

## SKULL AND VOCAL TRACT GROWTH: FROM NEWBORN TO ADULT

*Louis-Jean Boë<sup>1,2</sup>, Jean Granat<sup>2</sup>, Pierre Badin<sup>1</sup>, Denis Autesserre<sup>1</sup>  
David Pochic<sup>3</sup>, Nassim Zga<sup>3</sup>, Nathalie Henrich<sup>1</sup>, Lucie Ménard<sup>4</sup>*

<sup>1</sup>ICP – Depart Speech and Cognition, GIPSA, CNRS, Université Stendhal, Grenoble, France

<sup>2</sup>Muséum National Histoire Naturelle, CNRS, Paris, France

<sup>3</sup>École Nationale Supérieure d'Électronique, Grenoble, France

<sup>4</sup>Départ. Linguistique et Didactique des Langues, Univ. du Québec, Montréal, Canada

`louis-jean.boe, pierre.badin, nathalie.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.

**Keywords:** Skull, Vocal Tract, Ontogeny, Vowel Prototypes, Modelling.

### 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. 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. Goldstein's thesis [1] 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

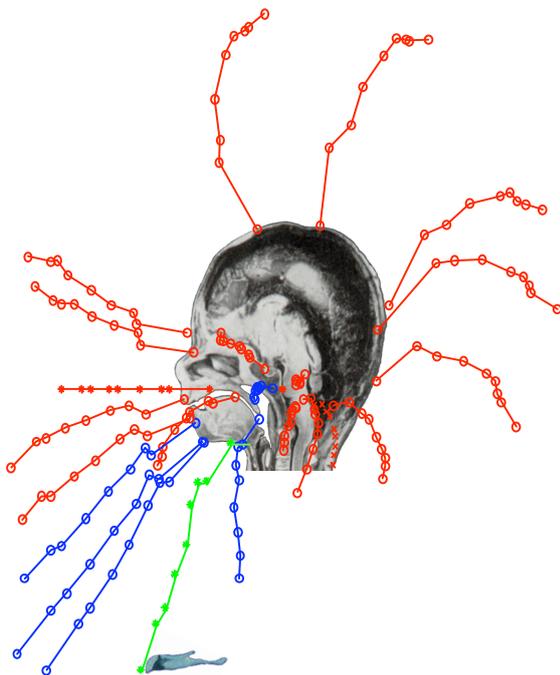
database, a model of vocal tract growth has been proposed [2] which has been systematically tested with real data [3] and has recently been improved. However, Fenart [4] recently published an anthropometric database allowing a description of the prototypical skull for nine growth stages, from the 5-month 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 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", including 13 mandible points. 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 (see Annex). 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 [5]. Figure 1 depicts the evolution of those 21 points across the 9 growth stages from birth to adulthood.

**Figure 1.** Ontogenesis trajectories of bony landmarks from 5 months of fetal life to adulthood, in the Frankfort plane (data from Fenart, 2003; we have added prediction of the hyoid bone).



For the evolution of each point, we calculated the amplitude and the age corresponding to 90% of the distance of each trajectory from birth to adulthood. 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 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 skull growth model.

### 3. ARTICULATORY MODEL

A version of the Variable Linear Articulatory Model, a growth model we have used in previous studies [2], has been modified by the addition of the new parameters. Those parameters allow a 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

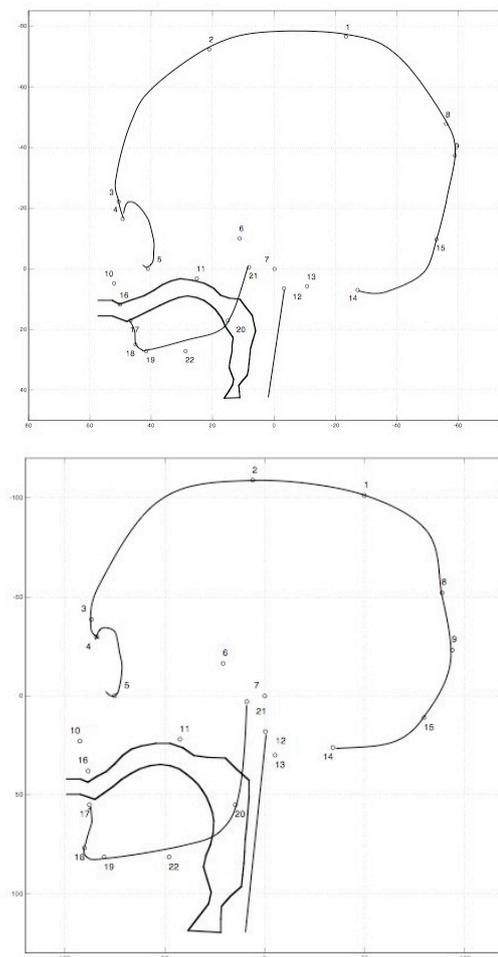
The glottal source can be described using a unified set of five parameters: fundamental frequency ( $F_0$ ), voicing amplitude ( $A_v$ ), open quotient ( $O_q$ ), asymmetry coefficient ( $\alpha_m$ ), and return phase quotient ( $Q_a$ ) (Henrich, 2001) [5]. The amplitude parameter  $A_v$  can be replaced 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 for babies [7] and rototypical data for children, teenagers and adults [8]. 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 published data [1, 8]. 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. Figure 2 presents the results for a newborn and an adult.

**Figure 2.** Modeling of the growing vocal tract positioned relative to Fenart's anthropometric landmarks (upper panel: newborn, lower panel: adult male). The straight line from the basion (point 12) roughly represents the position of the anterior part of the cervical vertebrae and its tilt relative to the skull. Note that the pharyngeal wall gets closer to the vertebrae during childhood.



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). One modification was done in order to adjust the palate height

relative to the posterior nasal spine (point 11). 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 F0 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 F1-F2-F3 vowel spaces as extensive as that of their parents, they simply need time to acquire and master the relevant control strategies [10, 11]. 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 (and presumably of phylogenesis) 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. 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.

**Annex.** 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. Opisthocranium, 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.

## 7. REFERENCES

- [1] Goldstein, U.G. 1980. An articulatory model for the vocal tract of the growing children. Thesis of Doct, MIT, Cambridge, Mass. <http://theses.mit.edu>.
- [2] Boë, L.-J., Maeda, S. 1998. Modélisation de la croissance du conduit vocal. *Journées d'Études Linguistiques, La voyelle dans tous ses états*, Nantes, 98-105.
- [3] Ménard, L., Schwartz, J.L., Boë, L.J. 2004. Role of vocal tract morphology in speech development. Perceptual targets and sensori-motor maps for French synthesized vowels from birth to adulthood. *J.S.L.H.R.*, 47, 1059-1080.
- [4] Fenart, R. 2003. Crâniographie vestibulaire. Analyse morphométrique positionnelle. *Biométrie Humaine et Anthropologie*, 21, 3-4, 231-284.
- [5] Granat, J., Peyre, E. 2004. La situation du larynx du genre Homo. Données anatomiques, embryologiques et physiologiques. *Biométrie Humaine et Anthropologie* 22 (3-4), 141-163.
- [6] Henrich N. 2001. *Étude de la source glottique en voix parlée et chantée: modélisation et estimation, mesures acoustiques et électroglottographiques, perception*. Ph.D. Thesis, Université Paris 6.
- [7] Beck, J. M. 1997. Organic variation of the vocal apparatus. In *The handbook of phonetic sciences*, W.J. Hardcastle, J., Laver (eds), 256-297, Blackwell: Oxford.
- [8] Lee, S., Potamianos, A., Narayanan, S. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *JASA.*, 105, 3, 1455-1468.
- [9] Fitch, W.T., Giedd, J. 1999. Morphology and development of the human vocal tract. *J.A.S.A.*, 106(3), 1511-1522.
- [10] Ménard, L., Schwartz, J.L., Boë, L.J., Kandel, S., Vallée, N. 2002. Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *J.A.S.A.*, 111(4), 1992-1905.
- [11] Serkhane, J.E., Schwartz, J.L., Boë, L.J., B. L. Davis, B.L., Matyear, C.L. 2007. Infants' vocalizations analyzed with an articulatory model: A preliminary report. *J. of Phonetics* (to appear), 2007.

This work was supported by *Origine de l'Homme du Langage et des Langues* program, CNRS, France, and *Origin of Man, Language and Languages* program, EUROCORES, Europe; and by the PPF Univ. Stendhal – MSH: *La propagation des idées scientifiques dans les sciences de la parole et du langage. Anthropologie, linguistique et génétique*.