CellPress
OPEN ACCESS

Trends in
Cognitive Sciences

Forum

# Arboreal origin of consonants and thus, ultimately, speech

Adriano R. Lameira [1,*]

**The world's spoken languages are universally composed of vowels and consonants, but the primate prototypical call repertoire is almost exclusively composed of vowel-like calls. What was the origin of consonant-like calls? Their prevalence across great apes suggests that an arboreal lifestyle and extractive foraging were ecological preconditions for speech evolution.**

## Whence consonant-like calls?

Spoken languages are universally composed of vowels and consonants. Vowels virtually always take the form of voiced utterances, whereas voiceless utterances the form of consonants. The former are produced by the larynx (predominantly, vocal folds), the latter in the mouth by the manoeuvring of supra-laryngeal articulators (i.e., lips, tongue, mandible).

However, nearly all (non-human) primates' call repertoires are composed exclusively or primarily of voiced vowel-like calls. The comparative study of primate vocal behaviour has therefore led theories of speech evolution to place special and exclusive attention on primate laryngeal anatomy and human vowels' forerunners. Nonetheless, sooner or later, any attempt to successfully reconstruct speech origins must explain the root cause(s) of voiceless consonant-like calls in the human lineage. Prevailing hypotheses offer very few helpful hints, but (non-human) great apes provide a crucial opening lead.

Unlike other primates, but similar to any spoken language, great ape call repertoires are composed of consonant- and vowel-like calls [1]. The presence of calls articulatorily and acoustically homologous to consonants in great apes allows a first dedicated effort to investigate their evolutionary origin.

Curiously, the prevalence of consonant-like calls is not consistent across great ape genera, *Pongo* (orangutans), *Gorilla* (gorillas), *Pan* (chimpanzees and bonobos), opening a hatch for the study of evolutionary mechanisms. In captivity, all great apes seem capable of socially learning or inventing new voiceless consonant-like calls [2–4]. This implies similar baseline capacities, probably shared from a common ancestor. However, exposure and enculturation to humans and relatively uniform settings in captivity mask between-genera variation, making it difficult to make strong inferences about evolution. Conversely, striking disparities occur in nature. Wild orangutans produce consonant-like calls across multiple contexts, from nest-building to mother–infant communication [5]. Their most frequent call type is a consonant-like alarm call [6]. Orangutan consonant-like calls also occur as traditions [5] and previously undescribed types are typically found in newly surveyed populations. In wild gorillas, chimpanzees, and bonobos, the prevalence and variety of consonant-like calls are contrastingly lower. In gorillas, a single putatively cultural consonant-like call occurs in some populations, but not others [7]. In *Pan*, some wild chimpanzee populations produce one or two consonant-like calls (with possible subvariants) within a single context (i.e., social grooming) [8], but these are notoriously uncommon in other populations. Only in wild orangutans are consonant-like calls universal, cultural, and occur across multiple contexts, as in speech (Box 1).

Here, I describe how wild orangutans' arboreal lifestyle and feeding ecology might help explain the ubiquity of their consonant-like calls compared with African apes.

## Accessing concealed foods without tabletop

All great apes are accomplished extractive foragers. They have developed motorically and cognitively complex mechanisms to access protected (e.g., nuts) or hidden foods (e.g., plant piths), which tend to be highly nutritious and essential fallbacks during scarcity. These techniques require meticulous manipulation and oftentimes tools. The way individuals handle foods and tools is, however, heavily constrained by their immediate physical settings. For example, gorillas can engage in intricate food processing while sitting on the forest floor. Similar stability for delicate hand movement and coordination cannot, however, be expected from orangutans standing or balancing on tree branches. For similar reasons, stone and/or hammering techniques occur in chimpanzees, but there are no stones up in the trees nor are there stable spaces for placing foods for impact. Orangutans would have to repeatedly descend trees to recover items until successful, whereas chimpanzees and bonobos can easily hold foods and tools with whichever combination of feet and hands suits best the task and skill level.

## The mouth as a fifth hand

Accessing valuable foods up the canopy is a different game altogether. Assuring body stability for behavioural action by large and heavy arboreal species, such as orangutans, requires constant use of one or more limbs. Depending on body position and pose, there are fewer limbs available for handling food and tools. Orangutans have bypassed some of these limitations by becoming apt users of their lips, tongue, and mandible as a 'fifth' hand to hold and process foods or position and manoeuvre tools. This fine oral neuromotoric control has become an integral part of their biology. For instance, when handed a pen by a human caretaker, a

### Box 1. At odds with observation biases

Present variation in great ape voiceless repertoires is all the more remarkable because an inverse pattern should be expected between ground-dwelling and arboreal genera. First, primatological research has been historically longer and more intense in Africa than insular refugia in Sumatra and Borneo. All else being equal, this ought to translate, after collective centuries of research time, in more voiceless calls known for African apes, not the opposite. Second, African apes occur in larger numbers and wider territories than *Pongo*. Accordingly, chimpanzees, for instance, might be expected to show more voiceless call types, produce them more often and across more contexts, not fewer. Third, African great apes are more sociable and voluble than Asian. For example, it is common for a solitary wild orangutan to produce no call during an entire nest-to-nest day, something unheard of with *Gorilla* or *Pan* groups. Finally, audio recordings are logistically easier with ground-dwelling apes. They can be collected closer to subjects, without obstruction or background noise of moving foliage, as it occurs with arboreal individuals. Ultimately, this should have made it easier to identify voiceless (often soft) calls in ground-dwelling apes, so it is striking that voiceless calls have been more extensively catalogued in orangutans.

captive orangutan will normally pass it back with her mouth, not her hand, and orangutans are known to peel oranges with their lips and no hand-aid [9], displaying overall higher oral skill and prowess than African apes, even in captivity, where they exhibit a mostly ground-bounded lifestyle [9].

## Orangutans with feeding manners smack

High prevalence and frequency of voiceless consonant-like calls in orangutans seem, thus, to have resulted from selection pressures driven by the demands of eating up the canopy. The fine control required for manipulating food and tools with their mouth has translated into enhanced motoric dexterity over the lips, tongue, and mandible. This degree of enhanced skilful control in orangutans appears to have translated into greater ease for the production of voiceless consonant-like sounds that depend directly on the same anatomical structures, ultimately resulting, compared with African apes, in a repertoire rich in smacks, raspberries, clicks, splutters, sizzles, and kiss sounds.

## Did speech come from above?

Currently, close to nothing is known about the origins of consonants in the human lineage. While largely absent in most branches of the primate order for 60 my, their arise in ancient hominids some 15 mya marked a point of no return in the process that ultimately moulded speech from an ancestral ape-like call repertoire. The lifestyle and feeding ecology of great apes have led to a striking variation in the number and type of consonant-like calls observed in arboreal versus ground-dwelling genera (Box 1). Accordingly, perch feeding may have also offered a potent built-in engine for consonant-like call production and diversification in ancestral preverbal hominids living up the trees versus on the ground. Arboreality seems, hence, to have been a preadaptation for speech evolution in human ancestors.

To test this evolutionary scenario, cross-/within-genus comparisons will help enlighten whether populations spending more time in trees produce more (often) consonant-like calls (e.g., *Gorilla* < *Pan* << *Pongo*). Equally, greater freedom of movement and position on the ground could hint at richer hand-to-tool, gestural, or postural repertoires on *terra firma*. Neuro-physiological characterisation of efferent innervations to oral articulators or standardised behavioural tests of oro-facial command could also offer a means to quantify and compare motoric scale and scope between genera.

For now, the proposed scenario for the evolutionary genesis of consonants aligns with evidence that human bipedalism arose in an arboreal context as an adaptation for locomotion on flexible branches [10] and that arborealism remained a crucial component in the lives of ancestral hominins [11] up until modern humans [12], often for accessing food. These converging data invite consideration of the knock-on effects that tree-living could have had on the communicative, cognitive, and cultural systems of human ancestors, inciting new research on the ecology of speech and language evolution (Box 2).

Continuing to chart the vocal repertoire of great apes across different habitats will shed new insights into how neuromotor capacities, ecological context, and path-dependent evolutionary processes worked in concert to generate among hominids a signal system as unique as speech.

### Box 2. Framing a phylogenetic family picture

Only three great ape genera survive from a once-diverse family. Despite this small sample size, observed differences in voiceless call repertoires (Box 1) are a compelling sign that there are selective forces driving hominid consonant proliferation and function [6] that science is yet unaware of. The proposed relationship between oro-facial control for feeding and call production is for now correlational but testing the predictions of this hypothesis (see main text) will help reveal potential causal mechanisms at work.

Compared with anatomical components and cognitive computations that make speech and language possible, the wider ecological context of these elements' evolution has remained virtually ignored. The likelihood that tree-living provided pre-adaptations necessary for spoken language suggests the intriguing possibility that human ancestors may have been more arboreal than African apes, despite close phylogenetic relatedness. Concurrently, invention and imitation of consonant-like calls by all great apes (from humans and other apes) in captivity [2–4], whose lives are mostly terrestrial, suggests that consonant-like call adoption could have been prompted during development, through learning and practice. Therefore, the rich interactions between innate, epigenetic, and social factors that underpin child language acquisition may have been already (at least partly) at play in preverbal ancestral hominids [13].

## Acknowledgments

I thank Nathaniel Dominy and three anonymous reviewers for constructive input on a previous version of the manuscript. I was funded by UK Research and Innovation, Future Leaders Fellowship, grant agreement number MR/T04229X/1.

## Declaration of interests

The author declares no competing interests.

[1]Department of Psychology, University of Warwick, Coventry, UK

*Correspondence:
adriano.lameira@warwick.ac.uk (A.R. Lameira).

https://doi.org/10.1016/j.tics.2022.11.012



## References

1. Lameira, A.R. *et al.* (2014) Primate feedstock for the evolution of consonants. *Trends Cogn. Sci.* 18, 60–62
2. Lameira, A.R. *et al.* (2013) Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: a replication and extension. *J. Acoust. Soc. Am.* 134, 1–11
3. Salmi, R. *et al.* (2022) A novel attention-getting vocalization in zoo-housed western gorillas. *PLoS One* 17, e0271871
4. Taglialatela, J.P. *et al.* (2012) Social learning of a communicative signal in captive chimpanzees. *Biol. Lett.* 8, 498–501
5. Wich, S.A. *et al.* (2012) Call cultures in orang-utans? *PLoS One* 7, e36180
6. Lameira, A.R. *et al.* (2017) Proto-consonants were information-dense via identical bioacoustic tags to proto-vowels. *Nat. Hum. Behav.* 1, 0044
7. Robbins, M.M. *et al.* (2016) Behavioral variation in gorillas: evidence of potential cultural traits. *PLoS One* 11, e0160483
8. Watts, D.P. (2015) Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: variation, social learning, and possible functions. *Primates* 57, 61–72
9. Walker, P.L. (1976) The adaptive significance of prehensile lips a functional analysis of pongid lip mobility. *Am. J. Phys. Anthropol.* 44, 213
10. Thorpe, S. *et al.* (2007) Origin of human bipedalism as an adaptation for locomotion on flexible branches. *Science* 316, 1328–1331
11. Kappelman, J. *et al.* (2016) Perimortem fractures in Lucy suggest mortality from fall out of tall tree. *Nature* 537, 503–507
12. Kraft, T.S. *et al.* (2014) A natural history of human tree climbing. *J. Hum. Evol.* 71, 105–118
13. Lameira, A.R. *et al.* (2022) Sociality predicts orangutan vocal phenotype. *Nat. Ecol. Evol.* 6, 644–652