



Human vocal tract growth: A longitudinal study of the development of various anatomical structures

Guillaume Barbier¹, Louis-Jean Boë¹, Guillaume Captier², Rafael Laboissière^{3,4}

¹Speech and Cognition Department, GIPSA-lab, CNRS & Grenoble University, Grenoble, France

²Anatomy laboratory, Montpellier University, Montpellier, France

³LPNC, CNRS & Grenoble University, Grenoble, France

⁴CRNL, INSERM, CNRS & Lyon 1 University, Lyon, France

guillaume.barbier@gipsa-lab.grenoble-inp.fr

Abstract

The growth of the head and neck and its components, including that of the vocal tract, is not homothetic but appears rather as an anamorphosis. The growth of various structures presents a phenomenon of heterochrony. Another important issue in vocal tract growth is sexual dimorphism. It was first claimed that sexual dimorphism appears at puberty, but a recent study has suggested that some prepubertal differences exist. To study these two phenomena, we used longitudinal radiographic data of sixty-eight typical subjects (966 radiographs, taken from 1 month to 25 years) and twelve fetuses (anatomical sections). In this study, we analyzed the growth curves and growth types of the hard and soft palate, the pharyngeal cavity and the estimated length of the whole vocal tract using non-linear mixed-effect models, in order to take advantage of our unique longitudinal dataset. Results indicate that most of the structures follow a neural/somatic growth type, while the pharyngeal cavity follows a more somatic growth type. As concerns sexual dimorphism, no prepubertal differences were found, suggesting that the sexual dimorphism is likely to begin at puberty. These results have implications for the acoustics of speech production during development and should lead to improvements in vocal tract growth modeling.

Index Terms: human vocal tract, growth, longitudinal study, speech production, acoustics

1. Introduction

Since the first studies on the growth of human head and neck, it has been shown that this development is not homothetic (linear), but appears rather as a true anamorphosis (complex non-linear transformation). Furthermore, the growth of different structures does not follow a unique growth type, highlighting a phenomenon called heterochrony. Since 1930 [1], it has been shown that the structures belonging to the skull follow a rapid growth in the first years of life to achieve the adult size quite early, which was named neural type of growth, while other structures follow a more continuous growth showing two inflections, which was named somatic type of growth.

Few studies have investigated human vocal tract growth [2, 3, 4, 5]. These studies have quantified the growth of different vocal tract structures, and have observed that different structures follow different growth rates and types. But none of these studies was done longitudinally nor included pre-natal data.

The first goal of the present study, based on a longitudinal data set and including fetal data, is to quantify the growth of

various human vocal tract structures and to define their type of growth, in order to explain their development in relation to the ones of the main structures of the head and the neck.

Another goal of this study is to assess sexual dimorphism (anatomical differences between males and females). Sexual dimorphism is an important issue in vocal tract growth since it could explain the differences in speech production between males and females. The age at which sexual dimorphism emerges is therefore important because it indicates the age from which the observed differences between males' and females' vocal productions could be due to anatomical differences, indicating that before the onset of sexual dimorphism, such differences could only be due to differences in the control of the oro-facial articulators.

Besides these two main issues addressed in this study, detailed knowledge of developmental vocal tract anatomy is a prerequisite in understanding how children learn to produce articulatory gestures with their own oro-facial apparatus and in understanding how the acoustic space for speech production is constrained during ontogeny. The overall growth of the vocal tract has important implications on how the acoustic space for speech production is constrained, since formant frequencies are, roughly, inversely proportional to the vocal tract length [6, 7]. Since the detailed vocal tract growth is not known, it is not possible to predict the maximal acoustic spaces of infants and children from those of adults with a simple scale factor [8]. Thus, it is necessary to have precise anatomical data in order to understand the development of the acoustic space for speech production. An interesting step towards the understanding of this relationship is the use of vocal tract growth models [2, 9]. We believe that anatomical data, as presented in this paper, should improve the accuracy of such models, leading to a more detailed understanding of the anatomic-acoustic relationship during ontogeny.

2. Materials and methods

2.1. Subjects

Four X-ray images databases were used in the present study: the Denver, Iowa, Michigan and Oregon Growth Studies, which are available from the American Association of Orthodontists Foundation in the AAOF Craniofacial Growth Legacy Collection (http://www.cril.org/aaof/aaof_home.asp). Measurements obtained from one of these databases (Denver growth study, 28 subjects) were already used for studying the ontogeny of hyoid and larynx de-

10.21437/Interspeech.2015-156

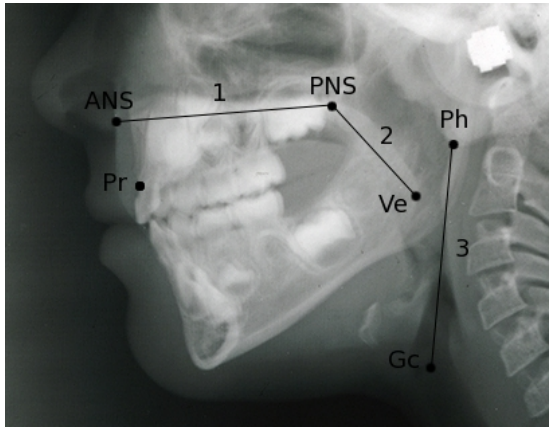


Figure 1: Landmarks and measurements used in the study. The landmarks are shown in dots and the three measurements of interest (1: hard palate, 2: soft palate, and 3: pharyngeal cavity) are shown with lines, superimposed on an example of radiographic image.

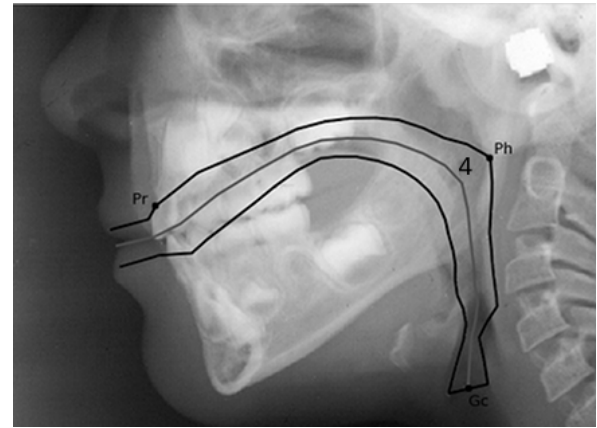


Figure 2: Vocal tract length estimation method. The contour of the articulatory model is superimposed to an example of radiographic image, adjusted to the landmark points Pr, Ph, and Gc. The VTL measurement of interested is indicated by the midline labelled 4.

scent in humans [3].

These sagittal radiographs, which have been digitized, have been taken between 1931 and 1970 from healthy typically developing Americans. The populations covered by these different databases are uniform and can be grouped together.

A total of 68 people, 35 males and 33 females were followed longitudinally from 1 month of age to 25 years of age, and 15 radiographs were taken per individual, on average, over this period. For each radiograph, the age, sex, a code that anonymously identified the subject, and the resolution (in mm/pixel) are included with the data. Altogether, 966 lateral radiographs of the head and neck (496 for males, 470 for females) were used for the present study.

To ensure a continuity of data around birth, we added to this longitudinal database, 12 anatomical sections of formalin-fixed human fetuses, which were taken from the laboratory collection at the Department of Anatomy, Montpellier University [10]. For each anatomical section, radiographs were performed in order to verify that the landmarks located on anatomical sections and on radiographs were comparable. The measurements are made on digitized images of these heads sectioned along the sagittal median plane. The sexes of these fetuses are unknown and their age is given in terms of amenorrhea weeks.

2.2. Landmarks

All measurements were made using CRANIOMAT, a software developed at GIPSA-lab under MATLAB® environment. Each image was calibrated using a scale visible on the radiograph or with the relative pixel/mm factor provided.

Six major landmarks, located manually on each radiograph and on each fetal anatomical section, were used in this study. These are Pr (prosthion, most anterior point of the anterior and inferior septum interalveolar maxillary central incisors), ANS (anterior nasal spine), PNS (posterior nasal spine), Ve (lowest point of the lower end of the soft palate), Ph (pharyngeal point, at the intersection of the occlusal plane and the pharyngeal wall) and Gc (glottis center, midpoint of the orifice of the larynx between the vocal folds along the anterior-posterior axis). The location of these landmarks is illustrated in Fig. 1.

2.3. Measurements of interest

In this paper, we present results for 4 main measurements of interest, which are defined as follow: (1) Hard palate length: Euclidean distance between ANS and PNS, (2) Soft palate length: Euclidean distance between PNS and Ve, (3) Pharyngeal cavity length: Euclidean distance between Ph and Gc, (4) Estimated vocal tract length (hereafter VTL). To give an estimation of the VTL in a rest position, we adapted a geometric model of the vocal tract (the VLAM model [9, 11] to the subject-specific anatomy thanks to three main landmarks: Pr, Ph and Gc. The VTL estimation is given as the curvilinear distance from the glottis to the lips measured as the midline of this subject-specific model. The first 3 measurements are illustrated in Fig. 1, the VTL estimation is illustrated in Fig. 2.

2.4. Growth curves and statistical analyses

In order to compute growth curves for each variable, we used non-linear mixed-effect models, to take advantage of our unique longitudinal data set (under the R environment [12], using the lme4 package [13]). For each variable, we fitted different functions that describe the growth, so we used various non-linear functions, with increasing complexity: simple logistic function, third degree polynomial function, third degree polynomial function with a logistic function, double logistic function and a fifth order polynomial with logistic function. We performed comparisons among those different models using the Akaike Information Criterion (AIC, [14]). For each fitted function, we considered the individual deviations of the parameters by introducing subjects random effects. Using a stepwise method, the optimal set of random effects was selected. Growth type is assessed for each variable thanks to the interpretation of the values obtained for the parameters.

As regards sexual dimorphism, its examination was done through to the comparison of the random terms obtained for males and females subjects.

3. Results

3.1. Overall growth of vocal tract structures

This section presents the quantitative results on vocal tract growth structures. In sum, these results indicate that, on average, the hard palate grows 22 mm from birth to adulthood, the soft palate grows only 12 mm from birth to adulthood, the pharyngeal cavity grows 55 mm, and the vocal tract length at birth is 80 mm and reaches the adult size of 175 mm for males and 155 mm for females. The population mean growth curves are shown in Fig. 3. Around those mean values, inter-individual differences exist, as well as differences between males and females. The analysis of sexual dimorphism is presented in section 3.4.

3.2. Growth curves

Among all functions that have been tested to explain growth, and for all measurements, the best model is the double logistic model, according to the AIC criterion (see section 2.4). The double logistic equation is:

$$f(t) = \frac{L_1}{1 + e^{-a_1(x-x_1)}} + \frac{L_2 - L_1}{1 + e^{-a_2(x-x_2)}}$$

in which t is time (logarithm of age), L_1 is the value of the first plateau, x_1 is the age of the first inflection, and a_1 is the slope at this inflection. The meaning of parameters L_2 , x_2 , and a_2 are equivalent as those for the first inflection.

To account for inter-individual variability, parameters L_1 , x_1 , a_1 , L_2 , x_2 , and a_2 were set as random parameters of the mixed-effects model. This accounts for the fact that there are inter-individual differences in the ages and slopes for the first and the second inflection, as well as for the two plateaus. One of the major advantages of the double logistic model is that the parameters are directly interpretable in terms of growth. By looking at the values of these parameters, we can induce which type of growth underlies each variable of interest. We can also assess sexual dimorphism by looking at the estimated values for the random parts of these parameters for both males and females. The growth curves of each vocal tract structure are presented in Fig. 3.

3.3. Growth types of vocal tract structures

The growth type of each structure was studied by looking at the values of the parameters of the double logistic model. For the hard palate, the growth is typically a somatic growth type, with two inflections, both substantial. For the soft palate, the growth is very rapid around birth and is very slow after 2 years. Its growth type can thus be characterized as being more neural than somatic. For the pharyngeal cavity length, the growth is very protracted and the first inflection is lesser than what has been observed for the other measurements. However, the second inflection is very substantial and happens quite late. Its growth is clearly somatic, with the particular fact that the growth at puberty is substantial. The growth of the VTL is a mixture of neural and somatic growths, because the whole vocal tract length depends on the dimensions of different structures, either having a more neural growth type (like the soft palate, and to a lesser extent, the hard palate) or having a more somatic growth type (like the pharyngeal cavity).

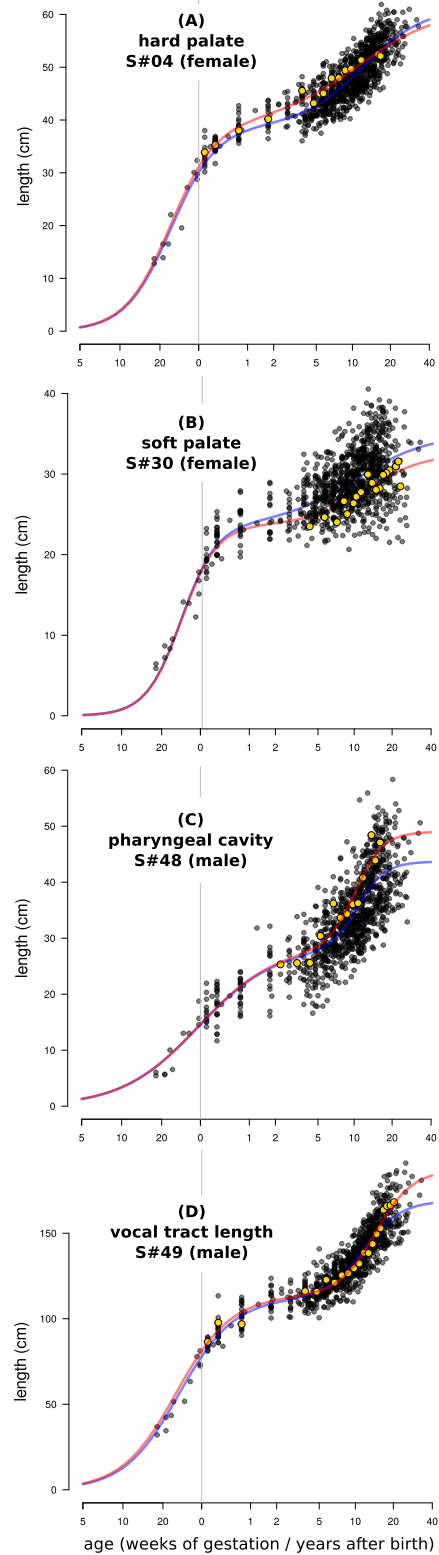


Figure 3: Raw data and fitted growth curves, four each of the four measurements of interest: (A) hard palate, (B) soft palate, (C) pharyngeal cavity, and (D) vocal tract length. The measurements for all subjects at all ages are shown as gray dots. The blue lines represent the population average growth curves. Measurements for individual subjects are shown with yellow dots, with their estimated growth curves shown with red lines. The vertical gray line represents birth.

3.4. Sexual dimorphism

Sexual dimorphism is assessed by comparing the values of the parameters obtained with the double logistic model for males and for females. This analysis reveals that there is no sexual dimorphism on parameters L_1 , x_1 , and a_1 , for none of the four measurements studied here. In contrast, in all four measurements, there are a statistically significant sexual dimorphisms for the parameters of the second inflection L_2 , x_2 , and a_2 (t-tests, for $p < 0.01$). This indicates that the sexual dimorphism appears at puberty. No pre-pubertal sexual dimorphism can be put in evidence from the data, according to our analysis. A transversal analysis, in which we grouped data by one-year age groups, confirmed this finding. According to this transversal analysis, sexual dimorphism appears at 12 years of age for the pharyngeal cavity, and at 15 years for the vocal tract length.

4. Discussion

This study has quantified the growth of various vocal tract structures. Our results were compared to previous anatomical studies on vocal tract growth [2, 3, 4, 5], with which they are highly consistent in terms of measurements. However, this study also includes the analysis of inter-individual differences, thanks to the longitudinal data, and adds an analysis of sexual dimorphism that takes into account these differences.

As regards this particular issue, we found no evidence for a pre-pubertal sexual dimorphism. Even though the differences in vowel production between adult males and females are well documented [15, 16], our results suggest that the observed gender differences in vocal production of young children [17] are likely to be due to differences in the control of oro-facial articulators. Indeed, it has been shown, for example, that for the production of /s/, girls as young as 4–5 years old have higher mean spectral frequencies than boys at the same age [18]. This is not the case for /j/, for example. This indicates that this difference in production is due to a difference in the articulatory control: If it would be due to a difference in anatomy, these differences would be reported for a large range of phonemes, which is not the case. These differences are very likely to be due to differences in the control of the tongue positioning, since it has been shown that adult females use a more anterior point of constriction than males while producing /s/ [19]. This particular gender difference is explainable in terms of sociophonetics, but is really difficult to explain if one tried to explain this phenomenon in terms of anatomical differences between 4-years-old boys and girls.

To finish, these data could be used to improve vocal tract growth models, in order to better understand the articulatory-acoustic relationships during ontogeny. Some work has already been done in this direction [20], and we believe that a more detailed model in terms of anatomy could lead to a better understanding of the constraints applied to the vowel acoustic space during ontogeny. It is also a necessary step in understanding how children learn to organize speech gestures in their own vocal tracts.

5. Conclusion

In sum, this study, the first using longitudinal data and fetal data, has quantified and qualified the anatomical changes of various vocal tract structures during ontogeny. It has been shown that different structures showed different growth rates and growth types, due to their belonging to different anatomical structures

(skull, neck, mandible or face). The emergence of sexual dimorphism has been studied, for which it was shown that no pre-pubertal sexual dimorphism emerges from these data. According to these results, we believe that the differences observed in the vocal productions of young children would be due to differences in the control of oro-facial articulators, such that these differences cannot be explained by anatomical differences.

6. Acknowledgements

The authors are grateful to the American Association of Orthodontists and Daniel Lieberman for providing access to the radiographic archives. This research was partly supported by the French ANR project SkullSpeech.

7. References

- [1] R. Scammon, "The measurement of the body in childhood," in *The Measurement of Man*, J. A. Harris, C. M. Jackson, D. G. Patterson, and R. E. Scammon, Eds. Minneapolis: University of Minnesota Press, 1930, p. 173–215.
- [2] U. G. Goldstein, "An articulatory model for the vocal tracts of growing children," Ph.D. thesis, MIT, Cambridge, MA, 1980.
- [3] D. E. Lieberman, R. C. McCarthy, K. M. Hiiemae, and J. B. Palmer, "Ontogeny of postnatal hyoid and larynx descent in humans," *Arch Oral Biol*, vol. 46, no. 2, p. 117–128, 2001.
- [4] W. T. Fitch and J. Giedd, "Morphology and development of the human vocal tract: a study using magnetic resonance imaging," *J Acoust Soc Am*, vol. 106, no. 3 Pt 1, p. 1511–1522, 1999.
- [5] H. K. Vorperian, S. Wang, M. K. Chung, E. M. Schimek, R. B. Durtschi, R. D. Kent, A. J. Ziegert, and L. R. Gentry, "Anatomic development of the oral and pharyngeal portions of the vocal tract: an imaging study," *J Acoust Soc Am*, vol. 125, no. 3, p. 1666–1678, 2009.
- [6] G. Fant, *Acoustic Theory of Speech Production*. The Hague: Mouton, 1960, p. 107–135.
- [7] H. Mol, *Fundamentals of Phonetics II: Acoustical Modes Generating the Formants of the Vowel Phoneme*. The Hague: Mouton, 1970, p. 62–71.
- [8] G. Fant, "A note on vocal tract size factors and non-uniform F-pattern scalings," *Speech Transactions Laboratory Quarterly Progress and Status Report*, Tech. Rep. 4, 22–30, 1966.
- [9] L.-J. Boë and S. Maeda, "Modélisation de la croissance du conduit vocal. espace vocalique des nouveau-nés et des adultes. conséquences pour l'ontogenèse et la phylogenèse," in *Journées d'Études Linguistiques. La Voyelle dans Tous ces États*, Nantes, 1998, p. 98–105.
- [10] G. Captier, J.-M. Faure, M. Bäumler, F. Bonnel, and J.-P. Daurès, "Anatomy and growth of the fetal soft palate: A cadaveric study to improve its ultrasonographic observation," *Cleft Palate–Craniofacial Journal*, vol. 45, no. 4, p. 439–445, 2008.
- [11] S. Maeda, "Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model," in *Speech Production and Speech Modelling*, W. J. Hardcastle and A. Marschal, Eds. Dordrecht: Kluwer Academic Publishers, 1990, p. 131–149.
- [12] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013, ISBN 3-900051-07-0. [Online]. Available: <http://www.R-project.org>
- [13] D. Bates, M. Maechler, and B. Bolker, *lme4: Linear mixed-effects models using Eigen and Eigen*, 2011, r package version 0.999375-41. [Online]. Available: <http://CRAN.R-project.org/package=lme4>
- [14] H. Akaike, "A new look at the statistical model identification," *IEEE T Automat Contr*, vol. 19, no. 6, p. 716–723, Dec 1974.

- [15] K. Johnson, E. Strand, and M. D’Imperio, “Auditory-visual integration of talker gender in vowel perception,” *J Phon*, vol. 27, p. 359–384, 1999.
- [16] K. Johnson, “Speaker normalization in speech perception,” in *The Handbook of Speech Perception*, D. Pisoni and R. Remez, Eds. Oxford: Blackwell Publishers, 2005, p. 363–389.
- [17] T. L. Perry, R. N. Ohde, and D. H. Ashmead, “The acoustic bases for gender identification from children’s voices,” *J Acoust Soc Am*, vol. 109, no. 6, p. 2988–2998, 2001.
- [18] M. Kinsman and F. Li, “The relationship between gender-differentiated productions of /s/ and gender role behaviour in young children,” in *Proceedings of InterSpeech 2013*, 2013, p. 1283–1286.
- [19] S. Fuchs and M. Toda, “Do differences in male versus female /s/ reflect biological or sociophonetic factors?” in *An interdisciplinary guide to turbulent sounds*, S. Fuchs, M. Todaand, and M. Zygis, Eds. Mouton de Gruyter, 2010, p. 281–302.
- [20] L. Ménard, J. Schwartz, L.-J. Boë, and J. Aubin, “Articulatory-acoustic relationships during vocal tract growth for french vowels: Analysis of real data and simulations with an articulatory model,” *J Phon*, vol. 35, p. 1–19, 2007.