PREDICTION OF THE ABILITY OF RECONSTITUTED VOCAL TRACTS OF FOSSILS TO PRODUCE SPEECH

Jean Granat[1], Louis-Jean Boë[1,2], Pierre Badin[2], David Pochic[3]
Jean-Louis Heim[1], Évelyne Peyre[1], Roland Benoît[4]

[1]Muséum National Histoire Naturelle, CNRS, Paris, France
[2]ICP – Depart Speech and Cognition, GIPSA, CNRS, Université Stendhal, Grenoble, France
[3]ENSERG, Grenoble, France
[4]Université René Descartes, Paris V, France

{louis-jean.boe,pierre.badin}@gipsa-lab.inpg.fr,jean@granat.name,pochicd@enserg.fr
heim@mnhn.fr, rolandbenoit1@free.fr

ABSTRACT

This work is part of a project in a quest of the origin of speech. From classical bony landmarks of the head and jaw used in anthropology, and using a generic model of the vocal tract we attempted to apply the prediction of geometric limits of the vocal tract for modern man to fossils covering a period from 10 ka until 2 Ma (Paleolithic period). We can infer that all the reconstituted vocal tracts could produce the same variety of speech sounds as we can today. But we do not know to what extent they mastered the control skills needed to produce speech.

Keywords: origin of speech, paleoanthropology, vocal tract, paleolithic period.

1. INTRODUCTION

In the present state of knowledge it is not possible to infer when our ancestors acquired the Faculty of Speech and Language in the Narrow sense [1]: control of speech articulators, coordination larynx and vocal tract, phonology, syntax, semantic and recursivity. It is not possible to infer which language or languages were acquired by our ancestors and if the languages spoken today derive from a common source. Among old unsolved questions: Why is our species alone in having speech and language? Why is language the way it is? How and when did language come to be this way? Some of these questions are ill posed problems: we do not have sufficiently data to answer. Perhaps these questions will remain unsolved. But we think that the following question can be solved: If we suppose that our ancestors (and distant cousins) controlled their larynx and vocal tract in the same way as present-day humans, did the geometry of their vocal tract allow them to produce the universal sound structures of the languages spoken today? To fully understand speech and language emergence and evolution it must be approached simultaneously from different disciplines: comparative anatomy between primates and modern men, physical and biological anthropology, embryology and genetics, articulatory and acoustic modeling of speech production and perception, knowledge of universal tendencies of the sound structures. We try in this communication to cover parts of some relevant domains. For the prediction of universal sound structures tendencies we link the Dispersion-Focalization Theory and the Maximum Utilization of the Available Distinctive Features principle in a Perception-for-Action-Control including child speech acquisition [2]. The typologies of phonological systems of the world's languages revealed relatively limited choices among all the syllables and phonemes possibilities. There is a strong bias in favor of systems with CV syllables with 5 vowels and 22 consonants. If a system has 3 vowels the are /i a u/. Systems add other plain vowels maintaining dispersion to ensure clear perceptual distinctiveness. For consonants it is possible to predict the emergence of series of stops (/B D G/), nasals and unvoiced then voiced fricatives. This path observed for speech acquisition sounds plausible for our ancestors.

In this paper we limit our goal. From a skull with mandible and cervical vertebrae, we attempt to: (1) Localize hyoid bone position; (2) Reconstitute a vocal tract model in a plausible way using training data and an articulatory model; (3) Quantify the acoustic capabilities of this reconstituted vocal tract. For this purpose, we combine phylogenesis and ontogenesis: Vocal tract growth from newborn to adult humans can allow us to infer information about the hypothesized evolution of the geometry of the vocal tract of fossil ancestors.

2. DATA

We analyzed 31 skulls from now (for training data) to 3.3 Ma (millions anni) BP (Before Present) for fossil hominids available at the *Musée de l'Homme* in Paris or in the literature: (1) 10-30 ka BP: modern humans: Paleolithic; (2) 90-200 ka BP: anatomically modern humans; (3) 45-90 ka BP: Neanderthals; (4) 1.5 Ma BP: *Homo ergaster*; (5) 3.1-3.3 Ma BP: *Australopithecus africanus, afarensis*. These skulls are all well kept and possess a jaw in the majority of cases. In the majority of case the vertebral column has been reconstituted.

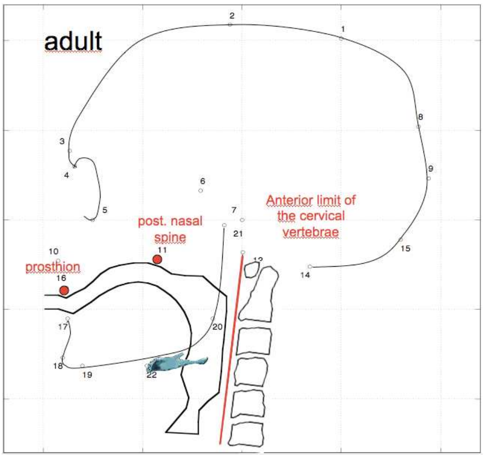
3. ARTICULATORY MODEL

To reconstitute the vocal tract inside the skull we adopted and improved an articulatory model derived from [3] and extensively tested [4]. The parameters are the followings: (1) anatomic : palate height, inclination of the head, ratio between palatal dimension and larynx height, tongue flatness (2) articulatory : lip (protrusion, height), jaw, tongue (apex, body, dorsum), and larynx (3) individual: age (from newborn to adult) and sexual dimorphism (woman and man).
Following [5], we used for our model a larynx height index (LHI) ratio between laryngeal height (LH: distance between the glottis center point 3 and the nasopharyngeal landmark, point 2), the "vertical" part of the vocal tract, roughly the height of the pharynx and palatal dimension (PD: distance between incisors (points 1-2), the "horizontal" part of the vocal tract.
With these parameters we could adjust the dimension and the shape of the vocal tract inside a given skull and taking into account vertebrae architecture. The model fit well with modern skulls (0, 1, 2, 4, 8, 14 years and adults). We verified the alignment of the model with key landmarks of a recent data base [6]: prosthion, posterior nasal spine, anterior line of the cervical vertebrae and hyoid bone (Figure 1).

4. METHOD

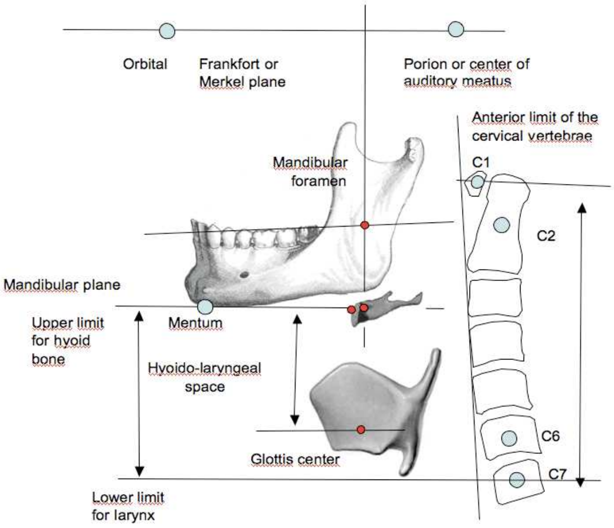
How to derive soft tissues of the vocal tract from bony architecture? Since estimation of the reverse relation, bony landmarks to vocal tract extremities, is a one-to-many problem, constraints must be added to the estimation of hyoid and larynx position. We can define: (1) the anterior limit of the vocal tract which corresponds to the lips. We have no information about the lips of our ancestors but prosthion (or upper incisors) can serve as a landmark; (2) the hard plate (on X-rays) and posterior nasal spine can be observed and used to adjust palate height; (3) the mandibular plane marks the upper limit for hyoido-laryngeal space; (4) the anterior limit of the cervical vertebrae can be considered as the the posterior limit of the pharyngeal wall space; (5) the position of the 7th cervical vertebra which is the lower limit of the hyoido-laryngeal space.

Figure 1. Modeling of the vocal tract positioned relative to Fenart's anthropometric landmarks (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.



In fact within these limits, larynx position moves with speech production (mandible movement, tongue displacement), vocal cord contraction and F0 variation, body position (vertical or supine) and swallowing. Linked by the thyro-hyoidian membrane the vertical position of the hyoid bone and the thyroid cartilage are significantly correlated, and the distance between these two parts have been measured from birth to adult [7]. Different methods have been proposed to derive the position of the hyoid bone (hence the glottis) from the bony architecture [8]. The vertical position of the anterior part of the body of the hyoid bone is located on a parallel to the Frankort (or Merkel) plane drawn from the mentum; and the horizontal position of the posterior part of the body is referred to the mandibular foramen opening on the medial surface of the ramus for transmitting the inferior alveolar artery and nerve. This foramen is important in anatomical and clinical studies, and there are a number of reports about such position (Figure 2). These locations can be derived from anatomic and genetic [9] considerations.

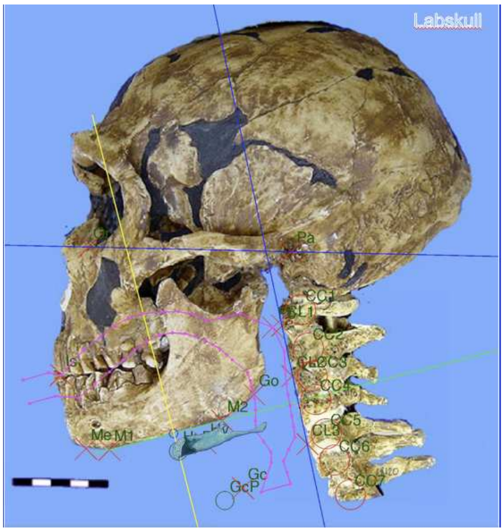
Figure 2. *The method to reconstitute the position of the hyoid bone and of the glottis in relation to the bony structure.*



5. RESULTS

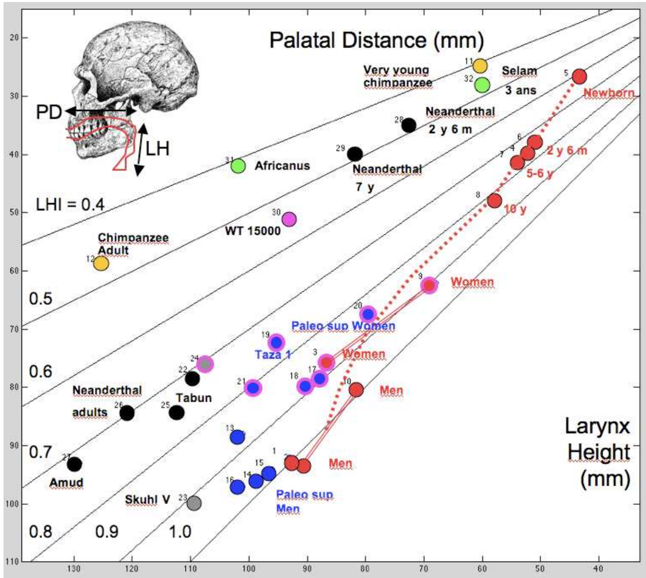
The figure 3 shows a reconstitution of the vocal tract (software *LabSkull*).

Figure 3. Reconstitution of the vocal tract of Neandertal (*La Ferrassie*).



The figure 4 allows to summarize the results with the palatal (horizontal) and laryngeal (vertical) dimensions (PD and LH) for all the skulls; the lines correspond to Larynx Height Index (LHI) varying from 0.4 to 1.0. Ontogenesis from modern men and Neandertal corresponds to an increase of PD and LH and LHI; phylogenesis can be characterized by a decrease of LHI associated to a relative increase of PD.

Figure 4. The position of the reconstituted vocal tract defined by their palatal distance and larynx height; the lines correspond to the different values of the ratio of these two dimensions.



6. ACOUSTIC CONSEQUENCES

We must keep in mind that vocal tract is a 3D volume of which the sagittal 2D contour is crucial to quantifying acoustic capabilities. Among the parameters which have an absence of influence or second-order influences on the acoustic result: (1) palate width: transverse waves are negligible; (2) variability of the palate height: it can be compensated by tongue control; (3) bend of the vocal tract due to the inclination of the head: the length of the vocal tract is only important; (4) the ratio between oral and pharyngeal anatomic cavities: the control of the tongue determines front and back cavities and constriction (or occlusion). Among the factors which have crucial consequences for speech production: (1) tongue position control for shaping the vocal tract; (2) lip area (and protrusion) control ; (3) length of the vocal tract, which shifts the maximal vowel space; (4) range of vibration of the vocal folds for inter-individual (ontogenesis) normalization.

It is important to remind that the articulatory configurations of vowels are due to control of the location and the area of one constriction that divides the vocal tract into 1 or 2 cavities, and to the interlip area. It is well known that the vocal tract can be modeled by four cylindrical tubes, which represented the back cavity, the constriction, the front cavity and the labial horn. For /i/ and /a/, this simplification can even be reduced to two tubes. The vowel /i/ can only be produced by controlling a palatal constriction at the front of the vocal tract. Once this anterior constriction is set, F2 depends only on the remaining length, *i.e*. that of the back cavity of the vocal tract (*λ/2)*. F1 depends on the Helmholtz resonator thus constituted, and F3 depends on the length of the constriction (*λ/2*). To produce the first two resonances typical of /i/ (λ/4 of front and back cavities). Can be generated by simply lowering the jaw. As the front cavity opens, a pharyngeal constriction is automatically produced. These two cavities are necessary For /u/, two Helmholtz resonators need to be produced, with similar volumes and necks. This configuration can be achieved regardless of the anatomical ratio between front and back cavity, since the tongue position alone divides the tract into two cavities and tongue and lip constrictions make the necks for these two resonances. With an adapted control all the vocal tracts we presented were able to produce speech sounds and then the Neanderthal vocal tracts.

7. CONCLUSION

With the anthropometric data that we currently have available for the skull and the vertebrae of fossils covering a period of 3 millions of years it is possible to estimate the limits of a plausible vocal tract. The typology of the Larynx Height Index for ontogenesis and phylogenesis shows a large range from 0.4 to 1.0 Compared to modern humans, these fossils possessed a vocal tract with a longer palatal section and likely a similar pharyngeal section. We are in a position to state that our ancestors and distant cousins were equipped with a vocal tract that could produce the same variety of speech sounds as we can today. But we do not know to what extent they mastered the control skills needed to produce speech. A low larynx (and large pharynx) can not be considered to be the *"anatomical prerequisites for producing the full range of human speech"* and there is no reason to believe that the lowering of larynx and in increase in pharynx size are necessary evolutionary pre-adaptations for speech. This result invalidate the well-known theory of Lieberman [10, 11] marred by a series of anatomic (hyoid position), articulatory and acoustic flaws (mainly confusion between anatomic parts of the vocal tract with front and back cavities of the vowel production [12, 13]. If Neanderthals could not speak, it is unlikely to have been for the articulatory and acoustic reasons advocated by Lieberman; they were physically able to do so.

With the emergence of speech, we are confronted with problems, with constraints and with 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.

Acknowledgements: Philippe Mennecier, director of the collections at the *Musée de l'Homme* (Paris, France), has kindly granted us the access to all the skulls. This research has been partly funded by the French CNRS (OHLL project) and by the European Community EUROCORES program (OMLL project); 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.*

8. REFERENCES

[1] Hauser, M.D., Chomsky, N., Fitch, W.T. 2002. The faculty of language. *Science* 298, 1569-1579.
[2] Schwartz, J.L., Boë, L.J., Abry, L.J. 2007. *Linking the DFT and the MUAF principle in a PAC Theory*. Hommage to J. Ohala. Berkeley.
[3] Boë, L.-J., Maeda, S. 1998. Modélisation de la croissance du conduit vocal. *Journées d'Études Linguistiques*, Nantes, 98-105
[4] Ménard, L. 2002. Production et perception de la croissance du conduit vocal. Ph.D. diss. Grenoble.
[5] Honda, K., Tiede, M. 1998. An MRI study on the relationship between oral cavity shape and larynx position. 5^{th} *ICSLP 2*, 437-440.
[6] Fenart, R. 2003. Crâniographie vestibulaire. Analyse morphométrique positionnelle. *Biométrie Humaine et Anthropologie*, 21, 3-4, 231-284.
[7] 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/
[8] Boë, L.J., Granat, J., Autesserre, D., Perrier, P., Peyre, E. 2006. Variation and prediction of the the hyoid bone position for modern man and Neandertal. *Biom. Hum. et Anthropol*., 257-271.
[9] Benoît, R. 2001. Development Biology, craniofacial genetics. *Edgewise J., 44, 9-40*.
[10] Lieberman, P., E. S. Crelin. 1971. On the speech of Neanderthal man. *Linguistic Inquiry.* 2:203-222.
[11] Lieberman, P. 2007. Current views on Neanderthal speech capabilities: A reply to Boë et al. *J. of Phonetics* (in press).
[12] Boë, L-J, J-L Heim, K. Honda, S. Maeda. 2002. The potential Neanderthal vowel space was as large as that of modern humans. *J. of Phonetics* 30, 465-484.
[13] Boë, L-J, J-L Heim, K. Honda, S. Maeda., Badin, P., Abry, C. 2007. The vocal tract of newborn humans and Neanderthals: Acoustic capabilities and consequences for the debate on the origin of language. A reply to Lieberman (2007). *J. of Phonetics*, in press.