

Editors: Louis-Jean Boë, Joël Fagot,
Pascal Perrier, Jean-Luc Schwartz

Introduction to “Origins of Human Language: Continuities and Discontinuities with Nonhuman Primates”

There have been a number of contributions in the past years about language origins from various points of view. In this book, we intend to contribute to establish a state-of-the-art of the knowledge about the continuities and ruptures between communication in primates and language in humans. A major strength of the present book is to explore a diversity of perspectives on the origins of language, including the description of vocal communication in animals, mainly in monkeys and apes, but also in birds, the study of vocal tract anatomy and cortical control of the vocal productions in monkeys and apes, the description of combinatorial structures and their social and communicative value, and the exploration of the cognitive environment in which language may have emerged from nonhuman primate vocal or gestural communication. Interestingly, this portrait emerges from a situation in which one long-standing hypothesis stating that a low larynx position was a prerequisite for the emergence of speech has been clearly discarded. Indeed, some contributors of this book have just participated to two papers showing that the monkey vocal tract was “speech ready” (Boë et al., 2017; Fitch et al, 2016). This renders the debates clearer, in that neurocognitive and social evolutions now unequivocally appear as the major potential sources of evolution towards language. The series of eleven chapters provides a rather complete portrait and elaboration on the facts, proposals, arguments and claims that pave the science way from animal communication to human language.

The book begins by a descriptive analysis of baboon calls by *Caralyn Kemp, Arnaud Rey, Thierry Legou, Louis-Jean Boë, Frédéric Berthommier, Yannick Becker and Joël Fagot*. In their study of the “*Vocal Repertoire of Captive Guinea Baboons (Papio papio)*”, the authors provide ethograms and a prototypical description of twelve kinds of vocalizations emerging

from the analysis of individual calls and call sequences in the vocal repertoire of a group of captive Guinea baboons. Typical sound examples of each type of vocalization are also provided in Supplementary Materials. This study will be of substantial value for students of primate vocalizations. More importantly in the context of the present book, it provides a concrete and significant example of the “phonetic” description of the vocal communication system in nonhuman primates, which contributes to the documentation of the precursors of human speech possibly enlightening the conditions of its emergence. Of importance here is the fact that exploitation of variations in various dimensions of the vocalizations appears as a possible way to increase the efficiency of communication without expanding the vocabulary of available units. Interestingly, the large co-variations between formants and fundamental frequency also suggest a non-independent mastery of vocal source and vocal tract configuration in baboons’ vocalizations.

The next chapter is in continuity with the previous one, providing a zoom on one of the twelve baboon vocalizations. *Louis-Jean Boë, Thomas R. Sawallis, Joël Fagot and Frédéric Berthommier* question “*What’s up with Wahoo? Exploring Baboon Vocalizations with Speech Science Techniques*”. Focusing on the “wahoo” vocalization, they analyze a corpus of 69 utterances of wahoo calls coming from the corpus of the previous chapter. Careful spectral analysis of these utterances provides major spectral peaks separately for the three proto-components {w}, {a} and {hoo}. These peaks are compared with those of a [wa.u] phonetic sequence uttered by a human speaker in various phonatory modes. In parallel, the authors propose an articulatory analysis of a film presenting a baboon uttering a wahoo vocalization. Altogether, they claim that these combined acoustic and articulatory analyses converge towards the assumption that baboon “wahoo” is rather similar to a human phonetic sequence that can be transcribed as [wa↓.u↑], with a first syllable chaining a back rounded semi-consonant /w/ and a front open /a/ produced in an ingressive way, and a back rounded /u/ produced in an egressive way.

The exploration of vocalizations in nonhuman primates continues with *Adriano R. Lameira* proposing a view on “*Origins of vowels and consonants: Articulatory continuities with nonhuman great apes*”. From his study of the call repertoire of orangutans, the author introduces the idea that there could exist an articulatory homology between voiceless calls and

human consonants on the one hand, and between voiced calls and human vowels on the other hand. Among the set of voiceless calls, Lameira focuses on whistles and shows clear learning abilities in captive orangutans, which relates to a number of reports of learning processes in other great apes. Concerning voiced calls, Lameira displays kinds of “babbling” vocalizations with rhythmic jaw movements similar to the ones displayed by infants, together with imitation games in which a captive orangutan is able to modify fundamental frequency in response to modulations of a human tutor. These plastic voiceless and voiced vocalizations could provide in the author’s view “proto-consonants and proto-vowels” in a kind of language precursor in a human ancestor.

Importantly, vocalizations in nonhuman primates are constrained by the anatomy of the orofacial system. This is at the core of the contribution by *Frédéric Berthommier, Louis-Jean Boë, Adrien Meguerditchian, Thomas R. Sawallis and Guillaume Captier* dealing with “*Comparative Anatomy of the Baboon and Human Vocal Tracts: Renewal of Methods, Data, and Hypotheses*”. This comparative anatomy aggregates a series of invaluable data enabling to qualitatively and quantitatively compare vocal tracts in baboons and humans. These data include (1) a dissection of two adult *Papio papio* heads, enabling detailed description of the vocal tract, the larynx and the tongue musculature, (2) fifty-six 3D MRI scans of *Papio anubis* baboons from 2 years to adulthood enabling authors to elaborate precise vocal tract biometry, (3) radiographic data for 127 human children from 3 to 7.5 years providing reference human biometry for comparison with the preceding set of Baboon data. This enables authors to claim that the hyoid bone would be placed one vertebra lower in human infants than in adult baboons – and also one additional vertebra lower in male human adults. The increase in the pharyngeal part of the vocal tract in humans would be accompanied by compensatory facial shortening, thus maintaining the vocal tract length similar in both species. On the basis of these data authors address the issue of how exaptation of articulatory patterns in feeding could have contributed to structure the articulation of speech sounds.

Vocalizations in nonhuman primates also depend of course a lot on the cortical and sub-cortical networks available for vocal and orofacial control. The question of cortical control is explored in the next two chapters. Firstly, *Veena Kumar and Kristina Simonyan* discuss in great detail the “*Evolu-*

tion of the laryngeal motor cortex for speech production". Their starting point is that, as already discussed in the first chapter by Kemp and coll., laryngeal control seems much more precise and stable in humans. Kumar and Simonyan analyze possible differences in laryngeal cortical control between humans and nonhuman primates. Firstly, they recapitulate several studies from their group leading to the conclusion that, while laryngeal motor control would be localized both in the primary motor cortex and in the premotor cortex in humans, localization would be reduced to the premotor cortex in apes and monkeys. Their hypothesis is that the premotor cortex would be sufficient for basic functions associated to e.g. breathing or physical effort, but the fine control in humans would require the additional involvement of the primary motor cortex. This evolution would be combined with the emergence in humans of direct cortical connections towards the brainstem, while they would be indirect in monkeys. Finally, the cortical network of connections between the laryngeal motor cortex and parietal and temporal regions necessary for learning and control would also be much more developed in humans.

William D. Hopkins then addresses the question of a potentially crucial cortical area for language, often considered as a potentially major piece in the emergence of language: Broca's area. His contribution, entitled "*Motor and Communicative Correlates of the Inferior Frontal Gyrus (Broca's Area) in Chimpanzees*", provides a rich synthesis of various types of comparative data about the Inferior Frontal Gyrus in monkeys, chimpanzees and humans. Firstly, he provides a detailed analysis of the literature on the morphology and cytoarchitectonics of Broca's area in primates and particularly in the species his group studied most, that is chimpanzees. While the Pars Opercularis (ParsO) and Pars Triangularis (ParsT) are difficult to define in the Inferior Frontal Lobe in monkeys, ParsO can be rather clearly defined in chimpanzees, but defining a ParsT homolog is less clear. Areas 44 and 45 present large volumetric expansions and more consistent leftward asymmetries in humans compared to chimpanzees, together with a larger amount of synaptic connections. The author then presents a number of results from his group displaying consistent correlations in chimpanzees between morphological properties of individual Inferior Frontal Gyrus and behavioral abilities associated with communicative actions (e.g. pointing manual gestures and attention-getting vocalizations) and tool-use. These

correlations seem partly genetically heritable. He concludes by discussing the implications of these findings in theories of language emergence.

The next two chapters explore the way vocalizations could indeed constitute a real communication system likely to open the road towards human oral language. Firstly, *Camille Coye, Simon Townsend and Alban Lemasson* discuss the question of combination and compositionality, in their chapter entitled: “*From animal communication to linguistics and back: insight from combinatorial abilities in monkeys and birds*”. From their analysis of the very wide literature on compositionality in bird songs and monkey calls, the authors attempt to carefully disentangle what could be a “phonological level” in which non-meaningful vocal units would be combined in various ways to provide meaningful sequences, and what could be a “morphosyntactic level”, in which meaningful units would be combined for producing larger meaningful structures. They argue that most reports in the literature do not provide convincing examples of nonhuman compositionality in one of these two strict senses. Then they present some “promising examples” of proto-phonology in the composition of flight calls in chestnut-crowned babblers (Australian birds), and protomorphosyntax in the composition of meaningful calls both in southern pied babblers (South-African birds) and in Diana monkeys from forests in West Africa. Finally, they suggest some possible social pressures driving the use of compositionality, in relation with the complexity of the social organization, and the habitat constraints on communication pushing for complex vocal communication with low ambiguity and long-range facilities of use.

Klaus Züberbühler then engages in a global evaluation of the ability of primate vocalizations to constitute “*The Primate Roots of Human Speech and Language*”. For this aim, he reviews the continuities between non-human primate vocal communication and oral language, but also some major limitations that can be traced in these continuities. First, the vocal tract seems speech-ready but cortical control is not sufficient to master the vocal source and the vocal learning and combination processes required in human speech. Second, the communicative content of the calls seems likely to be interpreted and even modulated by monkeys and apes in relation to context and audience. However, vocal call exchanges appear to convey low levels of intentionality – in reference to Denett’s scale – and nonhuman primates could lack the ability to share intentions and goals. The author also

addresses the question of referential communication, a crucial component of human language. Monkeys and apes do display elements of referential communication, but rather focused on themselves. The lack of clear view on the nature of their “mental concepts” sets severe limitations on establishing strong links with human language.

At stage, where the focus all over this book has been mostly put on vocal communication, the contribution by *Katja Liebal* provides a timely and important comparative overview of “*What gestures of nonhuman primates can (and cannot) tell us about language evolution*”. She begins by a review of arguments pros and cons for either vocal, gestural or orofacial communication as the possible unique precursor of human language, and she nicely shows that arguments in favor of one or the other are often partly incomplete or in some sense partial, and hence that no “unique precursor” theory is wholly convincing at this stage. Then, she focuses on what could be gained for a theory of language evolution by looking at gestural communication in monkeys and apes. Interestingly, this provides a number of echoes to the previous chapter by Zuberbühler, by discussing what aspects of gestural communication could display some continuity with human language. Intentionality is a basic component of communicative gestures, with clear evidence that both monkeys and apes monitor the attention of their partner and modulate communication accordingly. Flexibility – the ability to vary the context of use of a given stimulus – seems rather larger for gestures than for calls or orofacial productions. Gesture compositionality appears rather weak, with only one or two possible examples in the literature. Referentiality and iconicity are debated. Altogether, the author stays in a careful position in terms of the gestural vs vocal origin of human language.

The last two chapters open the angle of view even more widely by addressing the question of the cognitive environment required for the emergence of language. *Tecumseh Fitch* focuses on “*Dendrophilia and the Evolution of Syntax*”. Syntax is classically considered as a highly specific property of human language, and Fitch continues his exploration of the specific cognitive abilities that make humans special and could trace a major discontinuity in the emergence of language. He introduces the assumption that this ultimate human cognitive ability consists in the capacity to manipulate “supra-regular” grammars, thanks to a structural working memory providing generalization and elaborating hierarchies. This is what

he calls “dendrophilia” – a tendency to organize sensory flows into tree-like structures. The author reviews experimental data in which various animal species have been claimed to manipulate grammar-like structures. He raises objections to each of these studies, to argue that humans are the single species able to manipulate supra-regular grammars. He concludes on the likely implication of Broca’s area in the neural implementation of this uniquely human process.

Finally, *Joël Fagot, Raphaëlle Malassis, Tiphaine Medam and Marie Montant* adopt the inverse perspective by “*Comparing human and nonhuman animal performance on domain-general functions: towards a multiple bottlenecks scenario of language evolution*”. They propose an alternative to the search of a uniquely human capacity, and rather explore possible continuities and discontinuities in general cognitive abilities. They successively analyze integration in time and space, integration across sensory dimensions and sensory modalities, and various types of categorization processes. In each of these domains, they document resemblances between animals and humans, and aspects in which humans display a specific behavior. Humans appear better at processing and learning complex sequences, at extracting global aspects of visual scenes, at integrating sensory dimensions, at extrapolating perceptual properties in equivalence classes and elaborating qualitative rules and generalizing these rules across domains. This results in various types of “bottlenecks” that could have constrained the emergence of language. The authors conclude on the specific importance of attention and working memory in the bundle of factors that seem to have co-evolved in the route towards human language.

Although not exhaustive, we hope that the tour offered in this book will convey a clear sense of the progress that have been made in the field of language evolution, and also hope that this book will serve as resource for students and researchers in the field. We would like to thank all the contributors for their contributions.

Acknowledgements

Joël Fagot was supported by ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) and ANR-11-IDEX-0001-02 (A*MIDEX). Jean-Luc Schwartz was supported by a grant from the European Research Coun-

cil under the European Community's Seventh Framework Programme (FP7/2007–2013 Grant Agreement no. 339152, “Speech Unit(e)s”).

References

- 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(1), e0169321. doi.org/10.1371/journal.pone.0169321.
- Fitch, W.T., de Boer, B., Mathur, N., and Ghazanfar, A.A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2, 12, e1600723.