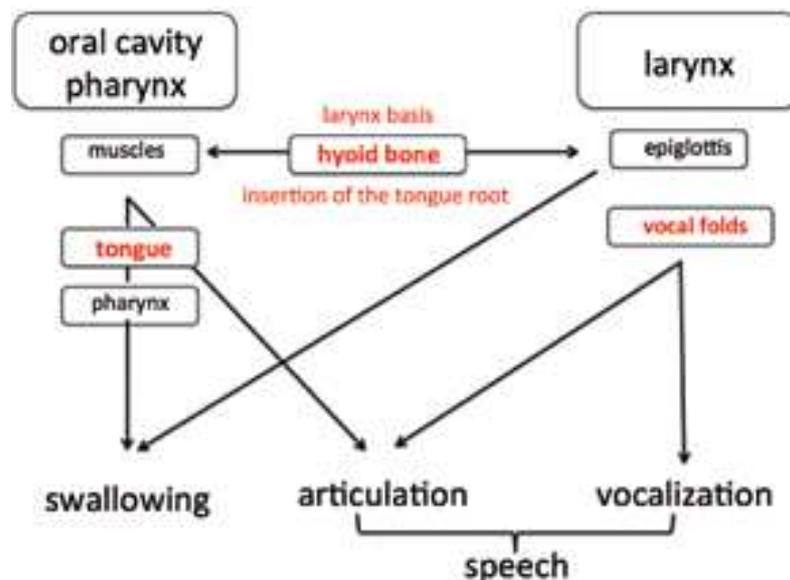
The functional requirements of vocalization involve mobilization of the air source, used for breathing, and of the vocal folds, which protect the airway during swallowing. Unlike chimpanzee or other mammalian larynges (Harrison, 1995; Kelemen, 1969), and contrary to general anatomy (Swindler and Wood, 1973), the baboon hyoid bone and larynx seem, with a few minor exceptions, to have been relatively little studied (Nishimura, 2003ab, 2005, 2006).

The functional imperatives of swallowing and breathing require a close anatomical relation between the oral cavity, the base of the tongue, the pharynx, and the larynx. Presumably, “spatial constraints related to deglutition impose greater restrictions on the rate and degree of hyo-laryngeal descent than do adaptations for vocalization” (Lieberman et al., 2001). The oral and pharyngeal phases in swallowing allow passage of a food bolus from the oral cavity to the esophagus while protecting the airways. The position of these anatomical structures is determined by their insertions on the skeleton, especially on the base

of the skull, the mandible, and the hyoid bone. The positions will be identified relative to the cervical spine (mainly C2, C3, and C4), which has been shown to be highly similar between baboons and humans (Tominaga et al., 1995).

Comparing baboons with humans reveals major morphological differences, in the less flexed base of the skull and in the face, that are associated with the arrangement of the muscle insertions and in the position of the hyoid bone in the baboon. It has been established that prognathism involves differences of the insertions of the muscles of the tongue and supra-hyoid muscles.

*Figure 2: Functional anatomy of the components involved in swallowing, vocalization, and articulation for speech. The two main composite structures, the vocal tract (left) and the larynx (right), are at the top, with their specific anatomical components immediately below, with explicit ties to their functions at the bottom. This shows the great overlap between structures for speech and swallowing. Note that the hyoid bone is a key to this overlap because it is the insertion of the tongue root at the same time as its anatomical relationship with the epiglottis is important for swallowing.*



The hyoid bone plays an important role in the functional anatomy of swallowing, vocalization, and speech production, since it supports the base of the tongue (Figure 2). It is an isolated and fragile bone, of which only a few fossilized specimens have been found: a complete Neanderthal hyoid (Kebara2, 60 kya), a partial Neanderthal hyoid (Asturias, Spain, 43 kya), two hyoids assigned to *Homo heidelbergensis* (Sierra de Atapuerca, Spain, 530 kya), and a complete “chimpanzee-like” hyoid assigned to *Australopithecus afarensis* (Dikika, Ethiopia, 3.3 mya). These discoveries have

renewed the interest in having such fossils for the debate on the origin of speech (Alemseged et al., 2006; Arensburg et al., 1989, 1990; D’Anastasio et al., 2013; Martinez et al., 2008; Rodriguez et al., 2003).

Indeed, the position of the hyoid bone relative to the cervical vertebrae has varied during phylogeny, and it also varies during ontogeny (Lieberman et al., 2001), both in humans and in non-human primates (Nishimura, 2006). The position of the hyoid is an important indicator to consider, perhaps more than that of the larynx itself, because it anchors the tongue root. Its position was considered as a marker of the speech production ability, according to the laryngeal descent hypothesis (Fitch, 2010, p. 312).

## 2. Descriptive anatomy of baboon vocal tract

An accurate comparison of the anatomy of the larynx and of the tongue musculature of baboons and humans is crucial for the discussion of the origin and evolution of speech, considering the crucial role played by these organs in speech. However, such comparative anatomy has been insufficiently described, notably less than for humans vs. chimpanzees (Hofer et al., 1990; Swindler and Wood, 1973; Takemoto, 2008).

The present description was based on two adult *Papio papio* heads, from one male and one female who died naturally in the UPS CNRS Primatology Station, Rousset, France, where various monkeys, including baboons, are kept. The two baboon heads were scanned at the Montpellier CHU in bone fenestration (General Electrics, cut 0.5 mm) when fresh, then sectioned in the strict median sagittal plane when frozen at the anatomy laboratory in Montpellier. Thawing was done in 10% formalin to perform the dissection that was conducted with binocular loupes in both *Papio papio* specimens.

### 2.1 General description of the vocal tract and larynx

The anatomical relations of the pharynx and the larynx are shown on a 3D reconstruction of a male *Papio papio* skull (Figure 3) incorporating a 3D reconstruction of the hyoid bone and the larynx (Figure 4). All 3D reconstructions are performed with Myrian® software. Additionally, a 3D reconstruction of the airways of the female *Papio papio* is superimposed on a sagittal section (Figure 5), and a zoomed portion of the larynx is compared to the median sagittal section (Figure 6). Details are provided in the figure captions.

The anatomy of the oral part of the baboon vocal tract is observed using a medial sagittal section of the female head (Figure 7). It is essentially equivalent to that of the human in its basic elements, but not in its proportions, as the tongue is longer and lower, the hard palate is flatter and the velum (soft palate) more horizontal. In humans, the genioglossal muscle (GG) has three groups of fibers (GGa, GGm, GGp) that form a fan-shaped structure with an anterior GGa vertical (Testut, 1897). In the baboon, the GG has the same structure with GGa composed of vertically oriented fibers and few or no fibers towards the tip of the tongue. The anterior portion of the tongue is proportionally larger, in concordance with the prognathism. The insertion on the superior mental spine is done by means of an aponeurosis on which the muscular fibers are fixed. The geniohyoid muscle, which is not a tongue muscle *per se*, is highly developed and accounts for more than half the tongue height. The styloglossus and digastric muscles were found at dissection to have a more horizontal orientation than those of humans. The hyoglossus presents, as in humans, two components, one inserting on the body of the hyoid bone and the other along the full length of the great horn.

*Figure 3: 3D reconstruction (Myrian® software) of a male Papio papio skull. Left, the skeletal profile shows the entire mandible with a strong ramus but a minimally pronounced coronoid process and mandibular notch. Right, the posterior part of the mandible has been digitally removed after segmentation to show the hyoid bone, the larynx and the trachea. The lowest projection of the hyoid bone occurs above the lower edge of the mandible. The body of the hyoid bone is located forward and above the thyroid cartilage with a portion that descends forward. The hyoid's greater horns follow the upper border of the thyroid cartilage towards the thyroid's upper horns.*

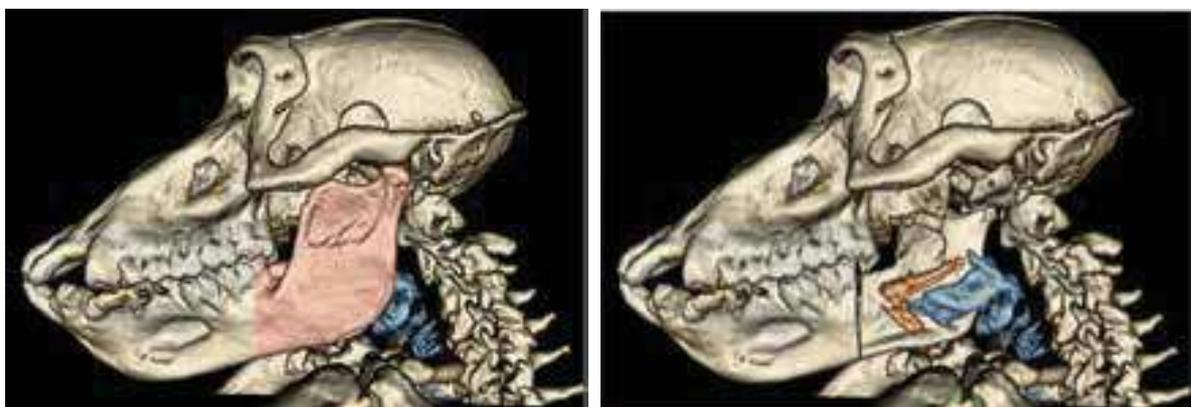


Figure 4: Comparing structure across species. Left panel, 3D reconstruction (Myrian® software) of the hyoid bone and the larynx of the male *Papio papio*. In this ¾ view, we find that morphology of the hyoid bone is grossly similar to that of the human with a median body, two greater horns and two lesser horns. The body of the hyoid bone has an inferior enlargement situated in front of the thyroid cartilage. The view of the larynx reveals a thyroid cartilage that is longer than it is high. It articulates with the cricoid cartilage and one can infer the position of the arytenoid cartilages. Right panel, in comparison to humans (right), the baboon larynx (left) shows that the thyroid cartilage is inset into a relatively larger hyoid bone, with the body enlarged inferiorly and total loss of the thyrohyoid membrane. This disposition is not found in great apes such as chimpanzee and gibbon, but it is found in the stump-tailed macaque, *Macaca arctoides*, and the white fronted capuchin, *Cebus albifrons* (Nishimura, 2003ab, 2005, 2006), shown in the inset (respectively left and right).

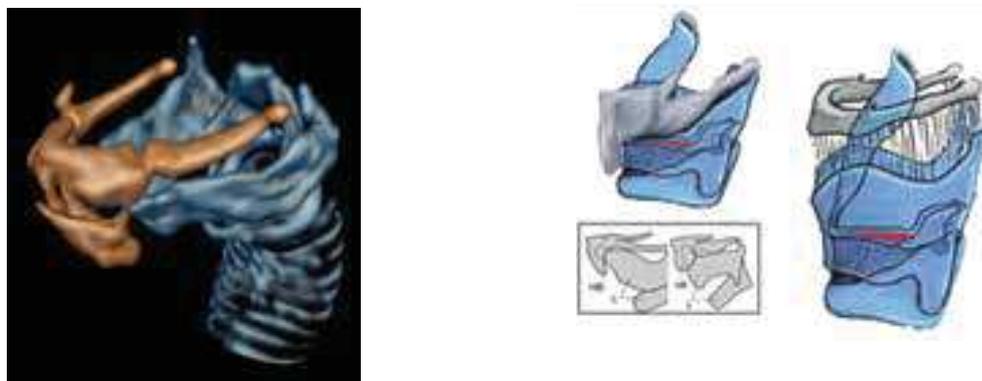


Figure 5: Sagittal section and 3D reconstruction (Myrian® software) of the airways of the female *Papio papio*. The airways were detected semi-automatically. The outlining of the upper airways and of the air in the oral cavity allows close study of their morphology. From behind the hyoid body and in front of the thyroid cartilage, the air sac and its conduit (circled) communicate with the larynx through a passage between the two cartilaginous plates of the thyroid cartilage. We see that the skull base is flattened: the angle shown is greater for all monkeys and apes than it is for humans. The axes of the hard palate and of the cervical vertebrae, both used in this study, are also displayed.

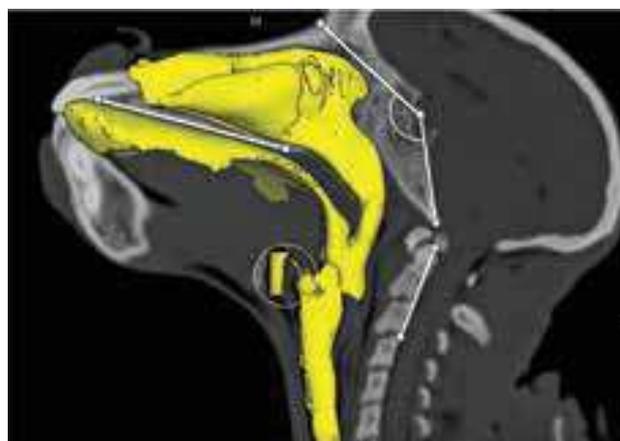


Figure 6: Laryngeal details of *Papio papio*. Left: The anatomical medial sagittal section of the larynx for a female *Papio papio* shows the hyoid bone (Hy) with the body anterior to the thyroid cartilage (Th). The air sac (AS) is between the two. The vocal folds (VF) are positioned at the middle of the thyroid cartilage's vertical dimension. The baboons' vocal folds measure 16.5 mm for the male and 11 mm for the female, in the same range as those of adult humans (Roers et al., 2009). The cricoid cartilage (Cr) is also found with its posterior part relatively high, as is the case for the epiglottis (E). The supraglottal portion of the larynx is very short. Right: Enlargement of the 3D reconstruction (Myrian® software) of the airway of the female *Papio papio* showing the connection from the air sac (AS) to the larynx, with the larynx divided into supraglottic (SG) and infraglottic (IG) portions. Connection is at the level of the glottis, and presents small laryngeal sacs (LS) laterally. The impression of the tongue (To) determines the oro-pharyngeal space that communicates with the piriform recess (PR).

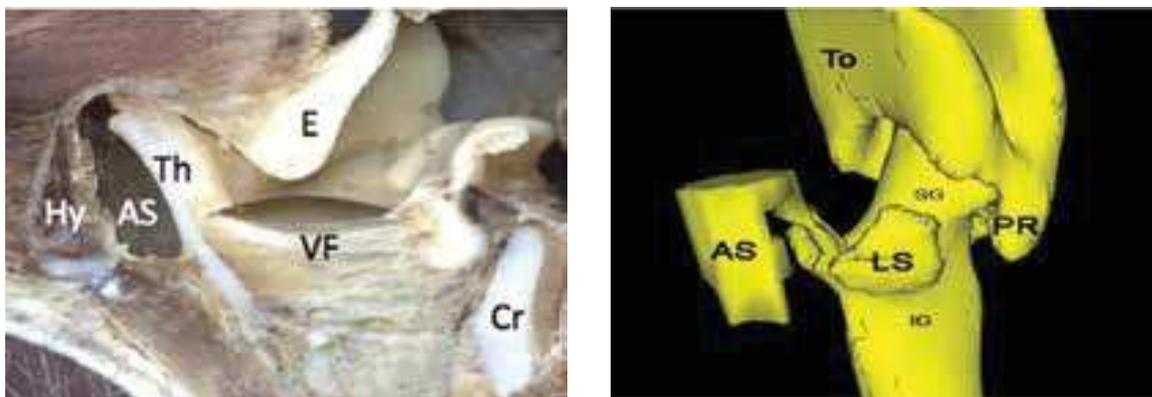
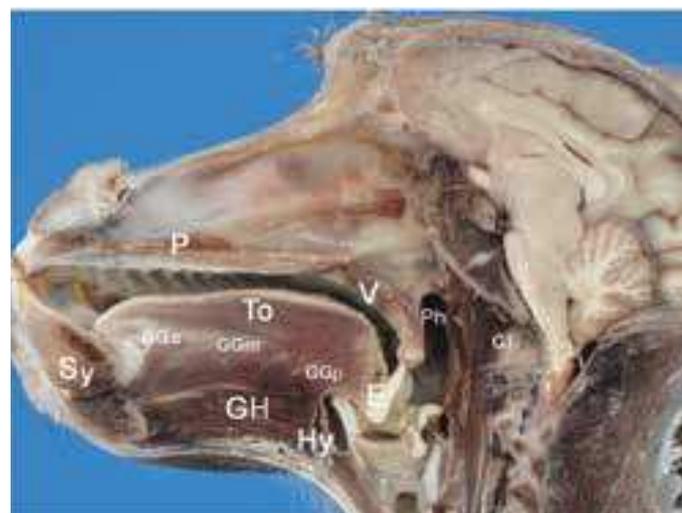


Figure 7: Anatomical medial sagittal section of a female *Papio papio*. The soft palate or velum (V) is at rest and disengaged from the pharyngeal wall (Ph), with the epiglottis (E) in contact with the uvula at the level of the atlas (C1). The anterior (GGa), middle (GGm) and posterior (GGp) parts of the genioglossus (GG) muscle of the tongue (To) are clearly discernible. The geniohyoid muscle (GH) is inserted from the symphysis (Sy) to the hyoid bone. (N.B. the left panel of Figure 6 is enlarged from this figure.)



## 2.2 Tongue musculature and consequences for vocalization

The tongue musculature in baboons was examined by a dissection protocol similar to that used in humans. It appears that tongue musculature is structurally similar in humans and baboons, with the styloglossus and the three parts of the genioglossus, although the external shapes differ: the baboon tongue is flatter while the human tongue is rounded. The muscular hydrostat theory of the tongue shape suggests that, as in chimpanzees, the primary actions available to the baboon tongue are protrusion and retraction (Takemoto, 2008). In addition, the extrinsic muscles raise the back of the tongue through the action of styloglossus, while jaw opening lowers the back of the tongue along with the mandible. This confers to the baboon tongue the necessary degrees of freedom of movement required for swallowing (Crompton and German, 1984; Green and Wang, 2003; Hiiemae, 1967, 2000; Hiiemae and Crompton, 1985; Hiiemae and Palmer, 1999; Hiiemae et al., 1995; Martin 1991; Serrurier et al., 2012), which can then be used to articulate distinctive vocalizations combining two axes (Figure 8). Taking into account the length and flat configuration of the hard palate, it is not clear whether the baboon is capable of sounds such as /i/, which in humans require a long apical constriction along the alveopalatal area.

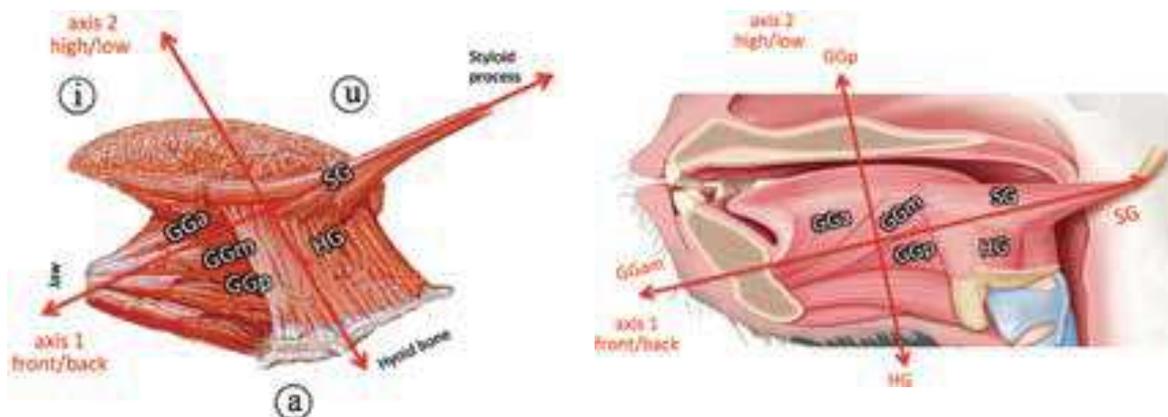
Overall, these considerations are nonetheless compatible with exaptation. Boë et al. (2017) discussed how “[t]he baboon’s muscle fiber orientation allows tongue motion along two main axes.” Antagonistic activation of the GGam, and SG tongue muscles produces changes in the vocal tract allowing both a front/back contrast homologous to the human [æ] ⇔ [u] along the first axis and the posterior constriction needed for [u]. The GGp and HG tongue muscles produce homologs to the human [ɑ] ⇔ [i] contrast through vertical tongue displacement along the second axis. These two axes do have different orientations in baboons and humans, due to the differences in the inclination of the styloglossus and the shape of the tongue (Badin and Serrurier, 2006; Buchaillard et al., 2010; Denny and McGowan, 2012; Honda, 1996).

The laryngeal descent hypothesis asserted that human-like speech production was not possible because of the lack of a posterior cavity allowing a second independent axis:

This lower, pharyngeal portion of the vocal tract provides a whole new dimension: by moving the tongue backwards and forwards this lower tube can be *independently* modified (this arrangement is thus dubbed a “two-tube” tract) ... Thus, the descent of the tongue root and larynx provides an additional degree of freedom, a new dimension of control, compared to the capabilities inferred for a normal mammalian tract. (Fitch, 2010, p. 312, italics in the original).

We disagree, and have concluded that the second axis is not related to laryngeal descent or to the increase in the posterior cavity, but to the tongue muscle structure itself.

*Figure 8: Human (left) and baboon (right) muscles and axes producing vowel contrasts. Note that the main axis of the styloglossus muscle (SG) is nearly horizontal in the baboon, with an action resulting in tongue retraction rather than elevation.*



### 3. Quantitative comparative anatomy of baboons and humans

We now proceed to quantitative analysis of anatomical landmarks associated with the hyoid bone, glottis, prosthion, palatal plane, oral and pharyngeal cavities, and the C2, C3, and C4 vertebrae. We are looking for an adapted anatomical description by: (1) measuring the length of the vocal tract; (2) decomposing it into an oral part (oral cavity length, OCL) and a

pharyngeal part (larynx height, LH); (3) locating the glottis and the hyoid relative to the cervical vertebrae; and (4) recapitulating these observations by age and sex for comparison of baboons with human females, males, and children.

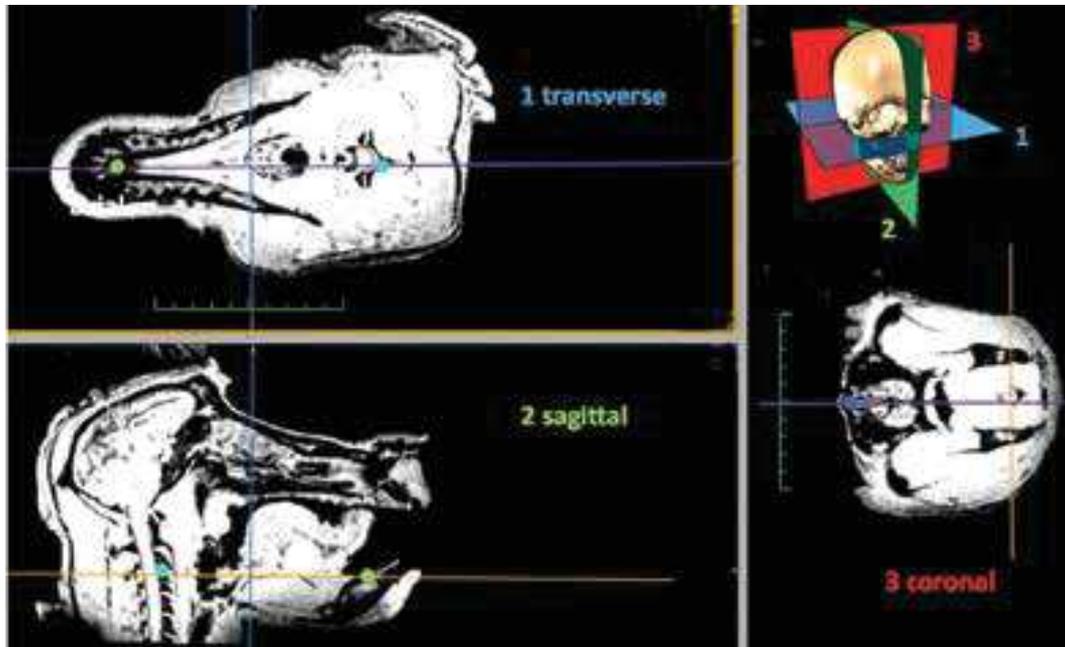
Vocal tract length, measured from the glottis to the lips, depends on the bony structure, soft parts, and their interactions during growth (Scammon, 1930). The vocal tract is composed of an oral part, from the lips to the pharyngeal point (defined below), and a pharyngeal part, from the pharyngeal point to the glottis. In humans, the growth of these parts is heterochronic and there is sexual dimorphism: during puberty, the pharyngeal part develops more than the oral part, especially in males (Goldstein, 1980).

Importantly, in modeling, vocal tract length happens to be a key parameter for the estimation of potential capacities for production of formant resonances in humans through constrictions and cavities (Boë et al., 1989; Bonder, 1983; Liljencrants and Lindblom, 1972) and also for characterizing the acoustic structure of vocalizations in nonhuman primates (Boë et al., 2017; Riede et al., 2005).

### 3.1 Vocal tract biometry of baboons from 3D MRI scans

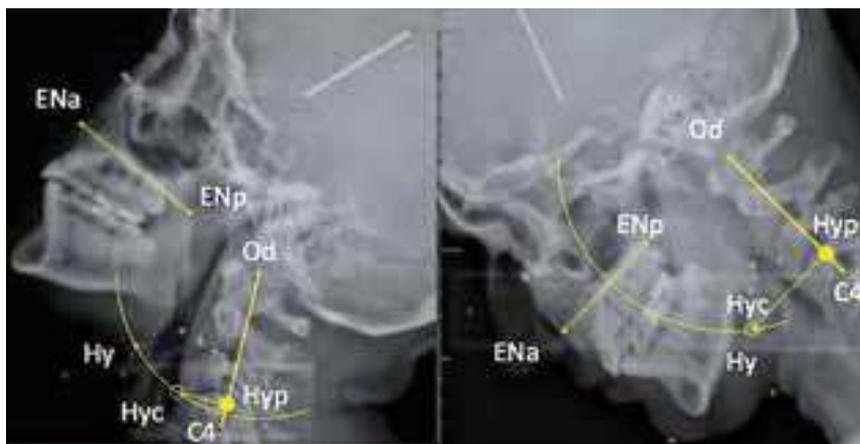
Fifty-six 3D MRI scans of *Papio anubis* baboons, consisting of 23 males and 33 females from 2 years to adulthood (at least 6 years), were performed at La Timone Hospital in Marseille. For the MRI scans the baboons were sedated. From the MRI files, a multiplanar analysis with OsiriX® Lite software (a DICOM viewer) served to locate the median sagittal plane. To do so, we first located median structures in an axial plane: we display in Figure 9, in front, the superior mental spine (EM; note that *spine* is *epine* in French), and in back, the middle of the vertebrae on the axial section. The median sagittal section passes through the mandibular symphysis at the level of the insertion of the genioglossus on the superior mental spine.

*Figure 9: Reference sections of the baboon head. The median sagittal plane must be determined using three points across which there is general bilateral symmetry. The plane thus determined invariably sections the hyoid. The final settings and sections are shown here.*



In the median sagittal plane, we located the following landmarks, using a program written in Matlab® (see Figure 10): the apex of the odontoid (Od), the posteroinferior edge of the vertebral bodies of C2, C3, and C4, the anterior nasal spine (ENa), the posterior nasal spine (ENp), the prosthion (Pr), the upper mental spine (EM), and the top of the upper edge of the hyoid bone (Hy). The following other references are obtained by construction (semi-landmarks): the occlusal plane (which is parallel to the palatal plane ENa-ENp) is evaluated and the pharyngeal point (PPh) is defined as the intersection of the occlusal plane with the pharyngeal wall in the sagittal plane. (Note that the occlusal and palatal planes are presumed perpendicular to the sagittal plane.) The position of the larynx in relation to the cervical spine changes with flexion and extension of the neck (Reidenberg and Laitman, 1991; Westhorpe, 1987).

*Figure 10: Hyoid positioning under head flexion. As explained in the text, a circle was struck at the hyoid position, Hy, with Od at the center. A correction based on the angle of Od-C4 and the palatal plane was then applied to correct for head inclination and find the corrected hyoid position, Hyc. The projected hyoid position, Hyp, was at the orthogonal projection of Hyc on the Od-C4 line. For adult humans Hyp is consistently found at the level of C4, regardless of the angle of head inclination. In the right-hand panel, we note that the angle is close to 90 degrees, so only a minimal correction is necessary from Hy to Hyc.*



In previous primate studies the inclination of their head relative to the cervical vertebrae has not always been well controlled, which has allowed vertical displacement of the hyoid bone according to the inclination of the head. To correct such differences, we adopted a new procedure and validated it first for humans. Working from a profile view, it consisted in locating the palatal plane (using the anterior and posterior nasal spines ENa and ENp) and then in drawing a circle centered at the apex of the odontoid (the estimated center of rotation) and passing through the top edge of the body of the hyoid bone. The angle between the palatal plane and the vertebral column (defined as Od-C4), which is compared to a theoretical angle of  $90^\circ$  (to make comparisons with humans), is then measured. The difference is the correction angle applied to the hyoid bone and after rotation (Hy => Hyc), this point Hyc is then projected orthogonally onto the vertebral column to get the final landmark Hyp (Figure 10). The procedure was tested on different radiographs and partly on MRIs of vowels /i a u/ pronounced by a French speaker. The position of Hyp along the Od-C4 axis was found constant. We do not apply this procedure for finding the position of the glottis, which descends slightly with the inclination of the head, but the larynx is generally much more stable than the hyoid bone.

In order to compare vocal tract and other anatomical dimensions across species, ages, and subjects, we have normalized all distances based on the distance between the apex of the odontoid and C3 inf (dOC3). Thus, we analyze the variations in vocal tract dimensions with respect to cervical vertebrae and their axis taken as a reference. The normalized measures will be expressed as ratios without units, and the position of C3 inf is always taken as 1. This choice is motivated by the consideration that the vertebral axis should be a stable reference in evolutionary terms because of its location in the HOX zone (Figure 1), which is known to have undergone few mutations over a very long time, including the period of mammalian emergence. Conversely the oral part of the vocal tract, located in the non-HOX areas, is considered variable, as well as its pharyngeal and laryngeal parts which are influenced by the hyoid bone position. The normalization operation makes it possible to quantify these variations, under the assumption that the contribution linked to isometric growth of the vocal tract cavities and the cervical vertebrae is thus suppressed, keeping only the relative variations with respect to the vertebral axis. The size differences between subjects are of course eliminated at the same time. This differentiates our study from those of the Japanese chimpanzee and macaque (Nishimura 2003a, 2003b, 2005, 2006; Nishimura et al., 2003, 2008). This approach allows us to evaluate laryngeal descent relative to the odontoid along the vertebral axis, and also indirectly relative to the palate.

Figure 11: Landmarks used to characterize baboon vocal tract geometry (15 year old adult female): Prosthion Pr; Pharyngeal Point PPh (located at the intersection of the occlusal plane and the pharyngeal wall); Glottis Gl; Hyoid Hy, corrected Hyoid Hyc, and projected Hyoid Hyp; Odontoid Od; inferior edges of the cervical vertebrae C2, C3, C4; upper mental spine EM; and anterior and posterior nasal spines ENa and ENp.

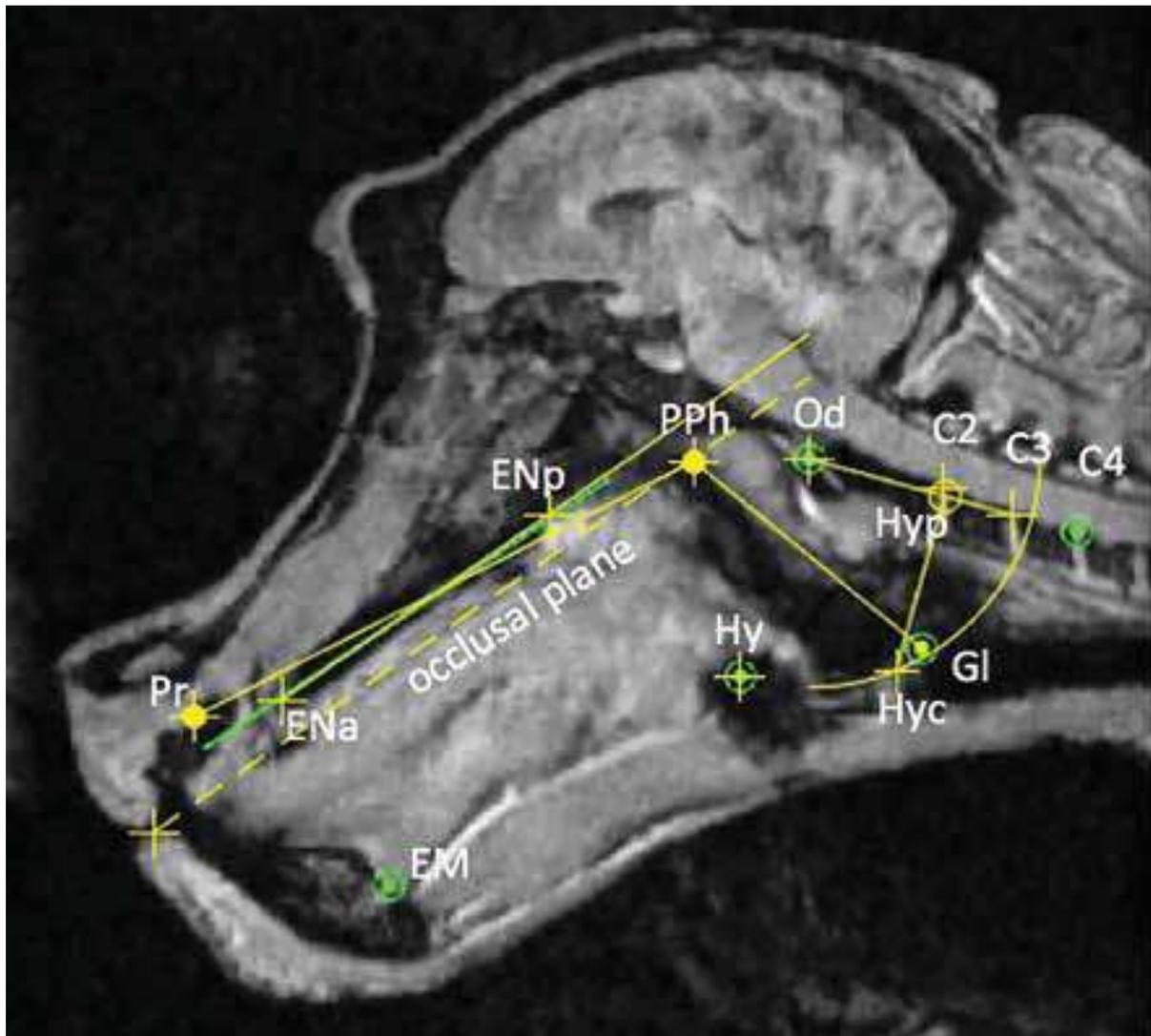
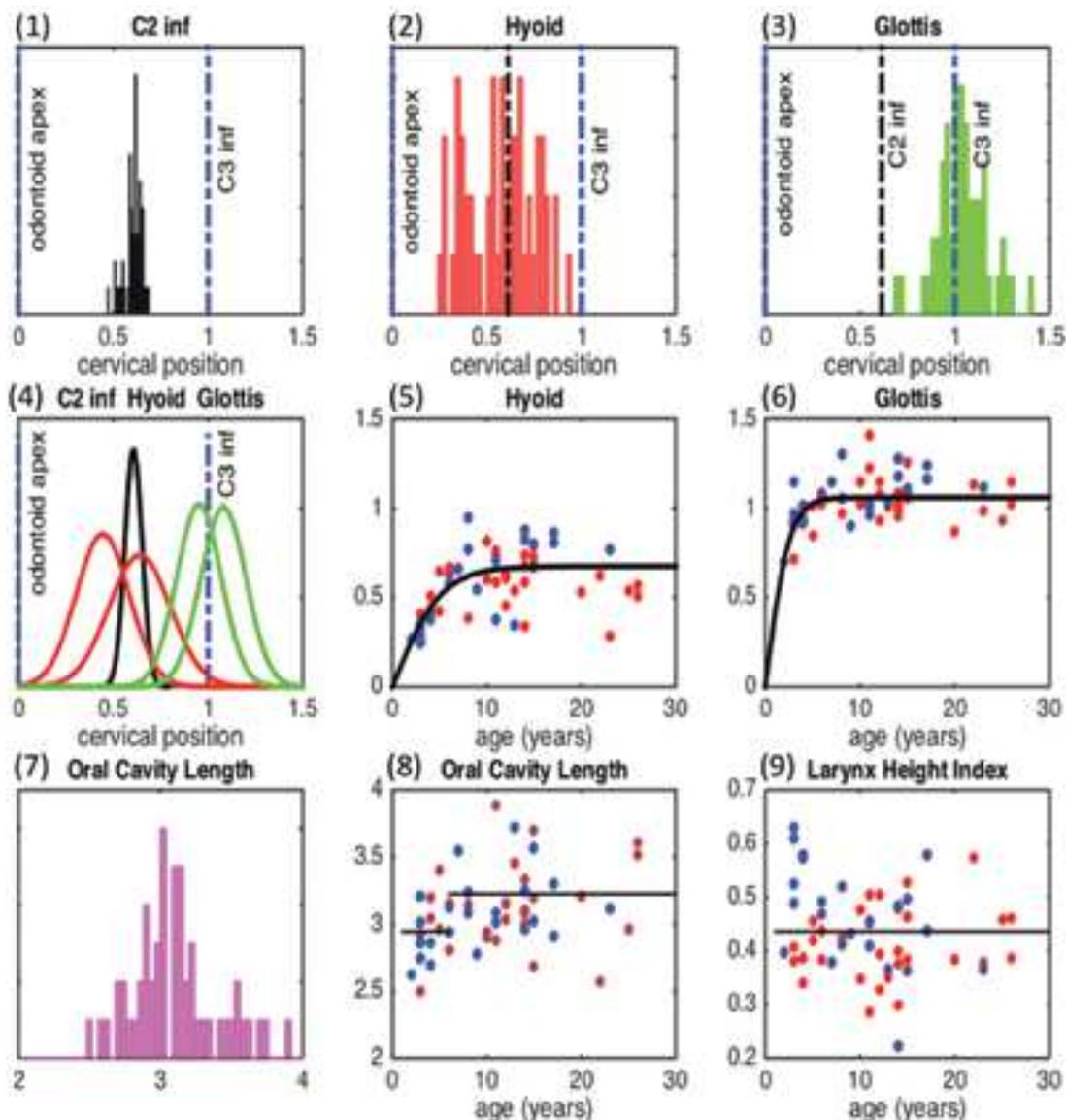


Figure 12: Biometric results for baboons. (See text for discussion of normalized measures in panels 1–4, where, on the abscissa, 0 corresponds to the odontoid and 1 to C3 inf.) (1) C2 inf position. (mean = 0.61); (2) Hyoid bone vertebral projection, Hyp, apparently bimodal; (3) Glottis position, which appears centered on C3 inf; (4) Gaussian modeling of data in panels 1, 2, & 3, with hyoid and glottis data split by age ( $\leq$  and  $>$  6 years); Note that the bimodal appearance of the panel 2 hyoid data is partially related to age, and that age also affects glottis position; (5) and (6) Single-sigmoid fits of projected Hyoid, Hyp, and Glottis, Gl, by age (red dots for females, blue for males); (7) Oral Cavity Length (OCL, the distance between Prosthion, Pr, and Pharyngeal Point, PPh); (8) OCL across age, with a variation represented by two levels, one up to 6 years and the other beyond; (9) LHI (ratio between Larynx Height, LH, defined as the distance between the pharyngeal point and the glottis, and the OCL) across age, which is approximately constant, with an average value of 0.43.



After normalization, the distribution of C2 inf position is narrow (Figure 12.1), providing a precise landmark. The distance between C2 inf and C3 inf (mean = 0.39) is useful for defining a metric expressed in fraction of a vertebra, knowing that one vertebra thus defined includes both the body and the intervertebral space. The distribution of (normalized) vertebral projections of the hyoid is bimodal with the main peak around C2 inf and a second smaller peak somewhat higher on C2 (Figure 12.2). The distribution according to age shows a clear increase which is well fitted with a sigmoid function (Figure 12.5) and the smaller peak apparently represents the young baboons. This is well shown thanks to Gaussian modeling (Figure 12.4) applied after decomposition in two age groups ( $\leq 6$  years and  $> 6$  years): the main peak corresponds to adults (mean = 0.64) and the smaller to young baboons (mean = 0.44), the narrow Gaussian indicating the position of C2 inf at 0.61. Using these means for the 2 age groups and our estimate above of a standard vertebra, we can estimate the amount of hyoid descent as  $(0.64 - 0.44) / 0.39 = 0.51$  vertebra. The glottis histogram is apparently monomodal, but we see some variation by age that we fit with a partial sigmoid function (Figure 12.6). This is also decomposed in two groups with Gaussian modeling (Figure 12.4) having means at 0.97 ( $\leq 6$  years) and 1.1 ( $> 6$  years). This is around C3 inf, and the distance between the hyoid bone projection and the glottis projection is approximately equal to one vertebra for both groups. The glottis descent,  $(1.1 - 0.97) / 0.39 = 0.33$  vertebra, is less than half a vertebra. Note that our determination of the location of the hyoid bone position is more precise than for the glottis, which is, ultimately, an empty space. Moreover, we do not apply any correction to the glottis position. Thus, the hypothesis of having approximately the same descent for hyoid bone and glottis, about half a vertebra, is reasonable. In summary, we quantify the process of laryngeal descent from baboon childhood to adulthood in the following manner: from the middle of C2 to C2 inf for the hyoid, and from the middle of C3 to C3 inf for the glottis, with a constant distance of one vertebra between hyoid and glottis. We have also found that dOC3 does not vary much with age (data not shown).

The oral cavity length (OCL) is normalized similarly in order to see if developmental laryngeal descent in baboons is associated with an increased prognathism. The histogram of OCL is monomodal with a peak

at about 3 (Figure 12.7). In other words, the distance between the prosthion and the pharyngeal point (resp. Pr and PPh in Figure 11) is on average about 3 times the distance between the glottis projection (centered on C3 inf) and the odontoid. The distribution of OCL according to age is divided in two groups (young and adults with limit at 6 years) to show a small increase in prognathism (Figure 12.8). Complementarily, the larynx height index (LHI) (Honda and Tiede, 1998) is defined as the ratio of the OCL and laryngeal height, itself defined as the pharyngeal-glottal distance (Figure 11). It averages a constant 0.43 across age in baboons (Figure 12.9), because it is a ratio between two similarly increasing values. In contrast, LHI varies from 0.5 at birth to 1.0 at adulthood for humans, since there is no increase in OCL while over time there is laryngeal descent.

### 3.2 Vocal tract biometry of human children, from radiography

The position of the hyoid bone and larynx in children has been reported in various studies (Amayeri et al., 2014; Coelho-Ferraz et al., 2006, 2007; Grant, 1965; Westhorpe, 1987). Our own radiographs of children were obtained as part of the SkullSpeech project (Perrier and Boë, 2009–2012). We also obtained radiographic data from M. J. Deshayes (127 children, girls and boys from 3 to 7.5 years, mean 5.2 years, standard deviation 0.95 year), an age range in which there is no sexual dimorphism.

We analyzed the radiographs with the same general procedure as for baboons, adding C4 inf, but eliminating the glottis, which was not visible. We applied the same normalization with dOC3. It appears that the distributions of both C2 inf and C4 inf are narrow. The hyoid projection is between the two, near C3 inf. The mean OCL value is 1.85, and the ratio of OCL between young baboons and infants is  $3/1.85=1.62$  indicating the high degree of prognathism in young baboons. The variations of both hyoid position and OCL seem minimal around 5 years of age. LHI cannot be measured since it requires glottis position, and the glottis was not visible.

Figure 13: Biometric results for human children. (See text for discussion of normalized measures.) (1) Positions of C2 inf and C4 inf, while C3 inf is at 1 by definition; (2) Hyoid bone vertebral projection, Hyp, mean = 1.08 centered on the space below C3 inf; (3) Gaussian models of distributions of the positions of C2 inf and C4 inf and of the hyoid projection, Hyp; (4) Oral Cavity Length (OCL), mean=1.85; (5) Hyoid projection, Hyp, across age, with no significant change; (6) OCL across age, with no significant change.

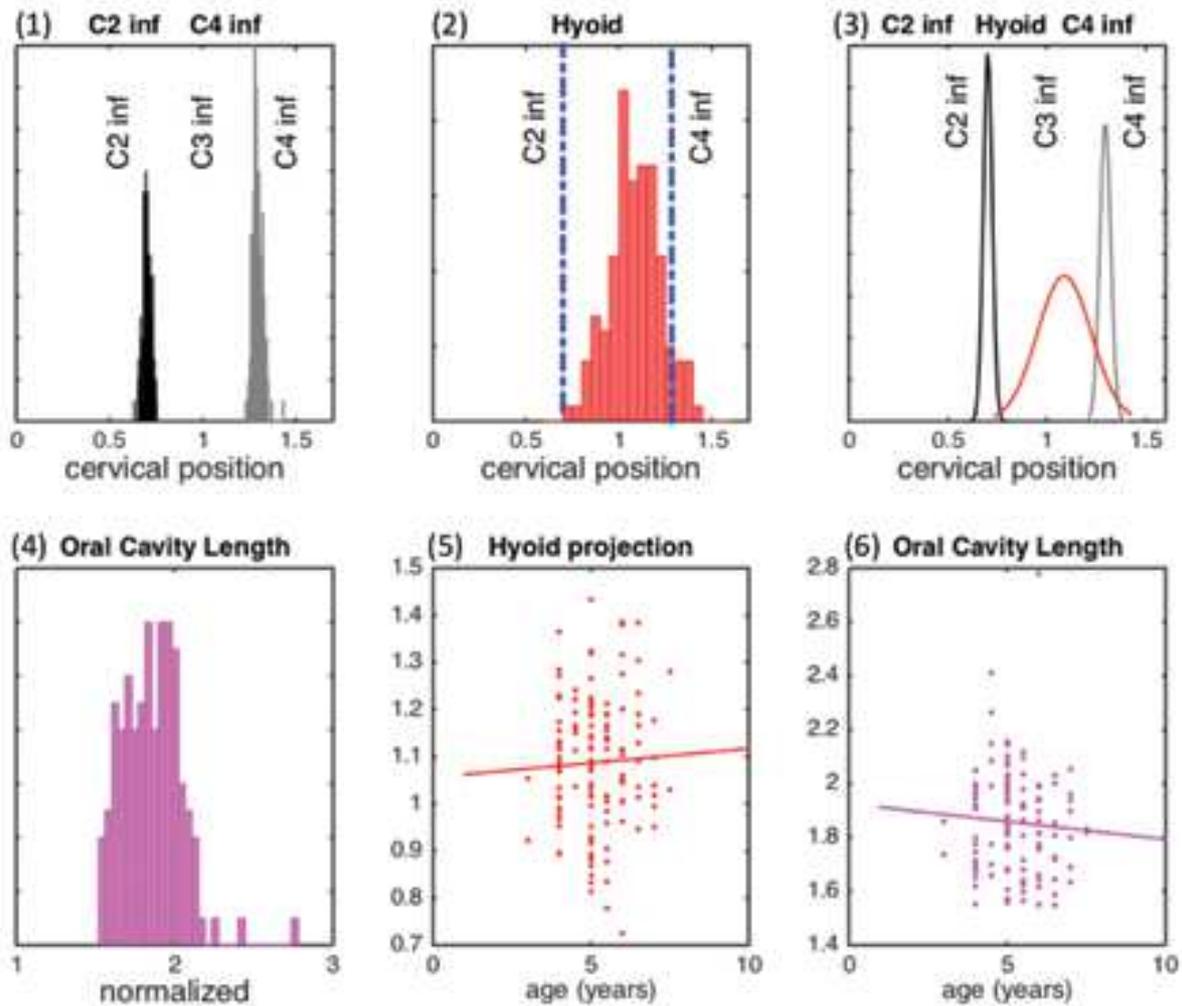
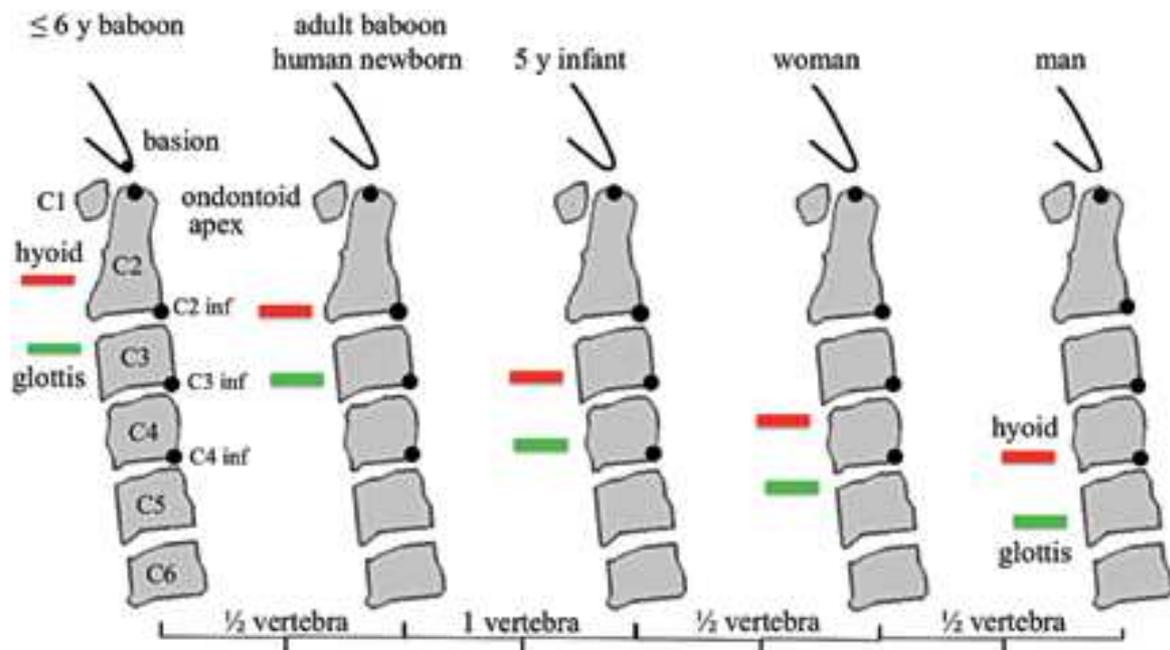


Figure 14: Position of hyoid and glottis compared to cervical vertebrae for young ( $\leq 6$  years) and adult baboons, and in humans, for newborns (Westhorpe, 1987; Barbier, 2010) 5 year-olds (present data; Barbier, 2010), and adult females (Barbier, 2010) and males (Westhorpe, 1987; Barbier, 2010).



Though we detail here only data for human children, our previous studies address growth patterns of the glottis, hyoid, and vertebrae in humans from childhood through adulthood (Barbier, 2010; Barbier et al., 2012, 2015). These studies show that in both males and females, there is considerable vertical growth during the first two years of life, then the growth of C3 and the hyoid position stabilize between 3 and 8 years of age. Glottis descent appears highly correlated with those of C3 and the hyoid. At around 10 years there begins a second surge of vertical growth, which only affects the hyoid and the glottis, C3 being extremely stable after the 8th year. This second surge, greater for the glottis than for the hyoid, seems to stabilize for the hyoid at about 15 years for women, but continues for the glottis until nearly 20 years for men (Barbier, 2010; Barbier et al., 2012, 2015). Figure 14 summarizes all these data in five diagrams comparing young baboons, adult baboons, and human children and adults, female and male. In all these diagrams we observe an offset of about one vertebra between the vertebral projections of the hyoid and the glottis.

We showed in the previous section that the hyoid in the young baboon (less than 6 years old) projects to the level of the body of C2, and for

the adult baboon to the level of C2 inf. Our data show that lowering in baboons takes place in a single step, without any clear sexual dimorphism. For comparison, the hyoid bone projects to the same level in the adult baboon, as in the human newborn (Barbier et al., 2012, 2015). However, the descent of the hyoid in humans takes place in two stages, and we found that in the 5-year-old child, the hyoid bone was around C3 inf (Figure 13). After adolescence, we estimate that the descent takes place down to C4 in females and C4 inf in males. It also appears that in baboons (Figure 12) as well as 5-year-old humans (Figure 13), and indeed in all cases (Figure 14), the glottis is situated about 1 vertebra below the hyoid bone.

It can be suggested that there might be an underlying morphological invariant, namely hyoid – glottis distance, allowing the epiglottis to play its role of protecting the airways while maintaining a constant relationship between its top and the base of the tongue. This hypothesis is reinforced by the fact that these anatomical elements are all regulated by HOX genes. In contrast, the oral anatomy would have escaped from HOX control (Chai and Maxson, 2006), leading to greater changes including a decrease in prognathism associated with a caudal displacement of the tongue and an increase in verticality. The consequence would be a caudal hyoid – glottis translation relative to the spinal column, maintaining the distance between them to enable swallowing and to protect the airways.

### 3.3 Vocal tract growth in humans and baboons

In a previous study (Barbier et al., 2012, 2015) a grouping of four American Association of Orthodontists (AAO) archives was used to quantify (human) vocal tract growth. These records contain 966 sagittal X-rays of the head and neck for 68 white North American subjects (33 women and 35 men), obtained approximately every year between 1 month and 25 years in order to study longitudinal growth of the dentition. For baboons, we were able to retain only 25 of our 56 subjects, the lips being sometimes absent from the MRI.

*Figure 15: Vocal tract length (VTL) measurement from glottis to lips: 10 landmarks positioned by hand and then joined with a spline curve. Baboon *Papio anubis*, male, 15 years, VTL = 13.2 cm.*

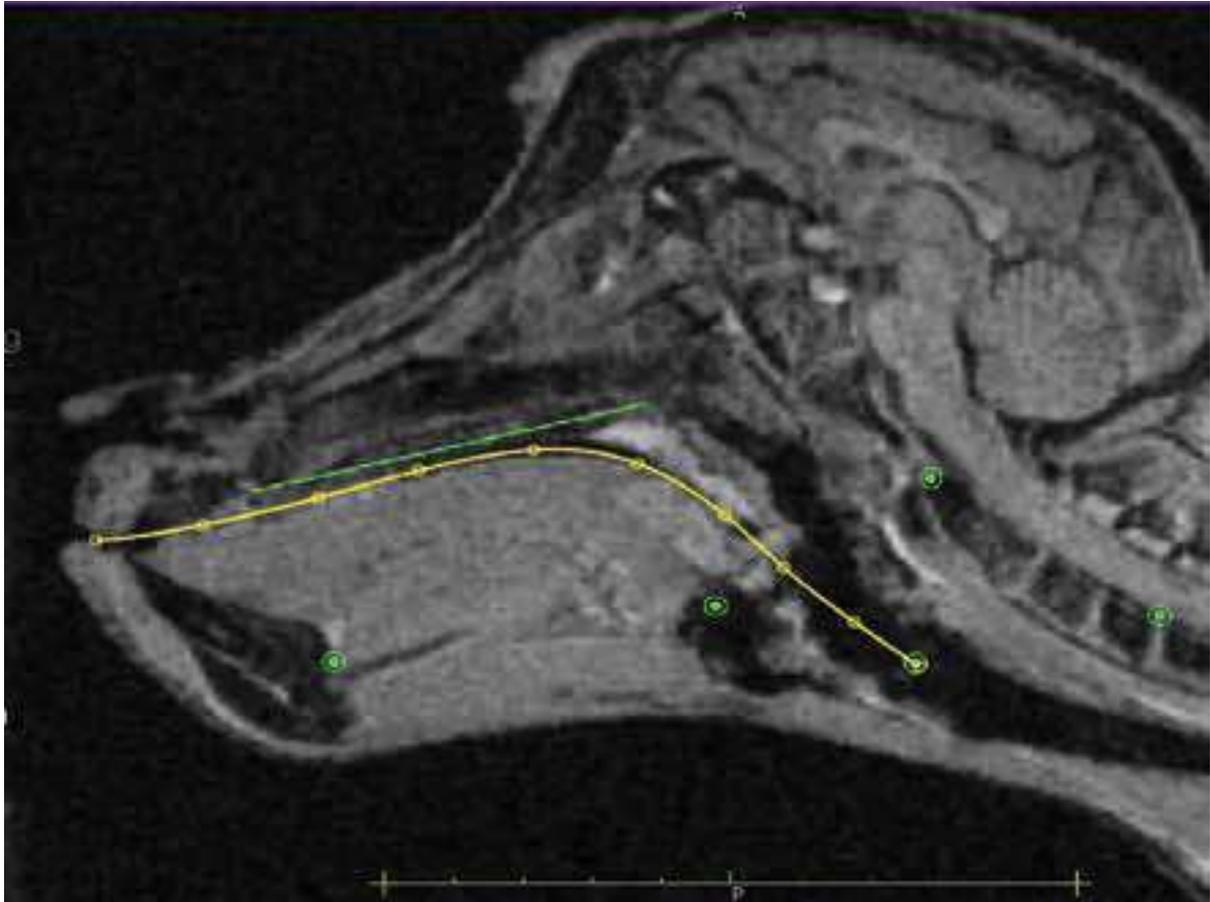
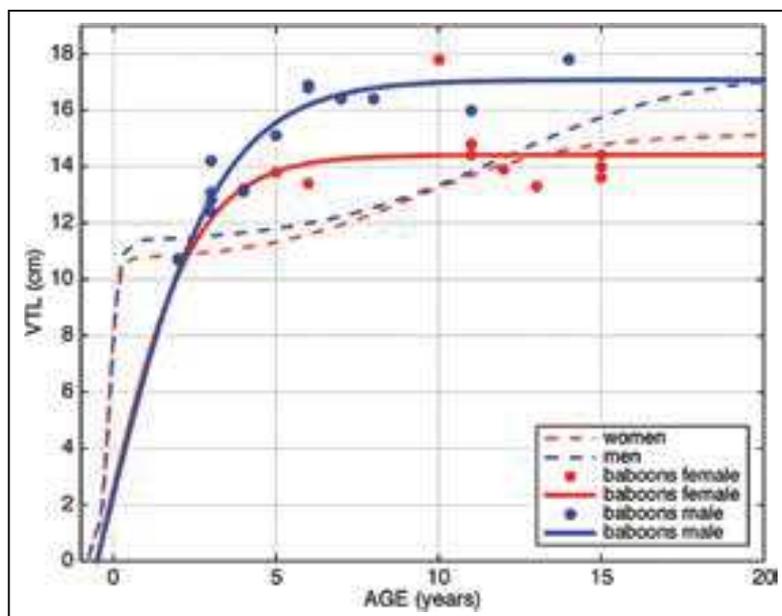


Figure 15 shows the method of VTL measurement used for the 25 baboons. The distance from lips to glottis is the length of a spline curve defined with ten points manually placed on the tongue surface. Figure 16 shows VTL growth obtained by fitting data with a double sigmoid function for human females and males (Barbier et al., 2012, 2015) and, for comparison, measurements modeled by a simple sigmoid for female and male baboons. The choice between simple and double sigmoids stems from the fact that baboons reach their maturity around 6–7 years without an adolescent phase. For adult baboons, as for humans, there is sexual dimorphism, so the growth profiles are not homogeneous and diverge after few years (Figure 16).

Figures 16: Development of vocal tract length from fertilization (assuming 9 months gestation) to adulthood for female and male humans (Barbier et al., 2012, 2015), as fitted with a double sigmoid, and for 12 female and 13 male baboons (assuming 6 months of gestation), as modeled with a simple sigmoid.



Length analysis of the vocal tract's oral segment as normalized to dOC3 shows that there is additional growth linked to prognathism in baboons from youth into adulthood (Figure 12.8, Oral Cavity Length). The fact that the vocal tract grows in both its OCL and LH dimensions results in stability of the Larynx Height Index (Figure 12.9). In addition, we measured VTL (Figure 15) using points manually placed on the back of the tongue and including the two oral and pharyngeal segments, which showed a one-step variation during baboon growth (Figure 16). In contrast, human data show a two-step VTL growth that is consistent with laryngeal descent, which takes place in two stages as well. We conclude that vocal tract growth in the baboon is less than humans in its pharyngeal region, but much greater than humans in the oral region (Goldstein, 1980). This results in a large difference in LHI between baboons and humans: 0.43 for baboons vs. 0.75 for female and 1.0 for male humans. Remarkably, laryngeal descent in humans compensates for the lack of prognathism, with the ultimate effect of preserving vocal tract length.

## 4. Conclusion

In this study, we propose a series of new qualitative and quantitative biometrical analyses of oropharyngeal anatomy adapted and normalized for the comparison between baboons and humans. Using the vertebral column as a fixed phylogenetic reference, we derive a measure of larynx height (the LHI), and also of variations of the oral cavity length expressed with a new metric. We show that the distance of one vertebra separating the hyoid bone and the glottis appears to be invariant, despite the great morphological differences illustrated in Figure 4. A new representation of the laryngeal descent process is summarized on Figure 14. In comparison to the young baboon, we find that the adult baboon has a single-stage laryngeal descent of only  $\frac{1}{2}$  vertebra. The oral cavity then grows an equivalent length, thus keeping LHI constant. The human newborn is at the level of the adult baboon, and humans undergo two descent stages, cumulatively amounting to  $1\frac{1}{2}$  vertebra for females and 2 for males. Since this is realized without the increased oral length from prognathism, it results in a large LHI increase.

Functionally, the distance of 1 vertebra between the hyoid bone and the glottis is highly constrained by the mechanical requirements of swallowing. As baboon tongue musculature is similar to that of humans, if we consider the hyoid bone as the tongue root, laryngeal descent in humans corresponds to a tongue shift toward the back, without modification of the relationship between the tongue and the oropharyngeal cavity because the length of the oral cavity is preserved. We assume this is also constrained by the swallowing function, because the role of the tongue is to drive the food, solid or liquid, from the lips to the esophagus. In baboons, the tongue tip appears to receive no fibers from GGa. This suggests that in humans, laryngeal descent divided the vocal tract in two cavities with a specialization of the anterior part in chewing and preparation of the food for ingestion. The corollary of this specialization was to acquire a better musculature of the tongue tip.

All these observations are compatible with the hypothesis of speech exaptation from feeding gestures. To begin with, tongue musculature in baboons and humans is similar, with the same two control axes enabling constrictions and cavities to form. Swallowing actually involves vocal tract constrictions, and even if the tongue is flat in baboons, the relationship between tongue and oropharyngeal cavity remains similar to that of humans.

Fitch (2010) recapitulates the argument that laryngeal descent is required to form the posterior cavity to produce the vowel /u/, and more generally, the human vowel triangle. While baboons can probably not produce a vocalization close to /i/ because their tongue tip musculature is not developed, a high larynx is not an obstacle preventing vocalizations with formants diverging significantly from the central vowel. More precisely, we consider that laryngeal descent is not required for the development of a posterior cavity. In fact, the stability of vocal tract length and the constant relationship of tongue to oropharynx ensure that the production of an /u/-like vocalization is possible for baboons through contraction of the styloglossus, the same muscle as in humans, allowing a retraction of the tongue and forming a posterior cavity (Figure 8; also see Boë et al., 2017). Constriction control, already present to allow swallowing, is the crucial factor in production of distinct vowel qualities, and this is clearly an example of exaptation from feeding gestures for speech.

## Acknowledgements

Thanks to Mathieu Labrunie who transcoded the files (\*.nii in \*.dcm) of the original MRIs so that we could process them with the OSIRIX software; to Guillaume Barbier for furnishing his data for children growth; to Pierre Badin for providing us with radiographs for the /i a u/ vowels with different inclinations of the head; and to Daniel Lieberman for access to the AAO Denver radiographic database (see Lieberman and McCarthy, 2001).

## References

- Alemseged, Z., Spoor, F., Kimbel, W. H., Bobe, R., Geraads, D, Reed, D., and Wynn, J.G. (2006). A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature*, 443, 296–301.
- Amayeri, M., Saleh, F., and Daleh, M. (2014). The position of hyoid bone in different facial patterns: A lateral cephalometric study. *European Scientific Journal*, 10, 15, 19–34.
- Andrew, R.J. (1976). Use of formants in the grunts of baboons and other nonhuman primates. In S. R. Harnad, H. D. Steklis, and J. Lancaster (Eds.), *Origins and evolution of language and speech* (pp. 673–693). New York: New York Academy of Sciences.

- Arensburg, B., Schepartz, L.A., Tillier, M.A., Vandermeersch, B., and Rak, Y. (1990). A reappraisal of the anatomical basis for speech in middle paleolithic hominids. *American Journal of Physical Anthropology*, 83, 137–146.
- Arensburg, B., Tillier, A.M., Vandermeersch, B., Duday, H., Schepartz, L.A., and Rak, Y. (1989). A middle palaeolithic human hyoid bone. *Nature*, 338, 558–760.
- Badin, P., and Serrurier, A. (2006). Three-dimensional linear modeling of tongue: Articulatory data and models. In *Proceedings of the 7th International Seminar on Speech Production*, 395–402.
- Ballard, K., Robin, D., and Folkins, J. (2003). An integrative model of speech motor control: A response to Ziegler. *Aphasiology*, 17, 37–48.
- Barbier, G. (2010). *Croissance du conduit vocal de la naissance à l'âge adulte. Étude radiographique longitudinale*. Mémoire de Master 1 Recherche. Université Stendhal, Grenoble.
- Barbier, G., Boë, L.J., and Captier, G. (2012). La croissance du conduit vocal du fœtus à l'âge adulte : une étude longitudinale. *Biométrie humaine et Anthropologie*, 1–2, 11–22.
- Barbier, G., Boë, L.J., Captier, G., and Laboissière, R. (2015) Human vocal tract growth: A longitudinal study of the development of various anatomical structures. In *Proceedings 16th Interspeech*, 364–368.
- Boë, L.J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T.R., Becker, Y., Rey, A., and Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS One*, 12, e0169321.
- Boë, L.J., Perrier P., Guérin, B., and Schwartz, J.L. (1989). Maximal vowel space. In *Proceedings Eurospeech*, 2, 281–284.
- Bonder, L. (1983). Equivalency of lossless n-tubes. *Acustica*, 53, 193–200.
- Buchillard, S., Perrier, P., and Payan, Y. (2010). A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning. *The Journal of the Acoustical Society of America*, 126, 2033–2051.
- Bunton, K. (2008). Speech versus nonspeech: different tasks, different neural organization. In *Proceedings Seminar Speech Language*, 29, 267–275.
- Chai, Y, and Maxson, R.E. Jr. (2006). Recent advances in craniofacial morphogenesis. *Developmental Dynamics*, 9, 2353–2375.

- Clark, W. E. L., Cooper D. M., and Zuckerman, S. (1936). The endocranial cast of the chimpanzee. *Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 66, 249–268.
- Coelho–Ferraz, M. J. P., Nouer, D. F., Bérzin, F., de Sousa, M. A., and Romano, F. (2006). Cephalometric appraisal of the hyoid triangle in Brazilian people of Piracicaba’s region. *Brazilian Journal of Oral Science*, 5, 1001–106.
- Coelho–Ferraz, M. J. P., Nouer, D. F., Teixeiras, J. R., and Bérzin, F. (2007). Cephalometric assessment of the hyoid bone position in oral breathing children. *Brazilian Journal of Otorhinolaryngology*, 73, 47–52.
- Couly, G., and Bennaceur, R. (1998). Biologie du développement de la face et du cou. Acquisitions récentes d’embryologie génétique. *Encyclopédie Médicale de Chirurgie et de Stomatologie*, 22–001-A-10, 7p.
- Couly, G., Coltey, P., and Le Douarin, N. (1993). The triple origin of skull in higher vertebrates: A study in quail-chick chimeres. *Development*, 117, 409–429.
- Couly, G., Creuzet, S., Bennaceur, S., Vincent, C., and Le Douarin, N. M. (2002). Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development*, 120, 1061–1073.
- Crompton, H. A. F, and German, R.Z. (1984). Mechanism of intraoral transport in macaques. *American Journal of Anthropology*, 65, 275–282.
- D’Anastasio, R., Wroe, S., Tuniz, C., Mancini, L., Cesana, D. T., Dreossi, D., Ravichandiran, M., Attard, M., Parr, W. C. H., Agur, A., and Capasso, L. (2013). Micro-biomechanics of the Kebara 2 hyoid and its implications for speech in Neanderthals. *PLoS One*, 8, e82261.
- Denny, M., and McGowan, R. S. (2012). Implications of peripheral muscular and anatomical development for the acquisition of lingual control for speech production: A review. *Folia Phoniatica et Logopaedica*, 64, 105–115.
- Falk, D. (2014). Interpreting sulci on hominin endocasts: old hypotheses and new findings. *Frontiers in Human Neuroscience*, 8, 134.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton and Co.
- Fischer, J., Hammerschmidt, K., Cheney, D., and Seyfarth R. (2002). Acoustic features of male baboon loud calls: influences of context, age, and

- individuality. *The Journal of the Acoustical Society of America*, 111, 1465–1474.
- Fitch, W.T. (2010). *The evolution of language*. Cambridge University Press: Cambridge.
- Folkins, J., Moon, J., Luschei, E., Robin, D., Tye-Murray, N., and Moll, K. (1995). What can nonspeech tasks tell us about speech motor disabilities? *Journal of Phonetics*, 23, 139–147.
- Goldstein, U.G. (1980). *An articulatory model for the vocal tract of the growing children*. Thesis of Doctor of Science, MIT, Cambridge, Massachusetts.
- Gould, S.J., and Vrba, E. (1982). Exaptation – a missing term in the science of form. *Paleobiology*, 8, 4–15.
- Grant, L.E. (1965). *A radiographic study of hyoid bone position in Angle's Class I, II, and III malocclusions* [Master's thesis]. University of Kansas City.
- Green, J.R., and Wang, Y.-T. (2003). Tongue-surface movement patterns during speech and swallowing. *The Journal of the Acoustical Society of America*, 113, 5, 2820–2833
- Hall, K. R. L., and DeVore, I. (1965). Baboon social behavior. In I. DeVore (Ed.) *Primate Behavior: Field Studies of Monkeys and Apes* (pp. 53–110). New York: Holt, Rinehart and Winston.
- Harrison, D. H. (1995). *The anatomy and physiology of the mammalian larynx*. New York: Cambridge Univ. Press.
- Hiiemae, K. M. (1967). Masticatory function in mammals. *Journal of Dental Research*, 46, 883–893.
- Hiiemae, K. M. (2000). Feeding in mammals. In K. Schwenk (Ed.), *Feeding in tetrapods* (pp. 411–448). San Diego: Academic Press.
- Hiiemae, K. M., and Crompton, A. W. (1985). Mastication, food transport and swallowing. In M. Hildebrand, D. Bramble, K. Liem, and D. Wake (Eds.), *Functional vertebrate morphology* (pp. 262–290). Cambridge, MA: Belknap Press, Harvard University Press.
- Hiiemae, K. M., and Palmer, J.B. (1999). Food transport and bolus formation during complete feeding sequences on foods of different initial consistency. *Dysphagia*, 14, 31–42.

- Hiiemae, K. M., and Palmer, J. B. (2003). Tongue movements in feeding and speech. *Critical Reviews in Oral Biology and Medicine*, 14, 413–429.
- Hiiemae, K. M., Palmer, J. B., Medicis, S. W., Hegener, J., Jackson, B. S., and Lieberman, D. E. (2002). Hyoid and tongue surface movements in speaking and eating. *Archives of Oral Biology*, 47, 11–27.
- Hiiemae, K. M., Reese, A., and Hayenga, S. (1995). Patterns of tongue and jaw movement: A cinefluorographic study of feeding in the macaque. *Archives of Oral Biology*, 40, 229–246.
- Hofer, H. O., Meinel, W., and Sauer, E. (1990). Comparative anatomic studies of the tongue of *Pan troglodytes* (Blumenbach, 1799) and other primates. I. The chimpanzee tongue. *Gegenbaurs morphologisches Jahrbuch*, 136, 455–492.
- Honda, K. (1996). Organization of tongue articulation for vowels. *Journal of Phonetics*, 24, 39–52.
- Honda, K., and Tiede, M. K. (1998). An MRI study on the relationship between oral cavity shape and larynx position. In *Proceedings 5th International Conference on Spoken Language Processing*, 2, 437–440.
- Kelemen, G. (1969). Anatomy of the larynx and the anatomical basis of vocal performance. In G. Bourne (Ed.), *The Chimpanzee*, 1 (pp. 165–186). Basel: Karger.
- Kuratani, S., Matsuo, A. I., and Aizawa, S. (1997). Developmental patterning and evolution of the mammalian viscerocranium: Genetic insights into comparative morphology. *Developmental Dynamics*, 209, 139–155.
- Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., and Palmer, J. B. (2001). Ontogeny and postnatal hyoid and larynx descent in humans. *Archives of Oral Biology*, 46, 117–128.
- Lieberman, P. (1975). *On the origins of language: An introduction to the evolution of speech*. New York: Macmillan.
- Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.
- Lieberman, P. (1998). *Eve spoke: Human language and human evolution*. New York: W.W. Norton; London: Picador, Macmillan.
- Lieberman, P. (2007). The evolution of human speech: Its anatomical and neural bases. *Current Anthropology*, 48, 39–66.

- Lieberman, P. (2015). Evolution of language. *Journal of Anthropological Sciences*, 94, 1–20.
- Lieberman, P., Klatt, D. H., and Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, 164, 1185–1187.
- Liljencrants, J., and Lindblom, B. (1972). Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language*, 48, 839–862.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioural and Brain Sciences*, 21, 499–546.
- MacNeilage, P. F., and Davis, B. (2000a). Deriving speech from nonspeech: a view from ontogeny. *Phonetica*, 57, 284–296.
- MacNeilage, P. F., and Davis, B. (2000b). On the origin of internal structure of word forms. *Science*, 288, 527–530.
- Martin, R. E. (1991). *A comparison of lingual movement in swallowing and speech production*. PhD dissertation, Madison, University of Wisconsin.
- Martínez, I., Arsuaga, J. L., Quam, R., Carretero, J. M., Gracia, A., and Rodríguez, L. (2008). Human hyoid bones from the middle pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *Journal of Human Evolution*, 54, 118–124.
- McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984). A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell*, 37, 403–408.
- Nishimura, T. (2003a) *Comparative morphology of the hyo-laryngeal complex*. Heineman, London.
- Nishimura, T. (2003b) Comparative morphology of the hyo-laryngeal complex in anthropoids: two steps in the evolution of the descent of the larynx. *Primates*, 44, 41–49.
- Nishimura, T. (2005). Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *American Journal of Physical Anthropology*, 126, 193–204.
- Nishimura, T. (2006). Descent of the larynx in chimpanzees : Mosaic and multiple steps evolution of the foundation for human speech. In T. Matsuzawa, M. Tomonaga, and M. Tanaka (Eds.), *Cognitive development in chimpanzees* (pp. 75–95). Tokyo: Springer-Verlag.

- Nishimura, T., Mikami, A., Suzuki, J., and Matsuzawa (2003). Descent of the larynx in chimpanzee infants. *Proceedings of the National Academy of Sciences*, 100, 6930–6933.
- Nishimura, T., Oishi, T., Mikami, A., Suzuki, J., Matsuda, K., and Takahashi, T. (2008). Development of the supralaryngeal vocal tract in Japanese macaques: Implication for the evolution of the descent of the larynx. *American Journal of Physical Anthropology*, 135, 1882–194.
- Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions. *The Journal of the Acoustical Society of America*, 101, 2951–2963.
- Perrier, P., and Boë, L.J. (2009–2012). *SkullSpeech –Acquisition et contrôle articulatoire de la parole, croissance de l’extrémité céphalique du nouveau-né jusqu’à l’âge adulte*. ANR project. [http://www.gipsa-lab.grenoble-inp.fr/crissp/projets.php?id\\_projet=125](http://www.gipsa-lab.grenoble-inp.fr/crissp/projets.php?id_projet=125).
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J., and Reby, D. (2016). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, 20, 304–318.
- Reidenberg, J. S., and Laitman, J. T. (1991). Effect of basicranial flexion on larynx and hyoid position in rats: An experimental study of skull and soft tissue interactions. *The Anatomical Record*, 230, 557–569.
- Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117, 944–955.
- Riede, T., Bronson, E., Hatzikirou, H., and Zuberbühler, K. (2005). Vocal production mechanisms in a non-human primate: Morphological data and a model. *Journal of Human Evolution*, 48, 85–96.
- Roers, F., Mürbe, D., and Sundberg, J. (2009). Predicted singers’ vocal fold lengths and voice classification: A study of x-ray morphological measures. *Journal of Voice*, 23, 408–413.
- Rodríguez, L., Cabo, L. L., and Egocheaga, J. E. (2003). Breve nota sobre el hioides neandertalense de Sidron (Pilona, Asturias). In M. P. Aluja, A. Malgosa, and R. M. Nogues (Eds.), *Antropología y Diversidad (vol. 1)* (pp. 484–493). Barcelona : Edicions Bellaterra.

- Scammon, R. E. (1930). The measurement of the body in childhood. In J. A. Harris, C. M. Jackson, D. G. Paterson, and R. E. Scammon (Eds.), *The Measurement of Man* (pp. 173–215). Minneapolis, USA: University of Minnesota Press.
- Serrurier, A., Badin, P., Barney, A., Boë, L.J., and Savariaux, C. (2012). Comparative articulatory modelling of the tongue in speech and feeding. *Journal of Phonetics*, 40, 745–763.
- Swindler, D. R., and Wood, C. D. (1973). *An atlas of primate gross anatomy. baboon, chimpanzee and man*. Seattle and London: University of Washington Press. 1982 Malabar R.E. Krieger Publ. Co.
- Takemoto, H. (2008). Morphological analyses and 3D modeling of tongue musculature of the chimpanzee (*Pan troglodytes*). *American Journal of Primatology*, 70, 965–975.
- Testut, L. (1897). *Traité d'anatomie humaine. Anatomie descriptive–histologie–développement*. Paris: Octave Doin.
- Tominaga, T., Dickman, C. A., Sonntag, V. K., and Coons, S. (1995). Comparative anatomy of the baboon and the human cervical spine. *Spine*, 20, 131–137.
- Vrolik, V. (1841). *Recherches d'anatomie comparée sur le Chimpanzé*. Amsterdam : J. Mueller.
- Westhorpe, R. N. (1987). The position of the larynx in children and its relationship to the ease of intubation. *Anaesthesia and Intensive Care*, 15, 384–388.
- Walker, E., and Fulton, J. F. (1936). The external configuration of the cerebral hemispheres of the chimpanzee. *Journal of Anatomy*, LXXI, 105–116.
- Youatt, T. W. (1835–1836). Account of the habits and illness of the late chimpanzee. *The Lancet*, 202–206.
- Ziegler, W. (2003). Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. *Aphasiology*, 17, 3–36.
- Zuberbühler, K. (2012). Primate communication. *Nature Education Knowledge*, 3, 83.