

Skull and vocal tract growth from newborn to adult

Louis-Jean Boë^{1,2}, Jean Granat², Pierre Badin¹, Denis Autesserre¹

David Pochic³, Nassim Zga³, Nathalie Henrich¹, Lucie Ménard⁴

¹Institut de la Communication Parlée, INPG, Université Stendhal, CNRS, Grenoble

²Muséum National Histoire Naturelle, CNRS, Paris

³École Nationale Supérieure d'Électronique, Grenoble

⁴Départ. Linguistique et Didactique des Langues, Univ. du Québec, Montréal, Canada

boe,badin,henrich@icp.inpg.fr, jean@granat.name,
denisaute@numericable.fr, pochicd,zgan@enserg.fr, menard.lucie@uqam.ca

***Abstract.** The objective of this work is twofold. First, a model of the vocal tract is positioned into the bony architecture of the male and female skulls from birth to adulthood. Second, vowel spaces are determined and vowel prototypes, for the cardinal vowels, are synthesized using a simulation of the laryngeal source. Results of this modeling study during ontogeny allow for a better understanding of speech acquisition processes in infants and vocal tract reconstruction of fossils' Hominids. New hypotheses with regard to the emergence of speech can also be proposed.*

1. Introduction

In manuals of speech sciences, the laryngeal source and the vocal tract are described in terms of soft tissues without relating them to the bony architecture: they are essentially composed of muscles, membranes, ligaments, fibrous lamella and adipose tissues. Upper incisors constitute the only visible bony landmark for the upper part of the vocal tract, as are the lower incisors for mandible position. The hyoid bone, visible on X-ray images, is unique in that it is not directly attached to any other bone in the skeleton. The cervical vertebrae are only referred to when it comes to describing the position of the larynx and the vocal folds. But in order to study vocal tract growth from an anatomical point of view and to model this process from birth to adulthood, it is crucial to determine the position of the vocal tract relative to the skull architecture and to the cervical vertebrae which are closely related to vocal tract configuration. Goldstein's thesis (1980) exemplified for the first time the possibilities of predicting vocal tract dimensions from bony landmarks in the skull. Her model incorporated data concerning influence of gender and growth on the vocal tract. Using this database, Maeda proposed a model of vocal tract growth (Boë, Maeda, 1998) which has been systematically tested with real data (Ménard et al., 2004) and has recently been improved. However, Fenart (2003) recently published an anthropometric database allowing a description of the prototypical skull for nine growth stages, from the 5-month fetal stage to adulthood. This database will be used in the modeling experiment presented in this paper. As regards the laryngeal source, we used a database allowing high-quality synthesis. The

goal of this work is to characterize, using an articulatory model, (i) the evolution of the vocal tract from birth to adulthood, positioned into the skull relative to the cervical rachis and (ii) to generate vowel spaces and prototypical configurations for cardinal vowels during growth taking into account modifications of the laryngeal source. Such results regarding ontogeny may contribute to vocal tract reconstruction during phylogeny and shed light on the acoustic possibilities during both temporal evolutions.

2. Growth of the bony structure of the vocal tract

We used here Fenart's (2003) data based on measurements made on French skulls to determine the average values of anthropometric landmarks for nine growth stages: 5 months and 7.5 months of fetal life, birth, one year old, 2 years old, 4 years old, 8.5 years old, 14 years old, and adulthood. This dataset consists of 3D coordinates of 87 points for a "hemiskull" (yielding 142 points for the whole skull), including 13 mandible points in vestibular plane based on inner ear labyrinth. In this landmark dataset, Fenart presents a superimposition of the skulls for various ontogenetic stages (Figure 1). It is noticeable that from 4 years of age, the volume of the upper part of the skull is almost similar to that of an adult. On the contrary, the size of the mandible, which plays an important role in the configuration of the oral part of the vocal tract and in the positioning of the larynx, is quite different in the 4-year-old compared to the adult. Among the landmarks presented by Fenart, 21 points were selected. The Frankfort plane was also used as a landmark for the orientation of the vocal tract. (Figure 2). These points allow the identification of the edges of the skull and of the vocal tract. We reconstructed the position of the hyoid bone using data published in recent studies (Granat, Peyre, 2004 ; Boë et al., 2005). Figure 3 depicts the evolution of those 21 points across the 7 growth stages from birth to adulthood. For the evolution of each point, we calculated the amplitude and the age corresponding to 90% of the whole spatial distance from birth to adulthood. Even though the exact values of this calculation depend on the chosen origin point, two general classes can be identified: data points associated to a steep growth curve (class 1, rapid growth) and data points characterized by a shallow growth curve (class 2, slow growth). Note that the points located on the mandible, which have a direct influence on the front-back dimension of the vocal tract and on the position of the hyoid bone, belong to the second class (slow growth). On the contrary, data points related to the upper part of the skull are characterized by a rapid growth curve. Those tendencies are also found with distances determined independently from the landmark. 90% of the total difference in the front-back dimension of the vocal tract from birth to adulthood is reached at 10 years and 7 months. However, 90% of the difference in terms of the vertical dimension of the vocal tract (prosthion-basion distance) is reached only at 18 years old.

A principal component analysis carried out on the coordinates of the points during vocal tract growth. The first factor corresponds to the radial growth of the skull and the second factor corresponds to a rotation of the upper-back region of the skull. Results of the analyses reveal that the first factor accounts for 84% of the total variance. The first two factors account for 96% of the total variance. The effects of those factors are presented in Figure 4. The rotation of the upper-back region of the skull is observed together with a decrease of the value of the sphenoidal angle which is found until the total sphenobasioccipital synostosis. The 14° decrease calculated from Fenart's data

correspond quite well to radiological data. Despite the fact that this value (14°) is lower than the value reported during phylogeny, this rotation permits a relative backward evolution of the face towards the cervix, yielding a backward evolution of the pharyngeal wall. Following this analysis, we developed a growth model.

3. Articulatory model

A version of the Variable Linear Articulatory Model, a growth model we have used in previous studies (Boë, Maeda, 1998), has been modified by the addition of the new parameters. Those parameters allow a modification of the palate height, the corresponding tongue flattening, the pharyngeal wall and the consequences of the inclination angle of the head. Those parameters play an important part in vocal tract reconstruction during ontogeny and phylogeny. The new version of the articulatory model is thus controlled by 14 parameters:

- 3 anatomical parameters: the palate height, the front-back dimension of the oral cavity and the pharyngeal-laryngeal height;
 - 1 positional parameter: the inclination of the head relative to the cervical rachis;
 - 8 articulatory parameters: lip opening and protrusion, the position of the tongue body, the position of the back of the tongue, the position of the tongue tip, the degree of tongue flattening, jaw opening and the position of the larynx;
 - age, which determines the ratio between the front-back dimension of the oral cavity and the pharyngeal-laryngeal height. This parameter also determines F0 value;
- Other parameters related to the glottal source will be added to this set of 14 parameters.

A set of coefficients was applied to the sagittal contour to obtain an estimation of the area function. From this function, an acoustic model was used to calculate the transfer function of the vocal tract, the related formant values and the sound wave. The prototypical configurations have been systematically compared to the data proposed in the literature.

4. Glottal source model

The glottal source can be described using a unified set of five parameters (Figure 5): fundamental frequency (F0), voicing amplitude (A_v), open quotient (Oq), asymmetry coefficient (α_m), and return phase quotient (Qa) (Henrich, 2001). The amplitude parameter A_v can be replaced by a parameter measured from the first derivative of the glottal flow, that is, the amplitude of maximum excitation E. It is also important to take into account the structural noise (jitter, shimmer) and breathiness, that is, the noise made by the air through the glottis in case of incomplete adduction of the vocal folds. Previous work carried out on 3 to 16-year-old children revealed that only the fundamental frequency parameter significantly varied with age, and that no gender difference could be observed before puberty. Jitter does not seem to be dependent of children age or gender ($0.76\% \pm 0.61\%$), nor does the open quotient parameter ($54.8\% \pm 3.3\%$). A similar pattern (regarding gender) can be found in adults. Indeed, differences related to gender and age primarily concerns voicing fundamental frequency. The continuous component of glottal flow does not differ among male and female speakers. Note however that male speakers are reported to have larger values of air flow and faster closing phases than female speakers. These differences result in

higher values of the amplitude parameters A_v and E in male speakers compared to female speakers. In general, the shape of the glottal pulse is more symmetrical in female speakers than in male speakers. Finally, female speakers are generally associated to a more breathy voice quality than male speakers, despite the fact that great between-speaker variability is found. In order to account for the age and gender effects on the fundamental frequency values, we adopted the prototypical variation curves proposed by Beck (1997) for babies. Prototypical data for children, teenagers and adults were found in Lee et al. (1999). As shown in Figure 6, double logistic functions were fitted to these data sets.

5. Key parameters to position the vocal tract into the skull

The total length of the vocal tract (from the glottis to the lips) during vocal tract growth is a key parameter since it determines the limits, on the frequency continuum, of the maximal vowel space. We used double logistics to fit length growth, combining Goldstein's (1980) data and Fitch et Giedd's (1999) data (Figure 7). It is important to note that the lengths of the oral cavity and of the pharyngeal cavity do not follow the same growth curves and vary with gender. This crucial difference is related to the growth of the mandible height and to the displacement of the hyoid bone and larynx during growth. The model has been readjusted such that those length variations are taken into account. For growth stages corresponding to those used in Fenart's study, we positioned the vocal tract generated by the model based on the position of the incisors relative to the prosthion and to the infradental (points 16 and 17 in Figure 2). One modification was done in order to adjust the palate height relative to the posterior nasal spine (point 11). Figure 8 presents the results for a newborn and an adult. Importantly, during growth, the posterior wall of the larynx is moving towards the cervical vertebrae. This is in line with previous anatomical observations made for the baby. Contrary to a widely accepted hypothesis, the baby's vocal tract is not more bent than that of an adult. The value of the angle of the oral cavity relative to the pharyngeal cavity is for the most part a consequence of the position of the head relative to the cervical rachis (newborns do not yet control this position). Note that this configurational difference does not have any acoustic consequences.

6. Results and perspectives

Maximal vowel spaces were generated for the various ontogenetic stages. Despite their linear translation in the frequency space (directly related to the stage-specific vocal tract length), their positions and shapes in acoustico-perceptual spaces (for example in Bark) are basically identical. From birth to adulthood, it is possible for the vocal tract to produce the point vowels [i a u] for which we shall present the corresponding stimuli. Their perceptual identification is very good since F_0 corresponds to the typical value for this growth stage (Ménard et al., 2002). If newborn infants had the same sensorimotor (control) capacities as adults, their vocal tracts would allow them to produce an F_1 - F_2 - F_3 vowel spaces as extensive as that of their parents, they simply need time to acquire and master the relevant control strategies (Ménard et al., 2002; Serkane et al., 2007). The palate height and the size of the oral and pharyngeal cavities are important for the anatomical description of ontogeny. But the geometric characteristics that have the most important consequences on formant values are the lip area and the position and area of

the constriction inside the part of the vocal tract that characterizes the place of articulation. It is important to note that back and front cavities do not necessarily correspond to an anatomical division into oral and pharyngeal parts. In fact, elementary knowledge of the basic acoustics of speech production and vocal tract modeling show that whatever the relationship between the pharyngeal and oral parts, the control of the tongue, jaw, and lips allows one to configure the vocal tract to produce the three vowels [i a u] found in all the world's spoken languages (Boë et al, 2007). This capacity for control, adapted throughout the course of ontogenesis to the dimensions of the speech production organs, permits children, adolescents, and adults of both gender and any age to produce a sound system that maximizes the perceptual distances between their vowels. Reconstructed vocal tracts for Neanderthals (45.000 BP) (Boë et al., 2005; Boë et al., 2007) show the same acoustic potential capacities as modern humans. With the acquisition and the emergence of speech, we are confronted with problems, constraints and limitations that are not fundamentally related to the geometry and the acoustics of the vocal tract, but which refer to the capacities of control and learning that are at the heart of the question of the emergence and structuring of language.

References

- Beck, J. M. Organic variation of the vocal apparatus. In *The handbook of phonetic sciences*, W.J. Hardcastle, J., Laver (eds), 256-297, Blackwell: Oxford, 1997.
- Boë L.J., Heim J.L., Honda K., Maeda S., Badin, P., Abry, C. The vocal tract of primates, newborn humans and Neanderthals: Acoustic capabilities and consequences for the debate on the origin of language. A reply to Philip Lieberman. *Journal of Phonetics*, in press, 2007.
- Boë, L.J., Heim, J.L., Autesserre, D., Badin, P. Prediction of geometrical vocal tract limits from bony landmarks: Modern humans and Neanderthals. In *Speech Production: Models, Phonetic Processes, and Techniques*. Harrington, J., Tabain, M. ed. New York: Psychology Press, 2005.
- Boë, L.-J., Maeda, S. Modélisation de la croissance du conduit vocal. *Journées d'Études Linguistiques, La voyelle dans tous ses états*, Nantes, 98-105, 1998.
- Fenart, R. Crâniographie vestibulaire. Analyse morphométrique positionnelle. *Biométrie Humaine et Anthropologie*, 21, 3-4, 231-284, 2003.
- Fitch, W.T., Giedd, J. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J. Acoust. Soc. of America*, 106(3), 1511-1522, 1999.
- Goldstein, U.G. An articulatory model for the vocal tract of the growing children. Thesis of Doctor of Science, MIT, Cambridge, Massachusetts. <http://theses.mit.edu/>, 1980.
- Granat, J., Peyre, E. La situation du larynx du genre Homo. Données anatomiques, embryologiques et physiologiques. *Biométrie Humaine et Anthropologie* 22 (3-4), 141-163, 2004.
- Henrich N. *Étude de la source glottique en voix parlée et chantée: modélisation et estimation, mesures acoustiques et électroglottographiques, perception*. Ph.D. Thesis, Université Paris 6, Paris, 2001.
- Lee, S., Potamianos, A., Narayanan, S. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.*, 105, 3, 1455-1468, 1999.
- Ménard, L., Schwartz, J.L., Boë, L.J. Role of vocal tract morphology in speech development: Perceptual targets and sensori-motor maps for French synthesized vowels from birth to adulthood. *J. Speech Language and Hearing Research*, 47, 1059-1080, 2004.
- Ménard, L., Schwartz, J.L., Boë, L.J., Kandel, S., Vallée, N. Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *J. Acoust. Soc. Am.*, 111(4), 1992-1905, 2002.
- Serkhane, J.E., Schwartz, J.L., Boë, L.J., B. L. Davis, B.L., Matyear, C.L. Infants' vocalizations analyzed with an articulatory model: A preliminary report. *J. of Phonetics* (accepted), 2007.

Acknowledgments

This research has been partly funded by the French *Centre National de la Recherche Scientifique* with the *Origine de l'Homme du Langage et des Langues* project, and by the European Community within *The Origin of Man, Language and Languages* project.

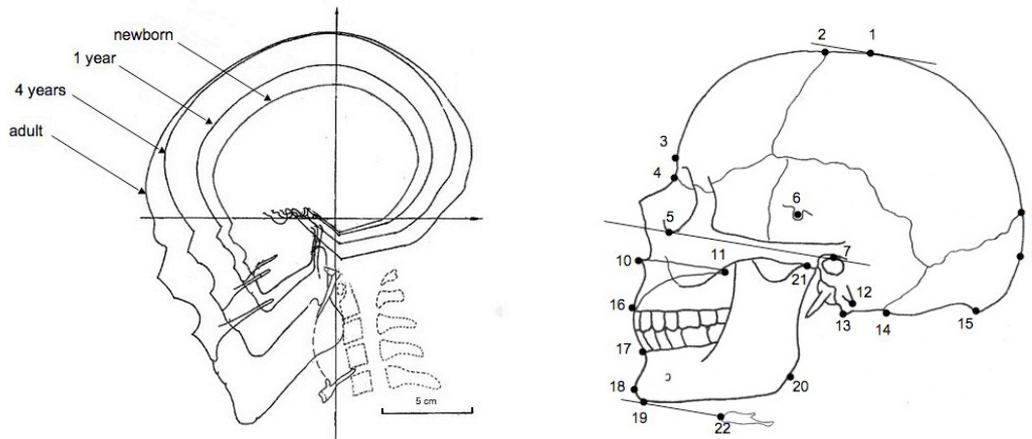


Figure 1. Superimposed sagittal views of the head of a newborn, a one-year-old child, a 4-year-old child and an adult, in the vestibular landmark (following Fenart, 2003).

Figure 2. Landmarks selected among those proposed by Fenart (2003), 1. Vertex, 2. Bregma, 3. Glabella, 4. Nasion, 5. Orbitale, 6. Center of sella turcica, 7. Porion, 8. Lambda, 9. Opisthocranion, 10. Anterior nasal spine, 11. Posterior nasal spine, 12. Basion, 13. Mastoid process, 14. Opisthion, 15. Inion, 16. Prosthion, 17. Infradentale, 18. Pogonion, 19. Menton, 20. Gonion, 21. Upper part of the condyle, 5-7. Frankfort plane and we have added 22 the anterior part of the hyoid bone.

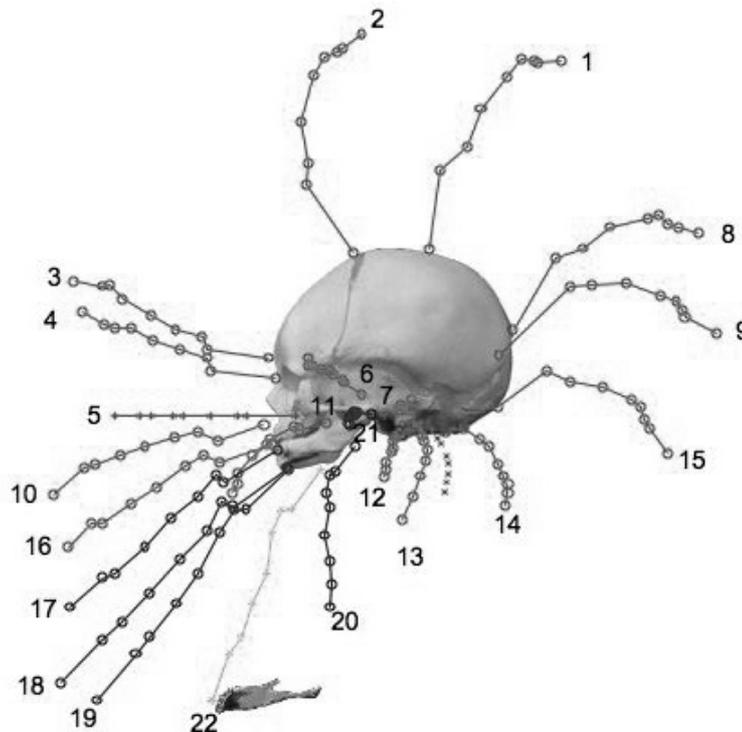
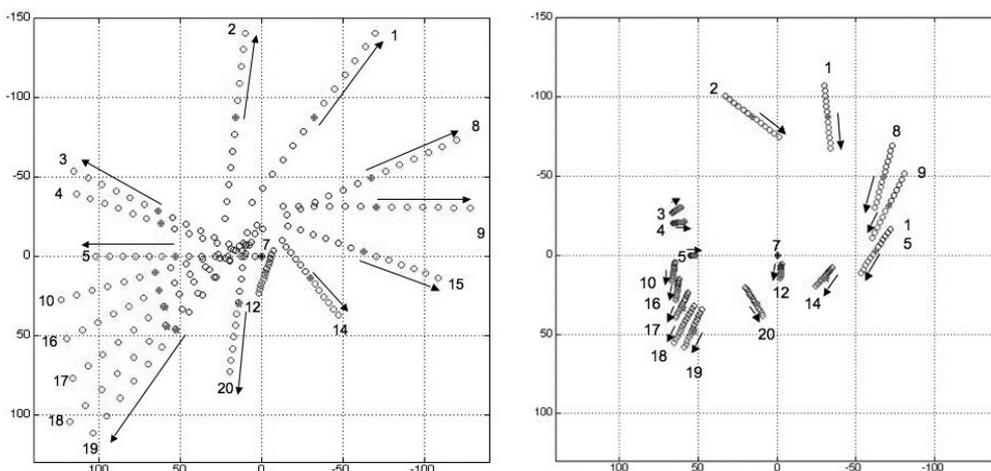


Figure 3. Ontogenesis trajectories of landmarks from birth to adulthood, in the Frankfort plane.



Figures 4. Modeling of the skull growth: variations by 0.5σ step of the mean for the first (left panel) and the second (right panel) parameters, for anthropometric landmarks.

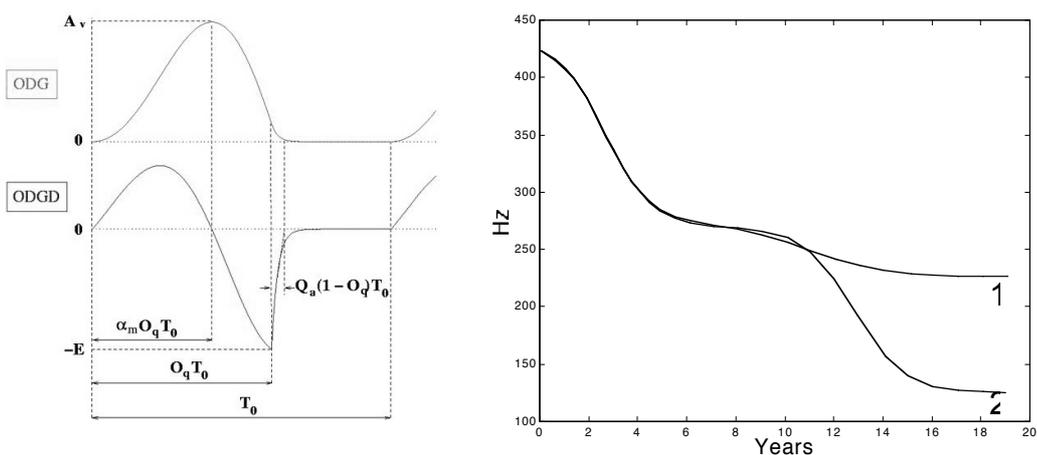
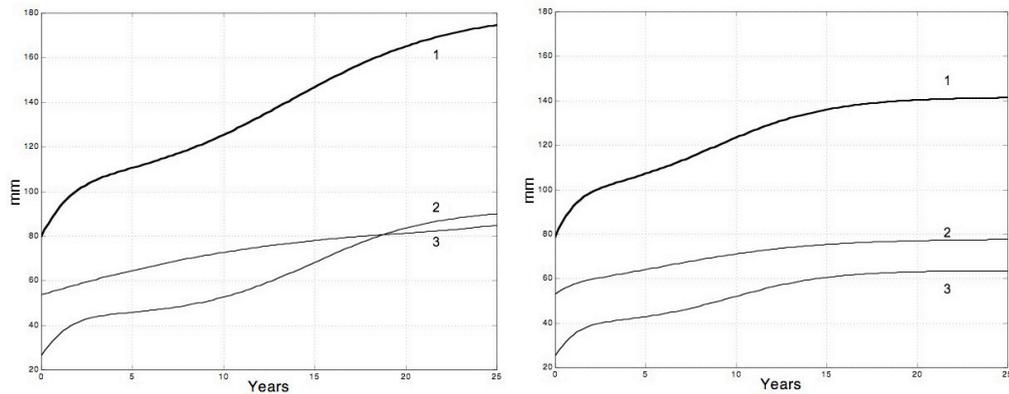


Figure 5. Left: Generic model of the glottal source and control parameters. T_0 , fundamental period; A_v , voicing amplitude; O_q , open quotient ; α_m , asymmetry coefficient; Q_a , return phase quotient; E , amplitude of the maximum excitation (following Henrich, 2001). Right: Prototypical F_0 variations for male and female speakers from birth to adulthood. Double logistic curves have been fitted to those variations.



Figures 7. Vocal tract length (1), pharyngeal cavity length (2), and oral (mouth) cavity length (3) for male (left panel) and female (right panel). It is noticeable that for male speakers, the pharyngeal cavity length reaches and even exceeds that of the oral cavity, which is not the case for female speakers.

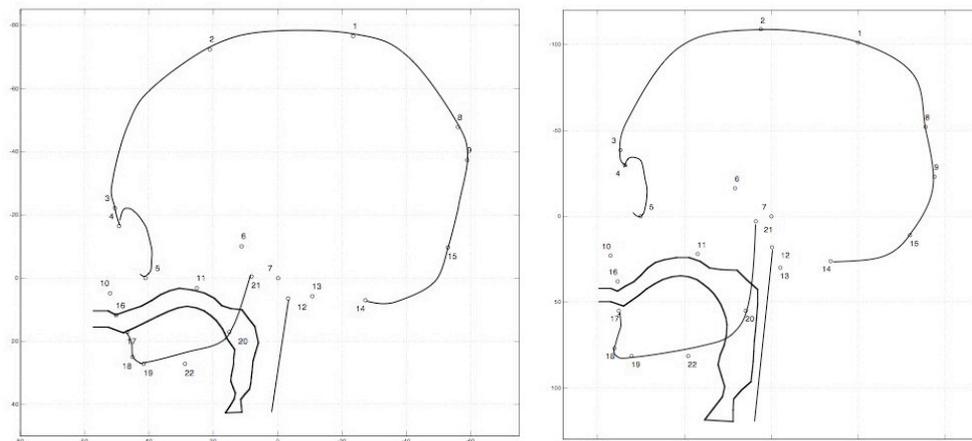


Figure 6. Modeling of the growing vocal tract positioned relative to Fenart's anthropometric landmarks (left panel: newborn, right panel: adult male). Anthropometric landmarks are those reported in Figure 2. The straight line from the basion (point 12) roughly represents the position of the anterior part of the cervical vertebrae and its inclination relative to the skull. Note that the pharyngeal wall gets closer to the vertebrae during childhood.