

How cooperation drove the evolution of vocal production learning in hominins

Sabine Stoll^{a,*}, Carel P. van Schaik^{a,b,c,*}, Konatsu Ono^b, Nikhil Phaniraj^b, Zitan Song^c,
Judith M. Burkart^{a,b,*}, and Balthasar Bickel^{a,*}

^aInstitute for the Interdisciplinary Study of Language Evolution, University of Zurich,
Affolternstrasse 56, 8050 Zurich, Switzerland

^bInstitute for Evolutionary Anthropology, University of Zurich, Winterthurerstrasse 190,
8057 Zurich, Switzerland

^cDepartment for the Ecology of Animal Societies, Max Planck Institute of Animal
Behavior, Bücklestrasse 5, 78467 Konstanz, Germany

*Contributed equally.

Abstract

Language evolved uniquely in the hominin lineage. Critical to its evolution was the emergence of **vocal production learning (VPL)**, the ability to acquire novel vocalizations. The mechanisms underlying VPL have been extensively studied across taxa and are convergent with hominins. However, the drivers of hominin VPL are still unknown. Reviewing the various functions of VPL across taxa, we propose that hominin VPL was selected for by the increased need for cooperation-induced communication in the new hunter-gatherer niche that formed after the move to the savanna at ca 2 Ma ago. The transition rests on two processes: (1) the amplification of an inherited mechanism of vocal accommodation, in which individuals align acoustic features of their vocalizations to enhance affiliative bonds; and (2) the emergence of babbling, in which infants train interactive learning mechanisms. Our perspective opens new directions for research on language evolution, focusing on the communicative needs of hunter-gatherer cooperation.

This manuscript was compiled on September 19, 2026

Author contributions: Conceptualization: S.S., C.v.S., J.B., B.B.; Methodology: S.S., C.V.S., Z.S., B.B.; Formal Analysis: Z.S., B.B.; Investigation: S.S., C.v.S., K.O., N.P.; Writing — original draft: S.S. C.v.S.; Writing — review and editing: S.S., C.v.S., K.O., N.P., J.B., B.B., Visualization: S.S., C.v.S., Z.S. B.B.; Project administration: S.S., C.v.S., Funding acquisition: S.S. C.v.S., J.B., B.B.

Introduction

Human language is among the most distinctive and transformative traits of our lineage, yet its evolutionary origins remain among the deepest unresolved problems in the biological and cognitive sciences. Although humans shared a common ancestor with chimpanzees and bonobos as recently as ~ 8 million years ago [1], our vocal communication system has diverged profoundly from those of our closest living relatives. Whereas great apes possess relatively fixed, phylogenetically conserved vocal repertoires [2], human language consists of diverse, rapidly evolving systems that are acquired through extensive learning rather than inherited as fixed repertoires [3]. This stark discontinuity raises a fundamental question: what processes enabled the transition from a small innate vocal repertoire to the learned and culturally transmitted system unique to humans? Explaining this transformation is a central challenge for evolutionary theory.

Previous accounts have often sought to explain this transition by invoking the absence of cognitive mechanisms necessary for language-like communication in early hominins. However, a growing body of evidence suggests that primates, particularly great apes, possess many of the cognitive capacities required to support key components of language. These include advanced auditory statistical learning, supported by homologous brain networks [4], aspects of theory of mind and imagination [5–7], pragmatic inference [8], and audience-aware, strategic communication [9–14]. In addition, great apes exhibit capacities for imitation and cultural transmission [15–17] and human-like conceptual structure [18–22]. Their communicative systems also show limited forms of compositionality and hierarchical organization in call combinations [23–28], sensitivity to non-adjacent dependencies and suprarregular grammars [29–31], and even modest forms of communicative creativity [32–36] that support group-specific variation in vocal repertoire (Figure 1). Taken together, these findings suggest that the cognitive foundation required for at least simple forms of language were already present in the last common ancestor of hominins. This challenges accounts that attribute the emergence of language primarily to the evolution of novel cognitive capacities, and instead point to other factors as the key drivers of this transition.

In light of this evidence, we adopt a functional perspective, focusing on why language evolved rather than how it evolved. If the initial step toward language involved an expansion of the vocal repertoire, the critical prerequisite was the ability to produce novel vocalizations. A gradualist scenario suggests that social and ecological pressures acted on preexisting vocal abilities, progressively shaping them into the flexible form of vocal learning observed in humans, i.e. vocal production learning (VPL) [37]. To identify the conditions favoring the evolution of VPL, we draw on convergent cases across taxa. We further consider the social functions of a more basic form of vocal flexibility, vocal accommodation (in the human literature also referred to as *vocal matching* [38], and as *interactive alignment* [39]), in which existing calls are modified without becoming qualitatively different [40]. We argue that this mechanism provided the evolutionary substrate from which VPL emerged. Integrating comparative evidence with archeological data, including changes in brain organization inferred from fossil endocasts, we identify the selective pressures shaping VPL in hominins, which increasingly relied on close cooperation in novel tasks introduced by the emerging hunter-gatherer niche. Building on these lines of evidence, we propose an evolutionary scenario in which language arose from a rudimentary form of vocal learning present in the hominin ancestor and was progressively elaborated to support the coordination and negotiation of novel cooperative action. This comparative functional framework offers a complementary perspective on language origins by linking convergent patterns across species to the unique evolutionary trajectory of human vocal production learning.

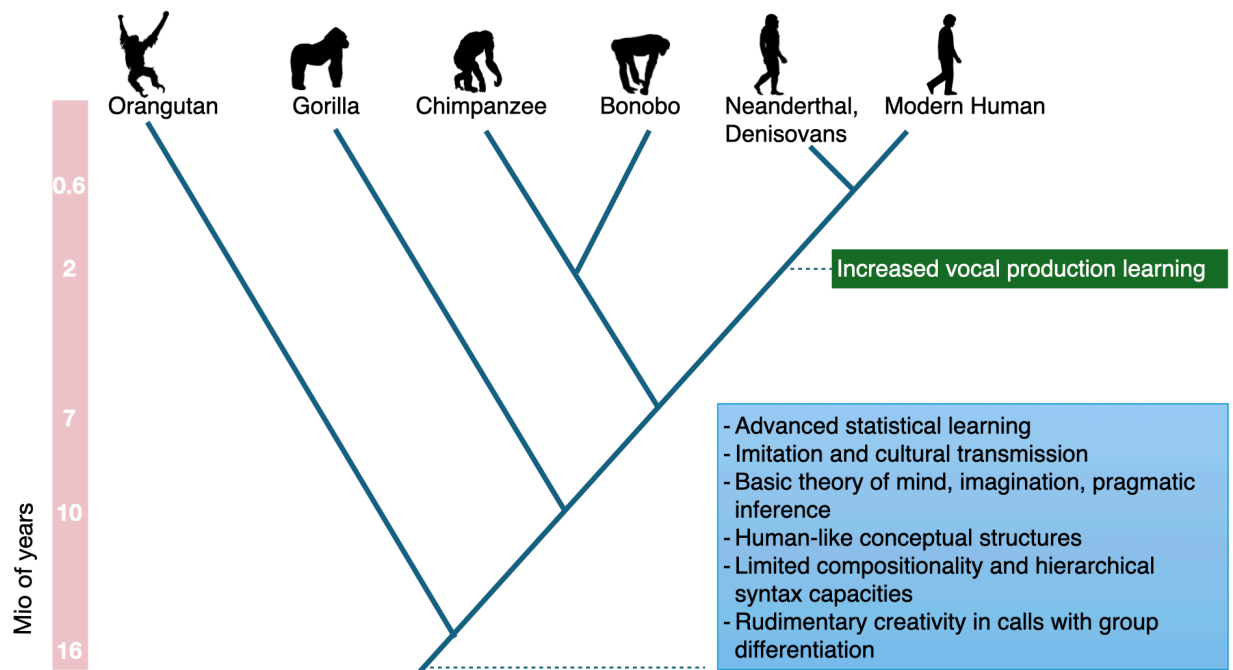


Figure 1: Many key elements of language are shared across great apes, often across primates more widely (see main text for references). What is distinctive in the *Homo* genus is extensive vocal production learning, the ability to learn new vocal signals in support of repertoire expansion and language diversification.

BOX 1 Definitions of terms used

Vocal learning	We use this as a cover term for all learned modifications of the vocal signal in the sense of [41], including vocal accommodation but excluding vocal usage learning (learning how to associate a given call with a specific context of use)
Vocal production learning	The ability to learn novel vocalization as a result of auditory experience with the vocalizations of others [37]. In terms of this definition only a few species with complex vocal abilities have been categorized as vocal learners.
Vocal accommodation	Vocal accommodation is the adaptation to the vocalization of a conspecific in auditory or structural features as a function of interaction. Accommodation can be short term, in direct interaction, or have long term effects leading to acoustic and structural changes [40, 42, 43].
Homophily	The term describes a social principle stating that social connections are more likely to form and be strengthened between interactional partners who are similar to one another, a concept first introduced by Lazarsfeld and Merton in the 1950s [44] and later formalized in social network theory [45].
Hominin	A lineage of great apes (hominids) characterized by being adapted to bipedal locomotion; it presumably arose in Africa during the late Miocene
Homology	A homologous trait is one that was inherited from a common ancestor.
Convergent evolution	A convergent trait is one that is functionally similar but evolved independently in different species.

The vocal learning spectrum and the origin of speech

Interest in vocal learning has grown substantially since Janik and Slater [37] defined it as the ability to modify vocalizations based on experience with vocalizations of others. Comparative research now reveals considerable variation in vocal learning abilities across taxa, suggesting a continuum from vocal non-learners, such as chickens, to high-capacity vocal learners, such as parrots and humans [41, 46–49]. Some also argue that this variation may be better characterized as a multi-dimensional spectrum than a simple linear scale [50, 51].

A fundamental distinction within this spectrum is that between vocal accommodation, referring to acoustic modifications of existing vocalizations during social interaction, and vocal production learning (VPL), the ability to produce and accurately imitate novel vocalizations [41, 46, 52]. VPL provides the key mechanism for extending the innate vocal repertoire. Species capable of VPL are generally also capable of vocal accommodation, [41, 53], whereas the reverse does not hold. In the following, we use the term *vocal learning* as an umbrella category encompassing both processes.

Because VPL is essential for the mastery of human speech, much research on language evolution has focused on species that exhibit this capacity [41, 54]. VPL has been documented in a

limited number of vertebrate lineages [51, 55, 56]: among birds, across at least three orders; among mammals, in bats, elephants, cetaceans, and pinnipeds, but it is generally assumed to be absent in nonhuman primates. Evidence suggests that VPL has evolved convergently in these lineages [55, 57]. Comparative analyses further reveal striking convergence in neuroanatomical mechanisms and associated genetic architectures [55, 58–60], potentially suggesting deep homologies and highlighting both the evolutionary constraints and opportunities that shape complex vocal learning.

Despite this, empirical data on taxonomic distributions remains incomplete. VPL is traditionally recognized using three criteria: (1) the absence of vocal development in young animals that are deaf or deprived of vocal input; (2) the spread of novel vocalizations in natural populations; or (3) the ability to imitate novel sounds outside the species’ innate repertoire. In practice, Criterion 1 is rarely applied, and Criterion 2 is largely restricted to cases in which new songs (rather than calls) spread in birds or cetaceans [61]. Closer examination of Criterion 3 indicates that it may generate numerous false negatives, suggesting that VPL could be more widespread than currently recognized. For example, immature bats and pinnipeds give isolation calls that are modified in response to vocal input, a form of accommodation, but they can also reproduce experimentally generated artificial calls [62, 63], qualifying them as VPL. Accordingly, the lack of targeted studies leaves numerous potential candidate species as non-VPL. Similarly, documented cases of VPL in mammals (elephants, cetaceans, pinnipeds) or parrots often involve mimicking human vocalizations well outside the range of species-specific repertoire [47, 49, 51]. In many instances, this evidence comes from captive individuals, even for well-studied species such as Asian elephants or most parrots [64]. By the same logic, hand-reared chimpanzees that learned a few English words may also show VPL [65], particularly in light of evidence that that primate vocal capacities are constrained more by neural and motivational factors than by vocal anatomy [66]. Indeed, rudimentary VPL capacities might have been underestimated in great apes [10, 67]. Excluding species for which only accommodation is documented may therefore underestimate the prevalence of VPL and limit efforts to reconstruct the contexts in which it evolved.

It is intriguing to note however that nonhuman primates show extensive evidence for vocal accommodation [43]. Chimpanzees, for instance, show both short-term [68] and long-term accommodation in captivity [69] and in the wild [12, 70, 71]. These accommodation abilities could readily have been turned into VPL when bolstered during hominin evolution, potentially exploiting rudimentary VPL capacities shared with other great apes. Second, while volitional control over vocalizations is a precondition for VPL, species with vocal accommodation may also be capable of volitional control to a substantial extent. This is indeed well-documented throughout life in common marmosets [72] and among juvenile rhesus monkeys [73]. Finally, the cytoarchitecture and connectivity of Broca’s area involved in human speech production show great similarities with the homologous structures in non-human primate brains [74, 75], with only gradual differences, chiefly in the form of prolonged, learning-friendly development [76, 77]. Together, these findings point to an absence of categorical constraints on the emergence of VPL in our hominin ancestors, underscoring the need to focus on functional aspects of vocal learning in the broad sense, including accommodation. We therefore compiled data on taxa with evidence for both VPL and vocal accommodation to identify the functional contexts in which VPL and accommodation are found.

Functions of vocal learning

VPL and vocal accommodation serve a wide range of functions [43, 78] (Figure 2A), which our survey groups into five broad classes (see Supplementary Information for sources). We begin with well-established cases of VPL and then turn to cases of vocal accommodation.

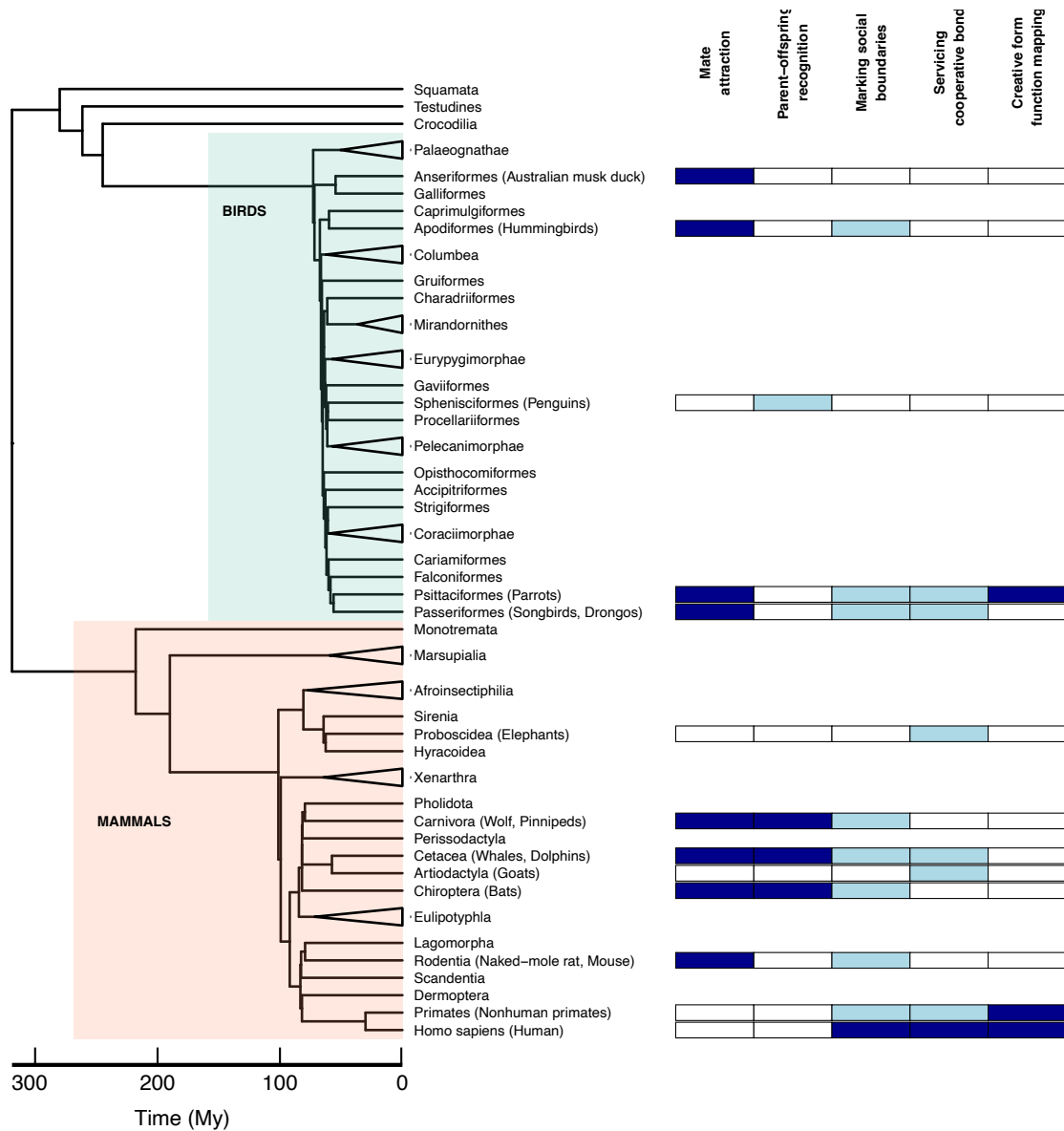


Figure 2: The known taxonomic distribution of two main kinds of vocal learning (vocal production learning in dark blue, and vocal accommodation in light blue) by function across amniotic vertebrates. Tip taxa are orders, but where vocal learning is recorded for isolated taxa within orders, these are mentioned after the order name (for details, see the Supporting Information). The evolutionary most important functions in humans revolve around social bonds and boundary marking, with other functions either absent in primates or shared (to different extents)

The first functional context of VPL is sexual selection, in which vocal learning is used to attract mates or defend territories. This function is widespread in birds, where one or both sexes produce songs that are often complex and vary across time and space [79]. Beyond passerines and parrots additional cases, such as hummingbirds and Australian musk ducks, are primarily found in lekking species [56], where male vocal displays serve to attract females. Similar functions have been documented in mammals such as whales [80] and harbor seals [81]. The learning process typically involves phases of imperfect imitation (e.g. subsong in birds), auditory feedback, and in some cases, adult tutoring, although direct imitation, as in vocal mimicry [82] also occurs. Such learning is often constrained to sensitive or critical periods during development [83].

Mate attraction is unlikely to have been the primary driver of VPL in language. Since Darwin [84], it is sometimes invoked as a driver for VPL in song and music. However, despite similarities in learning [85], these behaviors differ substantially from language. Most critically, they lack semantic compositionality, whereby novel meanings are expressed from binding together the meaning of components in strategic ways (syntax). As a result, song is used to store and transmit meaning only when accompanied by words. Unlike song or music, the primary functions of language are to think, inform, and instruct (for both technical and normative knowledge), to coordinate joint action, and to gossip, which helps assessing the reliability of cooperation partners. These differences imply that selective drivers of VPL in language rest on different ecological and behavioral contexts than those of VPL in song, dance, and music.

A second functional context of VPL involves vocal recognition between parents and offspring, particularly in contexts where numerous dependent young cluster together and must reunite with their parents after separation. This is most common in breeding colonies, especially among species with mobile young, such as sea birds, pinnipeds, or bats, where parents and young can be separated for extended periods. In such conditions, failure to reconnect is often fatal for the young. Accordingly, strong selection pressure is expected to favor modifications of existing calls that enhance individual distinctiveness and improve mutual recognition. This pressure promotes vocal learning, and more broadly, vocal production that exploits a larger acoustic space. Many examples from pinnipeds and bats fit this functional framework, and we hypothesize that parents-chick exchanges in colony-breeding birds, such as penguins, exhibit similar properties. The case of individualized signature whistles in dolphins may also fall within this domain. Although mother-calf bonds are strong, older calves increasingly undertake independent excursions. Because visual or gestalt recognition becomes unreliable at distance in aquatic environments, particularly in murky water, individuals rely heavily on vocal cues. By actively modifying calls away from modal patterns, pairs can expand their acoustic space, thereby reducing recognition errors and enhancing mutual identification.

Like mate attraction, this context is unlikely to have provided the foundation for linguistic VPL in hominins. First, during early infancy, primate young remain on or in close proximity to the mother, such that mutual recognition is ensured by a combination of vocal and olfactory cues, individual movement and gesture profiles, and most prominently: visual information, as demonstrated by numerous experiments on individual recognition from photographs or videos [86]. Second, by the time language is likely to have emerged, hominins no longer inhabited dense environments, and children were unlikely to be separated from other group members over substantial distances. Third, long-distance calls are facilitated by air sacs, which impair fine vocal control and thus speech production [87]. Consequently, simultaneous selection for long-distance calling and human-like speech is biomechanically incompatible.

A third function is the marking of social boundaries. This function enables individuals to distinguish friend from foe in low-visibility environments. Social boundary marking can operate at both dyadic and group levels. First, contact calls may facilitate coordinated movement; where kin or bonded partners are more likely to respond, vocal convergence can enhance vocal recognition (as

in the function of parent-offspring recognition) and thus be favored by selection. Second, convergence might be selected in contexts where fissioning groups are surrounded by hostile neighbors, making it advantageous to discriminate among conspecific callers rather than to respond indiscriminantly. This function is illustrated by differences between communities in chimpanzees [70, 88], which engage in lethal raiding of neighboring chimpanzee groups [89], as well as other nonhuman primates [90, 91]. We propose that these conditions may also explain vocal convergence of group members in other taxa, including elephants and cetaceans, and in pair-living species that forage apart, such as hornbills, parrots, or marmosets. The predominant mechanism here is likely vocal accommodation, as modification always occurs within the same call type, although species capable of VPL may exhibit more extensive changes. Importantly, social boundary marking may also emerge as a byproduct of dyadic vocal accommodation, particularly in small groups, and should therefore be demonstrated independently. The strongest evidence for this would be a pattern in which neighboring social units are more acoustically distinct in long-distance calls than in other calls, relative to more distant social units within the same population.

In humans, the marking of social boundaries through vocal accommodation is taken to an extreme. Accommodation serves not only to increase similarity within dyads or groups, but often also to accentuate differences between groups. This process, termed schismogenesis [92], is well documented [93–96] and widely regarded as a major driver of language change [3]. For this mechanism to have been playing a role in the emergence of language, however, interactions between social units must have been sufficiently frequent and competitive. Because even extant mobile hunter-gatherers live at very low population densities [97], such conditions were unlikely in the early stages. This suggests that schismogenetic processes became evolutionarily significant only later, as increasing densities intensified intergroup contact and competition [98–100].

Fourth, beyond the marking of social boundaries, vocal accommodation can also strengthen same-sex social bonds and between-sex pair bonds, thereby enhancing coordination and collaboration. This function has been identified in African penguins [101], chimpanzees, and various monkeys [43], most extensively in callitrichids, in which breeders form especially strong social bonds and newly formed pairs exhibit increasing vocal convergence as their bond strengthens [102, 103]. Strong convergence is also observed in systems where the recipient’s identity is encoded in contact calls, as in spectacled parrotlets [104], dolphins [105], African elephants [106] and marmosets [107]. These cases reflect contexts in which an individual label for both the sender and the intended recipient is adaptive [64], as the benefits of social interaction are dyad-specific. In humans, this bond-strengthening function of vocal accommodation, a form of vocal mimicry, is especially pronounced and pervasive. Humans universally engage in vocal accommodation. Interacting dyads converge across acoustic, syntactic and lexical features, producing short-term and in some cases also long-term effects [40, 42, 108]. In fact the amount of accommodation exhibited by interlocutors has been shown to directly contribute to the success in collaboration [109].

Vocal accommodation is typically accompanied by the broader mimicry of posture and gesture in both humans and other primates, contributing to homophily by fostering trust and strengthening social bonds [110]. In turn this facilitates cooperative exchange and the coordination of joint activities. We refer to this integrated phenomenon as *vocal homophily*. This function of vocal learning appears particularly relevant for language evolution. In species exhibiting this pattern, including humans, the two functions for vocal convergence, individual recognition and bond servicing, are closely intertwined, as both serve to establish and reinforce social relationships and signal shared group membership.

A final functional context is the creative mapping of form to function for transmitting information. This capacity is central to human language, but evidence in other species is extremely limited and often tentative. Among non-humans, parrots provide the clearest, but still not un-

equivocal, examples. Parrot males and females form stable, long-term pair bonds. Both sexes have been reported to produce novel, idiosyncratic calls used for referential communication, particularly to convey information about toxic or well-protected foods to their partners [64, 78]. In primates, creative form-function mapping is very rare. It has been documented only to a very limited extent and (so far) only in wild orangutans [32], and human-reared chimpanzees [65]. In all cases it is restricted to dyadic interactions, similar to what is observed in parrots. These examples suggest that, outside of humans, the emergence of truly referential, creatively mapped vocal signals is exceptionally uncommon and appears primarily tied to close social bonds.

In conclusion, the context in which VPL is most relevant for the emergence of language is one characterized by strong bonds within dyads or small groups. Here vocal homophily supports cooperation, while at the same time creating conditions in which novel form-function mappings are also advantageous (Figure 2). This capacity builds on a widespread foundation for reciprocity and joint action shared with other great apes. When these cooperative behaviors are intensified to better meet new challenges that require intensive cooperation, vocal homophily enhances their coordination. Simultaneously, they enable greater exploitation of capacities of creative form-function mappings. Building on this framework, we next propose a more detailed scenario linking vocal accommodation to the emergence of vocal production learning in the human lineage.

An evolutionary scenario: from accommodation to fully-fledged VPL

Our starting point is that a cooperative orientation is a critical precondition of the emergence of language [111, 112] rather than cooperation being a downstream product of language. On this view, language evolved to increase the efficiency of cooperation in both the social and technological dimensions of foraging, as well as in child rearing [72, 78, 113–118]. In light of the preceding argument that language builds on vocal homophily in cooperation in small units, we turn to the archaeological record to identify when such selective pressures emerged. We focus on the early Pleistocene, where evidence indicates the rise of intensified cooperation among hominins [119–122].

Specifically, we propose that the key functional context for the evolution of human language was the expansion and intensification of cooperation associated with a sustained shift into open savanna environments around 2 million years ago [123, 124]. Unlike other great apes, African *H. erectus* progressively occupied a unique hunter-gatherer niche, shaped by increasing demands for social transmission as complex technology accumulated [125], as well as by social organization that heightened interdependence both within and between the sexes [120, 126]. Multiple lines of evidence indicate that *H. erectus* engaged in high levels of ecological cooperation, supported by trust and durable social bonds [122, 127–129]. Ecologically, hominins thus evolved toward a highly interdependent foraging niche, with increasing reliance on vocal coordination, a feature that continues to characterize extant nomadic foragers [119, 130]. Taken together, paleoanthropological evidence supports the scenario outlined in Figure 3, in which intensified cooperation transformed vocal accommodation into vocal production learning.

Evidence for increased cooperation extends beyond the archeological indicators of cooperative foraging. Converging lines of evidence suggests that *H. erectus* adopted forms of extensive allomaternal care, or cooperative breeding, in which dependent offspring receive care from both parents and other group members [119]. First, fossil evidence of survival following debilitating injuries indicates sustained care and support for impaired individuals [134, 135]. Such prosocial behavior is rare among mammals and is typically confined to cooperative breeders [119]. Second, cooperative breeding in mammals is linked to increased brain size [136]. Notably, the brain size of *H. erectus*

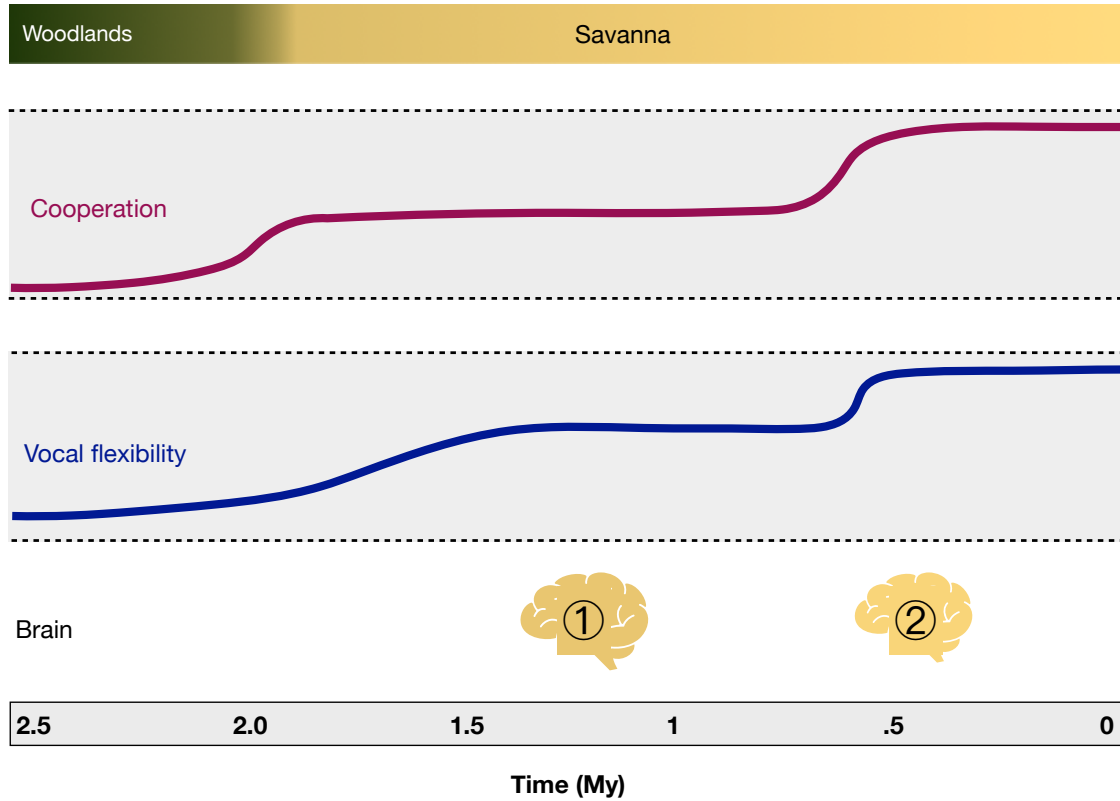


Figure 3: Proposed evolution in hominins. The expansion into the savanna intensified cooperation in the newly emerging hunter-gatherer niche, selecting for newly evolving brain structures around 1.25Mya [131], propelling accomodation for bonding in pairs or small groups into increased levels of VPL, and expanding creativity in vocal repertoires (1). VPL skills improved dramatically again in response to increased needs for social boundary marking when populations grew [99, 100, 120] and brain structures changed again after the split form Neanderthals and other lineages [132, 133] (2).

tus falls outside the demographically viable range observed in great apes [136], implying enhanced reproductive support, most plausibly through cooperative breeding [137]. Expansions in the prefrontal and temporal cortex [131] likely strengthened neural systems underlying social cognition and cooperation [138]. Third, *H. erectus* occupied highly seasonal habitats, a pattern commonly associated with cooperative breeders [139–141]. Fourth, increasing technological complexity in the African archaeological record is consistent with enhanced information sharing, including teaching [142, 143], and improved group-level problem solving [144], both characteristic of cooperative breeders.

Importantly, numerous features of communication in the only cooperatively breeding nonhuman primates, the callitrichids, provides converging support for the hypothesis that extensive allomaternal care favors more cooperative and more complex communication systems. Callitrichids engage in parental care and food provisioning, both of which rely on extensive communicative coordination. Consistent with this they exhibit particularly pronounced vocal accommodation compared to other primates. Their communicative profile includes relatively large repertoires, high volubility, vocal flexibility following pair formation as well as over short time scales, infant babbling, turn-taking in calling, vocal tutoring, and a degree of voluntary control over vocal productions [143]. Together, these features closely parallel key components of the “interaction engine” underlying human language [145], which is widely assumed to have preceded the emergence of language [146].

Notably, most of these features are largely absent in the great apes. We therefore suggest that their emergence in hominin evolution enabled the transformation of an ancestral great-ape-like communication system into a more flexible and dynamic system. This shift was likely reinforced by selection pressures for an expanded vocal repertoire, driven by the need to coordinate increasingly complex technologies and behaviors within a highly cooperative socioecological niche.

More general support for an effect of allomaternal care on enhanced vocal communication comes from comparative evidence across taxa. In birds, cooperatively breeding species exhibit larger vocal repertoires [147], particularly due to an expansion of alarm and contact calls. Among mammals, a similar pattern is observed in herpestids with variable rearing systems: the strictly cooperatively breeding meerkat shows the largest repertoire, again, largely driven by sentinel and alarm calls [148]. These patterns suggest that repertoire expansion associated with cooperative breeding primarily reflects increased demands for warning and movement coordination, that is, for cooperative (prosocial) activities.

An increase in repertoire size does not by itself constitute VPL. What is needed in addition is a mechanism that facilitates the acquisition of new vocalizations by immatures. We propose that the critical step here is the concomitant emergence of babbling. Babbling has been proposed as a prerequisite for VPL, providing a developmental substrate for the flexible acquisition and refinement of vocalizations [149–152]. A key function of babbling is vocal practice, trying out new sounds. Consistent with this, babbling emerges in human infants at around 6–8 months of age [150] when anatomical changes in the vocal tract first enable the production of more speech-like sounds [153, 154]. Over time, these vocalizations become increasingly similar to those of the ambient language [155].

Consistent with this, it has been reported in nearly all species exhibiting VPL, with research primarily focused on humans, songbirds, parrots and bats [156]. A core characteristic is the production of rhythmic, repetitive vocalizations derived from adult calls but produced out of context, often in prolonged and seemingly random sequences [157]. Based on this definition, Table 1 surveys babbling across species, with variable functions [156, 158].

Babbling is also reported in several species not known to exhibit VPL. These cases reveal a striking pattern: all documented non-VPL species that engage in babbling share extensive allomaternal care. This convergence suggests that babbling may serve yet another function, in addition to

Table 1: Taxa in which babbling has been recorded

Taxon	VPL	Rearing System
<i>Confirmed cases</i>		
Songbirds (subsong) [156, 160, 161]	yes	(mostly) biparental
Green-rumped parrotlet [162]	yes	biparental
Egyptian Fruit bat [163]	yes	mother-only
Greater sac-winged bat [164]	yes	mother-only
Gray mouse lemur [158]	no	allomaternal
Common marmoset [165]	no	cooperative
Pygmy marmoset [166, 167]	no	cooperative
Giant otter [168]	no	cooperative
<i>Suggested cases</i>		
Bottlenose dolphin [169]	yes	allomaternal
Beluga [170]	?	allomaternal
Naked mole rat [171]	no?	cooperative
Degu [172]	?	biparental
Norway rat [173]	no	allomaternal

practicing new sounds. It suggests that babbling is routinely used to elicit the attention of parents or caretakers [143, 159] and thereby to enhance social interaction in contingent vocal exchanges through feedback loops in both humans and nonhuman primates [150, 152, 156, 160?].

Given that allomaternal-only care is relatively rare among mammals, its consistent association with babbling is notable, even though the lack of reliable negative evidence precludes formal statistical evaluation. While the vocal-practice and care-eliciting functions are not mutually exclusive, the occurrence of babbling in many mammals lacking complex or variable vocal repertoires suggests that the care-eliciting function may be evolutionarily ancestral, primarily serving to attract the attention of selective caregivers.

Taken together, these patterns suggest that human babbling may not be functionally uniform across development, but instead comprises at least two partially overlapping phases. Early babbling appears to be primarily social in function, whereas later babbling increasingly reflects vocal practice. From this perspective, prelinguistic babbling is functionally analogous to that observed across prosocial species, particularly cooperative breeders such as marmosets, which are characterized by vocal accommodation and relatively limited, fixed vocal repertoires. By contrast, later linguistic babbling reflects an active form of vocal practice aimed at acquiring the native language, analogous to subsong in birds and perhaps other species exhibiting VPL.

This functional differentiation is consistent with the hypothesis that babbling in early hominins initially evolved to solicit care and was subsequently exapted for an additional role in language acquisition, while retaining its original function, particularly in its early form. Moreover, babbling independently contributed to the evolution of enhanced vocal control [174], thereby facilitating the emergence of vocal production learning. Notably, there is currently no evidence for babbling in great apes [175]. This absence suggests that babbling is a derived innovation in the human lineage, likely associated with shifts in child-rearing practices that distinguish humans from their closest relatives.

In conclusion, comparative evidence supports the idea that increasing demands for coopera-

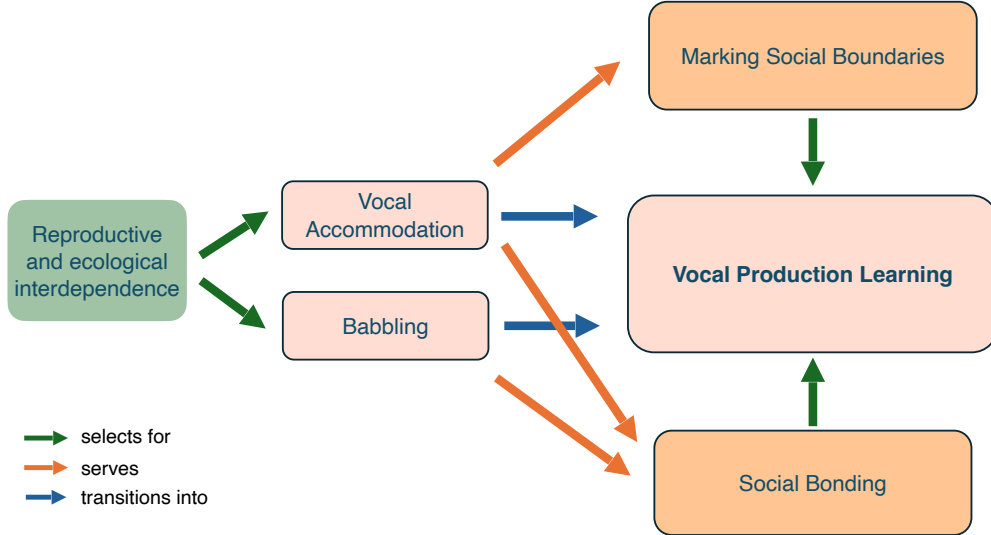


Figure 4: Proposed evolutionary scenario: Reproductive and ecological interdependence selectively favored vocal accommodation and babbling for their social functions. These in turn selectively favored the transition into fully-fledged vocal production learning.

tion, reflected in heightened ecological and reproductive interdependence, created social conditions conducive to the emergence of VPL and thus early language. In our lineage, cooperative breeding gave rise to changes in communication that fostered an interactional infrastructure, or “interaction engine” [146], which scaffolded the transformation of great-ape-like vocal communication, characterized by vocal accommodation, into VPL-based early forms of language. As illustrated in Figure 4, within this framework, babbling associated with allomaternal care likely provided an additional developmental scaffold, facilitating language learning as communicative systems became more complex [174], while also facilitating the emergence of the creative form-function mapping that characterizes human language.

The emergence of language

Language is sometimes assumed to be a relatively recent phenomenon, possibly even younger than our own species [176, 177]. Our framework departs from this view. In line with current genetic evidence [178], we adopt the perspective that language has an early evolutionary origin and emerged gradually over time. Within this framework, we propose that the first elements of language predate the split leading to anatomically modern humans, Neanderthals and Denisovans, typically dated to around 650 ka [179]. A plausible starting point lies much earlier, around 1.5 Ma, when hominin brains attained a more human-like size and organization [131], and Acheulean technology became widespread [180]. We interpret the archeological record as indicating that the widespread adoption of Acheulean toolmaking was facilitated by active demonstration and other forms of scaffolding, implying a degree of prosociality. This is consistent with evidence reviewed above suggesting that prosocial behaviors became increasingly prevalent from around ca 2 Ma, thereby creating conditions conducive to more complex communication.

Given the limited flexibility in vocal production and gestural repertoires observed in great apes [32, 36], we further suggest that early stages of increasingly prosocial communication relied heavily on ostensive and iconic gestures [181]. Over time, however, vocal signals, and eventually words,

came to predominate as they are more efficient: they free the hands, require no external props and function in low-visibility conditions. In this scenario, these conditions favored selection for increased vocal flexibility, enabling the transformation of vocal accommodation into VPL.

No other primates, not even highly cooperative species such as callitrichids, have evolved open-ended and locally variable vocal repertoires, let alone language. This raises the question: why did these capacities emerge uniquely in the hominin lineage? We suggest two non-exclusive explanations. First, the emergence of language likely depended on the specific brain architecture of hominins, characterized by enhanced control and decision-making, memory, and associative capacities, which enabled the transformation of expanded vocal repertoires into a productive and creative system. Second, hominin cooperation appears to have reached a level of complexity exceeding that of other cooperative species, including callitrichids and even highly social taxa such as African wild dogs, which engage in cooperative hunting and cooperative breeding [182]). In this context, language may have evolved to manage increasingly demanding social and ecological tasks, including the planning and coordination of foraging movements, extensive and delayed food sharing, teaching and allomaternal care. This explanation is consistent with archaeological evidence for an emerging hunting and gathering lifeway [119, 120, 183]. The growing importance of social transmission, for complex technological skills, foraging strategies, and social negotiation skills in multilevel societies [184], would have intensified selection for a more efficient communication. This pressure is particularly salient given the slow acquisition of learned skills in humans [185, 186]. Experimental evidence from modern humans further indicated that such transmission becomes more efficient when linguistic instruction is available [187].

In this light, the centrality of social transmission in humans is consistent with the evolution of pedagogy [188], which extends forms of active teaching observed in cooperative breeders [189] into the communication domain [190]. Together with babbling, this helps explain why human language evolved into a system that could rely to an unprecedented extent on learning.

VPL was likely established in *H. erectus*, giving rise to what could be considered proto-language [116]. Genetic evidence, however, suggests that VPL capacities in our lineage increased further after the split from the Neanderthals and Denisovans [132, 133, 191, 192] (Figure 3). We propose that this secondary increase reflects an evolutionary response to a novel cooperative challenge associated with rapid growing population sizes in *Homo sapiens* [99, 100, 193]: How to identify trustworthy partners for cooperation, particularly in contexts such as food sharing. One solution to this challenge may have been the introduction of group identity markers, which distinguish in-group from out-group strangers. In-group strangers can be trusted because any violations of trust would likely be transmitted through social networks via gossip, whereas cooperation with out-group individuals carries greater risk [100]. Communciation provides one of the most ubiquitous group identity markers [93, 194, 195], potentially driving the diversification of human communication system into thousands of dialects and languages [3]. In turn, this process may have selected for increasingly refined VPL abilities.

If this interpretation is correct, it suggests new directions for comparative research on the origin of language. In particular, future work could focus on species that face similar demands for communication in contexts of complex, coordination-intensive cooperation. Especially promising is the search for novel form-function mappings that emerge in cooperative contexts that develop into local traditions, potentially paralleling early stages of language evolution (see Figure 1), as been suggested in parrots [64, 78]. Identifying and systematically comparing such systems across taxa will be critical for testing the hypothesis that language emerged from the interaction between increasingly complex cooperation and the progressive expansion of vocal production learning capacities.

References

- [1] Langergraber, K. E. *et al.* Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. *Proceedings of the National Academy of Sciences* **109**, 15716–15721 (2012).
- [2] Fischer, J. & Hammerschmidt, K. Towards a new taxonomy of primate vocal production learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20190045 (2019).
- [3] Bickel, B., Giraud, A.-L., Zuberbühler, K. & Van Schaik, C. P. Language follows a distinct mode of extra-genomic evolution. *Physics of Life Reviews* **50**, 211–225 (2024).
- [4] Wilson, B. *et al.* Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature Communications* **6**, 8901 (2015).
- [5] Devaine, M. *et al.* Reading wild minds: A computational assay of Theory of Mind sophistication across seven primate species. *PLOS Computational Biology* **13**, e1005833 (2017).
- [6] Krupenye, C. & Call, J. Theory of mind in animals: Current and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science* **10**, e1503 (2019).
- [7] Bastos, A. P. M. & Krupenye, C. Evidence for representation of pretend objects by Kanzi, a language-trained bonobo. *Science* **391**, 583–586 (2026).
- [8] Seyfarth, R. M. & Cheney, D. L. Precursors to language: Social cognition and pragmatic inference in primates. *Psychonomic Bulletin & Review* **24**, 79–84 (2017).
- [9] Kato, A., Masataka, N., Koda, H. & Oyakawa, C. Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour* **144**, 681–692 (2007).
- [10] Hopkins, W. D., Taglialatela, J. P. & Leavens, D. A. Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour* **73**, 281–286 (2007).
- [11] Wich, S. A. & De Vries, H. Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society B: Biological Sciences* **273**, 735–740 (2006).
- [12] Slocombe, K. E. & Zuberbühler, K. Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences* **104**, 17228–17233 (2007).
- [13] Crockford, C., Wittig, R. M., Mundry, R. & Zuberbühler, K. Wild chimpanzees inform ignorant group members of danger. *Current Biology* **22**, 142–146 (2012).
- [14] Heesen, R., Austry, D. A., Upton, Z. & Clay, Z. Flexible signalling strategies by victims mediate post-conflict interactions in bonobos. *Philosophical Transactions of the Royal Society B* **377**, 20210310 (2022).
- [15] Whiten, A., McGuigan, N., Marshall-Pescini, S. & Hopper, L. M. Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 2417–2428 (2009).
- [16] Van Leeuwen, E. J. C., Cronin, K. A. & Haun, D. B. M. A group-specific arbitrary tradition in chimpanzees (*Pan troglodytes*). *Animal Cognition* **17**, 1421–1425 (2014).
- [17] Van Leeuwen, E. J. C. *et al.* Getting to the bottom of social learning: Chimpanzees copy arbitrary behavior from conspecifics (2024).
- [18] Chemla, E., Dautriche, I., Buccola, B. & Fagot, J. Constraints on the lexicons of human languages have cognitive roots present in baboons (*Papio papio*). *Proceedings of the National Academy of Sciences* **116**, 14926–14930 (2019).
- [19] Wilson, V. A. D. *et al.* Humans and great apes visually track event roles in similar ways. *PLOS Biology* **22**, e3002857 (2024).
- [20] Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K. & Bickel, B. A universal preference for animate agents in hominids. *iScience* 109996 (2024).
- [21] Brocard, S., Berton, C., Wilson, V., Bickel, B. & Zuberbühler, K. The perception of prosocial agents by chimpanzees and humans. *Royal Society Open Science* **12**, 250916 (2025).
- [22] Brocard, S., Voinov, P. V., Bickel, B. & Zuberbühler, K. Spontaneous Encoding of Event Roles in Hominids. *Open Mind* **9**, 559–575 (2025).
- [23] Leroux, M. *et al.* Call combinations and compositional processing in wild chimpanzees. *Nature Communications* **14**, 2225 (2023).

- [24] Berthet, M., Surbeck, M. & Townsend, S. W. Extensive compositionality in the vocal system of bonobos. *Science* **388**, 104–108 (2025).
- [25] Bosshard, A. B. *et al.* Beyond bigrams: call sequencing in the common marmoset (*Callithrix jacchus*) vocal system. *Royal Society Open Science* **11**, 240218 (2024).
- [26] Girard-Buttoz, C. *et al.* Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology* **5**, 410 (2022).
- [27] Girard-Buttoz, C. *et al.* Versatile use of chimpanzee call combinations promotes meaning expansion. *Science Advances* **11**, eadq2879 (2025).
- [28] Ouattara, K., Lemasson, A. & Zuberbühler, K. Campbell’s monkeys use affixation to alter call meaning. *PloS One* **4**, e7808 (2009).
- [29] Watson, S. K. *et al.* Nonadjacent dependency processing in monkeys, apes, and humans. *Science Advances* **6** (2020).
- [30] Jiang, X. *et al.* Production of supra-regular spatial sequences by macaque monkeys. *Current Biology* **28**, 1851 – 1859.e4 (2018).
- [31] Ferrigno, S., Cheyette, S. J., Piantadosi, S. T. & Cantlon, J. F. Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. *Science Advances* **6**, eaaz1002 (2020).
- [32] Wich, S. A. *et al.* Call cultures in orang-utans? *PLoS ONE* **7**, e36180 (2012).
- [33] Watts, D. P. Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: variation, social learning, and possible functions. *Primates* **57**, 61–72 (2016).
- [34] Robbins, M. M. *et al.* Behavioral variation in gorillas: Evidence of potential cultural traits. *PLOS ONE* **11**, e0160483 (2016).
- [35] Van Schaik, C. P. *et al.* Orangutan cultures and the evolution of material culture. *Science* **299**, 102–105 (2003).
- [36] Fröhlich, M. & Van Schaik, C. P. Must all signals be evolved? A proposal for a new classification of communicative acts. *WIREs Cognitive Science* **11**, e1527 (2020).
- [37] Janik, V. M. & Slater, P. J. Vocal Learning in mammals. In *Advances in the Study of Behavior*, vol. 26, 59–99 (Elsevier, 1997).
- [38] King, S. L. & McGregor, P. K. Vocal matching: the what, the why and the how. *Biology Letters* **12**, 20160666 (2016).
- [39] Pickering, M. J. & Garrod, S. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* **27** (2004).
- [40] Giles, H., Coupland, N. & Coupland, J. Accommodation theory: Communication, context, and consequence. In Giles, H., Coupland, J. & Coupland, N. (eds.) *Contexts of Accommodation*, 1–68 (Cambridge University Press, 1991).
- [41] Petkov, C. & Jarvis, E. Birds, primates and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience* **4**, 12 (2012).
- [42] Pardo, J. S., Pellegrino, E., Dellwo, V. & Möbius, B. Special issue: Vocal accommodation in speech communication. *Journal of Phonetics* **95**, 101196 (2022).
- [43] Ruch, H., Zürcher, Y. & Burkart, J. M. The function and mechanism of vocal accommodation in humans and other primates. *Biological Reviews* **93**, 996–1013 (2018).
- [44] Lazarsfeld, P. & Merton, R. K. Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society* **18**, 18–66 (1954).
- [45] McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: homophily in social networks. *Annual Review of Sociology* **27**, 415–444 (2001).
- [46] Arriaga, G. & Jarvis, E. D. Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language* **124**, 96–116 (2013).
- [47] Janik, V. M. & Knörnschild, M. Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200244 (2021).
- [48] Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F. & Hammerschmidt, K. Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20202531 (2020).
- [49] Tyack, P. L. A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B:*

- Biological Sciences* **375**, 20180406 (2020).
- [50] Martins, P. T. & Boeckx, C. Vocal learning: Beyond the continuum. *PLOS Biology* **18**, e3000672 (2020).
 - [51] Vernes, S. C. *et al.* The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200236 (2021).
 - [52] Fitch, W. T. Three meanings of “recursion”: key distinctions for biolinguistics. In Larson, R. K., Déprez, V. & Yamakido, H. (eds.) *The evolution of human language: biolinguistic perspectives*, 73–90 (Cambridge University Press, Cambridge, 2010).
 - [53] Tyack, P. L. Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology (Washington, D.C.: 1983)* **122**, 319–331 (2008).
 - [54] Fitch, W. The evolution of speech: a comparative review. *Trends in Cognitive Sciences* **4**, 258–267 (2000).
 - [55] Jarvis, E. D. Evolution of vocal learning and spoken language. *Science* **366**, 50–54 (2019).
 - [56] Ten Cate, C. Re-evaluating vocal production learning in non-oscine birds. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200249 (2021).
 - [57] Wirthlin, M. E. *et al.* Vocal learning–associated convergent evolution in mammalian proteins and regulatory elements. *Science* **383**, eabn3263 (2024).
 - [58] Scharff, C. & Petri, J. Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2124–2140 (2011).
 - [59] Fisher, S. E. Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review* **24**, 34–40 (2017).
 - [60] Casten, L. G. *et al.* Ancient regulatory evolution shapes individual language abilities in present-day humans. *Science Advances* **12**, ead5260 (2026).
 - [61] Garland, E. C. & McGregor, P. K. Cultural transmission, evolution, and revolution in vocal displays: Insights from bird and whale song. *Frontiers in Psychology* **11** (2020).
 - [62] Knörnschild, M. Vocal production learning in bats. *Current Opinion in Neurobiology* **28**, 80–85 (2014).
 - [63] Schusterman, R. J. Vocal learning in mammals with special emphasis on pinnipeds. In Oller, D. K. & Gribel, U. (eds.) *The evolution of communicative flexibility: complexity, creativity and adaptability in human and animal communication*, 41–70 (MIT Press, Cambridge, MA, 2008).
 - [64] Bradbury, J. W. & Balsby, T. J. S. The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology* **70**, 293–312 (2016).
 - [65] Ekström, A. G., Gannon, C., Edlund, J., Moran, S. & Lameira, A. R. Chimpanzee utterances refute purported missing links for novel vocalizations and syllabic speech. *Scientific Reports* **14**, 17135 (2024).
 - [66] Fitch, W. T., de Boer, B., Mathur, N. & Ghazanfar, A. A. Monkey vocal tracts are speech-ready. *Science Advances* **2** (2016).
 - [67] Lameira, A. R. Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews* **83**, 429–439 (2017).
 - [68] Mitani, J. & Gros-Louis, J. Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**, 1041–1064 (1998).
 - [69] Marshall, A. J., Wrangham, R. W. & Arcadi, A. C. Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* **58**, 825–830 (1999).
 - [70] Crockford, C., Herbinger, I., Vigilant, L. & Boesch, C. Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**, 221–243 (2004).
 - [71] Townsend, S. W., Deschner, T. & Zuberbühler, K. Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS ONE* **3**, e2431 (2008).
 - [72] Ghazanfar, A. A., Liao, D. A. & Takahashi, D. Y. Volition and learning in primate vocal behaviour. *Animal Behaviour* **151**, 239–247 (2019).
 - [73] Hage, S. R. & Nieder, A. Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications* **4**, 2409 (2013).
 - [74] Hage, S. R. & Nieder, A. Dual neural network model for the evolution of speech and language. *Trends*

- in *Neurosciences* **39**, 813–829 (2016).
- [75] Becker, Y. *et al.* Long arcuate fascicle in wild and captive chimpanzees as a potential structural precursor of the language network. *Nature Communications* **16**, 4485 (2025).
 - [76] Petanjek, Z. *et al.* Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences* **108**, 13281–13286 (2011).
 - [77] Cerrito, P., Gascon, E., Roberts, A. C., Sawiak, S. J. & Burkart, J. M. Neurodevelopmental timing and socio-cognitive development in a prosocial cooperatively breeding primate (*Callithrix jacchus*). *Science Advances* **10**, eado3486 (2024).
 - [78] Carouso-Peck, S., Goldstein, M. H. & Fitch, W. T. The many functions of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200235 (2021).
 - [79] Marler, P. Some ethological implications for neuroethology: the ontogeny of birdsong. In *Advances in Vertebrate Neuroethology*, 21–52 (Plenum Press, New York, 1983).
 - [80] Herman, L. M. *et al.* Humpback whale song: who sings? *Behavioral Ecology and Sociobiology* **67**, 1653–1663 (2013).
 - [81] Duengen, D. & Ravignani, A. The paradox of learned song in a semi-solitary mammal. *Ethology* **129**, 445–453 (2023).
 - [82] Dalziell, A. H., Welbergen, J. A., Iqic, B. & Magrath, R. D. Avian vocal mimicry: a unified conceptual framework. *Biological Reviews* **90**, 643–668 (2015).
 - [83] Sakata, J., Woolley, S., Fay, R. & Popper, A. *The Neuroethology of Birdsong*. (Springer International Publications, Cham, 2020).
 - [84] Darwin, C. *The descent of man, and selection in relation to sex*. (Murray, London, 1871).
 - [85] Wilbrecht, L. & Nottebohm, F. Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Reviews* **9**, 135–148 (2003).
 - [86] Parr, L. A. & De Waal, F. B. M. Visual kin recognition in chimpanzees. *Nature* **399**, 647–648 (1999).
 - [87] Nishimura, T. *et al.* Evolutionary loss of complexity in human vocal anatomy as an adaptation for speech. *Science* **377**, 760–763 (2022).
 - [88] Watson, S. K. *et al.* Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology* **25**, 495–499 (2015).
 - [89] Sandel, A. A. *et al.* Lethal conflict after group fission in wild chimpanzees. *Science* **392**, 216–220 (2026).
 - [90] Lemasson, A., Hausberger, M. & Zuberbühler, K. Socially Meaningful Vocal Plasticity in Adult Campbell’s Monkeys (*Cercopithecus campbelli*). *Journal of Comparative Psychology* **119**, 220–229 (2005).
 - [91] Zürcher, Y. & Burkart, J. M. Evidence for dialects in three captive populations of Common marmosets (*callithrix jacchus*). *International Journal of Primatology* **38**, 780–793 (2017).
 - [92] Bateson, G. Culture Contact and Schismogenesis. *Man* **35**, 178 (1935).
 - [93] Cohen, E. The evolution of tag-based cooperation in humans: The case for accent. *Current Anthropology* **53**, 588–616 (2012).
 - [94] Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J. & Pagel, M. Languages evolve in punctuational bursts. *Science* **319**, 588–588 (2008).
 - [95] Mansfield, J., Leslie-O’Neill, H. & Li, H. Dialect differences and linguistic divergence: A crosslinguistic survey of grammatical variation. *Language Dynamics and Change* **13**, 232–276 (2023).
 - [96] Graff, A. *et al.* Patterns of genetic admixture reveal similar rates of borrowing across diverse scenarios of language contact. *Science Advances* **11**, eadv7521 (2025).
 - [97] Stutz, A. J. A niche of their own: population dynamics, niche diversification, and biopolitics in the recent biocultural evolution of hunter-gatherers. *Journal of Anthropological Archaeology* **57**, 101120 (2020).
 - [98] Stiner, M. C., Munro, N. D., Surovell, T. A., Tchernov, E. & Bar-Yosef, O. Paleolithic Population Growth Pulses Evidenced by Small Animal Exploitation. *Science* **283**, 190–194 (1999).
 - [99] Foley, R. A. & Mirazón Lahr, M. The evolution of the diversity of cultures. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 1080–1089 (2011).
 - [100] Cerrito, P. *et al.* Pleistocene origins of cultural and linguistic diversification: How Homo sapiens and

- Neanderthals differed. *EcoEvoRxiv* (2025).
- [101] Baciadonna, L. *et al.* Vocal accommodation in penguins (*Spheniscus demersus*) as a result of social environment. *Proceedings of the Royal Society B: Biological Sciences* **289**, 20220626 (2022).
 - [102] Zürcher, Y., Willems, E. P. & Burkart, J. M. Trade-offs between vocal accommodation and individual recognisability in common marmoset vocalizations. *Scientific Reports* **11**, 15683 (2021).
 - [103] Phaniraj, N., Wierucka, K. & Burkart, J. M. Dynamic vocal learning in adult marmoset monkeys (2023).
 - [104] Wanker, R., Sugama, Y. & Prinage, S. Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. *Animal Behaviour* **70**, 111–118 (2005).
 - [105] Janik, V. M., Sayigh, L. S. & Wells, R. S. Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences* **103**, 8293–8297 (2006).
 - [106] Pardo, M. A. *et al.* African elephants address one another with individually specific name-like calls. *Nature Ecology & Evolution* **8**, 1353–1364 (2024).
 - [107] Oren, G. *et al.* Vocal labeling of others by nonhuman primates. *Science* **385**, 996–1003 (2024).
 - [108] Pardo, J. S. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* **119**, 2382–2393 (2006).
 - [109] Sagi, E. & Diermeier, D. Language Use and Coalition Formation in Multiparty Negotiations. *Cognitive Science* **41**, 259–271 (2017).
 - [110] Chartrand, T. L. & Lakin, J. L. The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology* **64**, 285–308 (2013).
 - [111] Lachmann, M. & Bergstrom, C. T. The disadvantage of combinatorial communication. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 2337–2343 (2004).
 - [112] Smith, E. A. Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior* **31**, 231–245 (2010).
 - [113] Burkart, J. M., Hrdy, S. B. & Van Schaik, C. P. Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews* **18**, 175–186 (2009).
 - [114] Zuberbühler, K. Cooperative breeding and the evolution of vocal flexibility. In Tallerman, M. & Gibson, K. R. (eds.) *The Oxford handbook of language evolution*, 72–81 (Oxford University Press, Oxford, 2012).
 - [115] Tomasello, M. *The Origins of Human Communication* (MIT Press, 2008).
 - [116] Fitch, W. T. *The Evolution of Language* (Cambridge University Press, Cambridge, 2010).
 - [117] Nowicki, S. & Searcy, W. A. The evolution of vocal learning. *Current Opinion in Neurobiology* **28**, 48–53 (2014).
 - [118] Sewall, K. B., Young, A. M. & Wright, T. F. Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour* **120**, 163–172 (2016).
 - [119] Hrdy, S. B. *Mothers and Others: The Evolutionary Origins of Mutual Understanding* (Harvard University Press, Cambridge, MA, 2009).
 - [120] Tomasello, M., Melis, A. P., Tennie, C., Wyman, E. & Herrmann, E. Two key steps in the evolution of human cooperation: The Interdependence Hypothesis. *Current Anthropology* **53**, 673–692 (2012).
 - [121] Potts, R. Environmental hypotheses of hominin evolution. *American Journal of Physical Anthropology* **107**, 93–136 (1998).
 - [122] Willems, E. P. & Van Schaik, C. P. The social organization of *Homo ergaster*: Inferences from anti-predator responses in extant primates. *Journal of Human Evolution* **109**, 11–21 (2017).
 - [123] Pobiner, B. L. The zooarchaeology and paleoecology of early hominin scavenging. *Evolutionary Anthropology: Issues, News, and Reviews* **29**, 68–82 (2020).
 - [124] Plummer, T. W. *et al.* Oldest Evidence of Toolmaking Hominins in a Grassland-Dominated Ecosystem. *PLoS ONE* **4**, e7199 (2009).
 - [125] Migliano, A. B. & Vinicius, L. The origins of human cumulative culture: from the foraging niche to collective intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377**, 20200317 (2022).
 - [126] Ringen, E. J., Duda, P. & Jaeggi, A. V. The evolution of daily food sharing: A Bayesian phylogenetic analysis. *Evolution and Human Behavior* **40**, 375–384 (2019).

- [127] Aiello, L. C. & Key, C. Energetic consequences of being a *Homo erectus* female. *American Journal of Human Biology* **14**, 551–565 (2002).
- [128] Yravedra, J., Rubio-Jara, S., Courtenay, L. A. & Martos, J. A. Mammal butchery by homo erectus at the lower pleistocene acheulean site of Juma’s korongo 2 (JK2), bed III, Olduvai Gorge, Tanzania. *Quaternary Science Reviews* **249**, 106612 (2020).
- [129] Pobiner, B. Meat-Eating Among the Earliest Humans. *American Scientist* **104**, 110 (2016).
- [130] Hill, K. & Hurtado, A. M. Cooperative breeding in South American hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences* **276**, 3863–3870 (2009).
- [131] Ponce De León, M. S. *et al.* The primitive brain of early *Homo*. *Science* **372**, 165–171 (2021).
- [132] Gunz, P. *et al.* Neandertal introgression sheds light on modern human endocranial globularity. *Current Biology* **29**, 120–127.e5 (2019).
- [133] Tajima, Y. *et al.* A humanized NOVA1 splicing factor alters mouse vocal communications. *Nature Communications* **16**, 1542 (2025).
- [134] Spikins, P., Rutherford, H. & Needham, A. From homininity to humanity: Compassion from the earliest archaics to modern humans. *Time and Mind* **3**, 303–325 (2010).
- [135] Haeusler, M. Spinal pathologies in fossil hominins. In Been, E., Gómez-Olivencia, A. & Ann Kramer, P. (eds.) *Spinal Evolution*, 213–245 (Springer International Publishing, Cham, 2019).
- [136] Isler, K. & Van Schaik, C. P. How our ancestors broke through the gray ceiling: comparative evidence for cooperative breeding in early *homo*. *Current Anthropology* **53**, S453–S465 (2012).
- [137] Burkart, J. M., Cerrito, P., Natalucci, G. & Van Schaik, C. P. Cooperative breeding as a likely early catalyst of human evolution. *Evolutionary Anthropology: Issues, News, and Reviews* **34**, e70016 (2025).
- [138] Deen, B., Schwiedrzik, C. M., Sliwa, J. & Freiwald, W. A. Specialized Networks for Social Cognition in the Primate Brain. *Annual Review of Neuroscience* **46**, 381–401 (2023).
- [139] Lukas, D. & Clutton-Brock, T. Climate and the distribution of cooperative breeding in mammals. *Royal Society Open Science* **4**, 160897 (2017).
- [140] Cornwallis, C. K. *et al.* Cooperation facilitates the colonization of harsh environments. *Nature Ecology & Evolution* **1**, 0057 (2017).
- [141] Griesser, M., Drobniak, S. M., Nakagawa, S. & Botero, C. A. Family living sets the stage for cooperative breeding and ecological resilience in birds. *PLOS Biology* **15**, e2000483 (2017).
- [142] Rapaport, L. G. & Brown, G. R. Social influences on foraging behavior in young nonhuman primates: Learning what, where, and how to eat. *Evolutionary Anthropology: Issues, News, and Reviews* **17**, 189–201 (2008).
- [143] Burkart, J. M. *et al.* A convergent interaction engine: vocal communication among marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377**, 20210098 (2022).
- [144] Sehner, S. *et al.* Problem-solving in groups of common marmosets (*Callithrix jacchus*): more than the sum of its parts. *PNAS Nexus* **1**, pgac168 (2022).
- [145] Levinson, S. C. Turn-taking in human communication – origins and implications for language processing. *Trends in Cognitive Sciences* **20**, 6–14 (2016).
- [146] Levinson, S. C. *The Interaction Engine: Language in Social Life and Human Evolution* (Cambridge University Press, Cambridge, 2025).
- [147] Leighton, G. M. Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171508 (2017).
- [148] Manser, M. B. *et al.* Vocal complexity in meerkats and other mongoose species. In *Advances in the Study of Behavior*, vol. 46, 281–310 (Elsevier, 2014).
- [149] Fitch, W. T. Sequence and hierarchy in vocal rhythms and phonology. *Annals of the New York Academy of Sciences* **1453**, 29–46 (2019).
- [150] Oller, D. K. *The Emergence of the Speech Capacity* (Psychology Press, 2000).
- [151] Elmlinger, S. L., Schwade, J. A. & Goldstein, M. H. The ecology of prelinguistic vocal learning: parents simplify the structure of their speech in response to babbling. *Journal of Child Language* **46**, 998–1011 (2019).
- [152] Goldstein, M. H. & Schwade, J. A. Social feedback to infants’ babbling facilitates rapid phonological learning. *Psychological Science* **19**, 515–523 (2008).

- [153] Sasaki, C. T., Levine, P. A., Laitman, J. T. & Crelin, E. S. Postnatal descent of the epiglottis in man: A preliminary report. *Archives of Otolaryngology - Head and Neck Surgery* **103**, 169–171 (1977).
- [154] Vorperian, H. K. & Kent, R. D. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* **50**, 1510–1545 (2007).
- [155] Boysson-Bardies, B. D., Sagart, L. & Durand, C. Discernible differences in the babbling of infants according to target language. *Journal of Child Language* **11**, 1–15 (1984).
- [156] Ter Haar, S. M. *et al.* Cross-species parallels in babbling: animals and algorithms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200239 (2021).
- [157] Fernandez, A. A., Burchardt, L. S., Nagy, M. & Knörnschild, M. Babbling in a vocal learning bat resembles human infant babbling. *Science* **373**, 923–926 (2021).
- [158] Langehennig-Peristenidou, A., Romero-Mujalli, D., Bergmann, T. & Scheumann, M. Features of animal babbling in the vocal ontogeny of the gray mouse lemur (*Microcebus murinus*). *Scientific Reports* **13**, 21384 (2023).
- [159] Snowdon, C. & Elowson, A. M. 'Babbling' in Pygmy marmosets: Development after infancy. *Behaviour* **138**, 1235–1248 (2001).
- [160] Goldstein, M. H., King, A. P. & West, M. J. Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences* **100**, 8030–8035 (2003).
- [161] Lipkind, D. *et al.* Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* **498**, 104–108 (2013).
- [162] Eggleston, R. *et al.* Vocal babbling in a wild parrot shows life history and endocrine affinities with human infants. *Proceedings of the Royal Society B: Biological Sciences* **289**, 20220592 (2022).
- [163] Prat, Y., Taub, M. & Yovel, Y. Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances* **1**, e1500019 (2015).
- [164] Knörnschild, M., Nagy, M., Metz, M., Mayer, F. & von Helversen, O. Complex vocal imitation during ontogeny in a bat. *Biology Letters* **6**, 156 (2009).
- [165] Pistorio, A. L., Vintch, B. & Wang, X. Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*). *The Journal of the Acoustical Society of America* **120**, 1655–1670 (2006).
- [166] Elowson, A. M., Snowdon, C. T. & Lazaro-Perea, C. 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends in Cognitive Sciences* **2**, 31–37 (1998).
- [167] Elowson, A. M., Snowdon, C. & Lazaro-Perea, C. Infant 'babbling' in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour* **135**, 643–664 (1998).
- [168] Mumm, C. A. S. & Knörnschild, M. The vocal repertoire of adult and neonate giant otters (*Pteronura brasiliensis*). *PLoS ONE* **9**, e112562 (2014).
- [169] McCowan, B. & Reiss, D. The fallacy of 'signature whistles' in bottlenose dolphins: a comparative perspective of 'signature information' in animal vocalizations. *Animal Behaviour* **62**, 1151–1162 (2001).
- [170] Vergara, V. & Barrett-Lennard, L. G. Vocal Development in a Beluga Calf (*Delphinapterus leucas*). *Aquatic Mammals* **34**, 123–143 (2008).
- [171] Buffenstein, R. Colony-specific dialects of naked mole-rats. *Science* **371**, 461–462 (2021).
- [172] Okanoya, K. Language, evolution, and an emergent property. *Current Opinion in Neurobiology* **17**, 1–6 (2007).
- [173] Wöhr, M. & Schwarting, R. K. W. Maternal care, isolation-induced infant ultrasonic calling, and their relations to adult anxiety-related behavior in the rat. *Behavioral Neuroscience* **122**, 310–330 (2008).
- [174] Hrdy, S. B. & Burkart, J. M. The emergence of emotionally modern humans: implications for language and learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20190499 (2020).
- [175] Wegdell, F. *et al.* The evolution of infant-directed communication: Comparing vocal input across all great apes. *Science Advances* **11**, eadt7718 (2025).
- [176] Klein, R. *The Human Career: Human Biological and Cultural Origins*. (University of Chicago Press, Chicago, 2009).
- [177] Bolhuis, J. J., Tattersall, I., Chomsky, N. & Berwick, R. C. How could language have evolved? *PLoS Biology* **12**, e1001934 (2014).

- [178] Casten, L. G. *et al.* Ancient regulatory evolution shapes individual language abilities in present-day humans (2026).
- [179] Dediu, D. & Levinson, S. C. On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* **4** (2013).
- [180] Lepre, C. J. *et al.* An earlier origin for the Acheulian. *Nature* **477**, 82–85 (2011).
- [181] Tomasello, M. *Origins of human communication* (MIT Press, Cambridge, MA, 2008).
- [182] Creel, S. & Creel, N. M. *The African wild dog: behavior, ecology, and conservation*. Monographs in behavior and ecology (Princeton University Press, Princeton, N.J, 2002).
- [183] Burkart, J. M., Hrdy, S. & Van Schaik, C. P. Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology* **18**, 175–186 (2009).
- [184] Migliano, A. B. & Vinicius, L. The origins of human cumulative culture: from the foraging niche to collective intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377**, 20200317 (2021).
- [185] Meulman, E. J. M., Seed, A. M. & Mann, J. If at first you don’t succeed... Studies of ontogeny shed light on the cognitive demands of habitual tool use. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20130050 (2013).
- [186] Koster, J. *et al.* The life history of human foraging: Cross-cultural and individual variation. *Science Advances* **6**, eaax9070 (2020).
- [187] Caldwell, C. A., Renner, E. & Atkinson, M. Human teaching and cumulative cultural evolution. *Review of Philosophy and Psychology* **9**, 751–770 (2018).
- [188] Csibra, G. & Gergely, G. Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 1149–1157 (2011).
- [189] Kline, M. A. How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences* **38**, e31 (2015).
- [190] Schick, J. *et al.* The function and evolution of child-directed communication. *PLOS Biology* **20**, e3001630 (2022).
- [191] Gokhman, D. *et al.* Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nature Communications* **11**, 1189 (2020).
- [192] Buisan, R., Moriano, J., Andirkó, A. & Boeckx, C. A brain region-specific expression profile for genes within large introgression deserts and under positive selection in homo sapiens. *Frontiers in Cell and Developmental Biology* **10**, 824740 (2022).
- [193] Cousins, T., Scally, A. & Durbin, R. A structured coalescent model reveals deep ancestral structure shared by all modern humans (2024).
- [194] Moffett, M. W. Human Identity and the Evolution of Societies. *Human Nature* **24**, 219–267 (2013).
- [195] Efferson, C., Lalive, R. & Fehr, E. The Coevolution of Cultural Groups and Ingroup Favoritism. *Science* **321**, 1844–1849 (2008).

Supporting Information for

How cooperation drove the evolution of vocal production learning in hominins

Sabine Stoll, Carel P. van Schaik, Konatsu Ono, Nikhil Phaniraj, Zitan Song, Judith Burkart & Balthasar Bickel

Table S1: Evidence for functions across taxa

Function	Mechanism	Taxa [references]
Mate attraction	VPL	Australian musk duck [1] Hummingbirds [1] Songbirds [2–4] Whales [5–9] Pinnipeds [10–13] Bats [14, 15] Mice [16, 17]
Parent-offspring recognition	VPL Accommodation	Penguins [18, 19] Pinnipeds [13, 20] Dolphins [21] Bats [22–25]
Marking social boundaries	Accommodation	Hummingbirds [26] Parrots [27, 28] Songbirds [29–32] Wolves [33] Whales [34, 35] Bats [36, 37] Naked-mole rat [38] Nonhuman-primates [39–42] Homo sapiens
Servicing cooperative bonds	VPL Accommodation	Parrots [27, 43, 44] Songbirds [32, 45] Bats [22, 46, 47] Elephants [48, 49] Dolphins [50, 51] Whales [52] Goats [53] Non-human primates [54, 55] Homo sapiens
Creative form/meaning mapping	VPL	Parrots [27, 56] Orangutans [57] Homo sapiens

References

- [1] Ten Cate, C. Re-evaluating vocal production learning in non-oscine birds. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200249 (2021).
- [2] Catchpole, C., Slater, P. J. B. & Catchpole, C. *Bird Song: Biological Themes and Variations* (Cambridge Univ. Press, Cambridge, 2003).
- [3] Marler, P. Some ethological implications for neuroethology: the ontogeny of birdsong. In *Advances in Vertebrate Neuroethology*, 21–52 (Plenum Press, New York, 1983).
- [4] Kroodsma, D. E. & Byers, B. E. The Function(s) of Bird Song. *American Zoologist* **31**, 318–328 (1991).
- [5] Stafford, K. M. Singing behavior in the bowhead whale. In Clark, C. W. & Garland, E. C. (eds.) *Ethology and Behavioral Ecology of Mysticetes*, 277–295 (Springer International Publishing, Cham, 2022).
- [6] Herman, L. M. The multiple functions of male song within the humpback whale (*Megaptera novaeangliae*) mating system: review, evaluation, and synthesis. *Biological Reviews* **92**, 1795–1818 (2017).

- [7] Oleson, E. *et al.* Behavioral context of call production by eastern North Pacific blue whales. *Marine Ecology Progress Series* **330**, 269–284 (2007).
- [8] Garland, E. C. & McGregor, P. K. Cultural transmission, evolution, and revolution in vocal displays: Insights from bird and whale song. *Frontiers in Psychology* **11** (2020).
- [9] Garland, E. C. *et al.* Population structure of humpback whales in the western and central South Pacific Ocean as determined by vocal exchange among populations. *Conservation Biology* **29**, 1198–1207 (2015).
- [10] Schusterman, R. J. Vocal learning in mammals with special emphasis on pinnipeds. In Oller, D. K. & Gribel, U. (eds.) *The evolution of communicative flexibility: complexity, creativity and adaptability in human and animal communication*, 41–70 (MIT Press, Cambridge, MA, 2008).
- [11] Sjare, B. & Stirling, I. The breeding behavior of Atlantic walruses, *Odobenus rosmarus rosmarus*, in the Canadian High Arctic. *Canadian Journal of Zoology* **74**, 897–911 (1996).
- [12] Duengen, D. & Ravignani, A. The paradox of learned song in a semi-solitary mammal. *Ethology* **129**, 445–453 (2023).
- [13] Insley, S., Phillips, A. V. & Charrier, I. A review of social recognition in pinnipeds. *Aquatic Mammals* **29**, 181–201 (2003).
- [14] Knörnschild, M., Blüml, S., Steidl, P., Eckenweber, M. & Nagy, M. Bat songs as acoustic beacons - male territorial songs attract dispersing females. *Scientific Reports* **7**, 13918 (2017).
- [15] Fernandez, A. A., Burchardt, L. S., Nagy, M. & Knörnschild, M. Babbling in a vocal learning bat resembles human infant babbling. *Science* **373**, 923–926 (2021).
- [16] Arriaga, G., Zhou, E. P. & Jarvis, E. D. Of mice, birds, and men: The mouse ultrasonic song system has some features similar to humans and song-Learning birds. *PLOS ONE* **7**, e46610 (2012).
- [17] Arriaga, G. & Jarvis, E. D. Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language* **124**, 96–116 (2013).
- [18] Jouventin, P., Aubin, T. & Lengagne, T. Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Animal Behaviour* **57**, 1175–1183 (1999).
- [19] Baciadonna, L. *et al.* Vocal accommodation in penguins (*Spheniscus demersus*) as a result of social environment. *Proceedings of the Royal Society B: Biological Sciences* **289**, 20220626 (2022).
- [20] Charrier, I. Mother–offspring vocal recognition and social system in pinnipeds. In Aubin, T. & Mathevon, N. (eds.) *Coding Strategies in Vertebrate Acoustic Communication*, vol. 7, 231–246 (Springer International Publishing, Cham, 2020).
- [21] Sayigh, L. S. *et al.* Bottlenose dolphin mothers modify signature whistles in the presence of their own calves.
- [22] Esser, K. & Schmidt, U. Mother-infant communication in the lesser spear-nosed bat *phylostomus discolor* (chiroptera, phyllostomidae) — evidence for acoustic learning. *Ethology* **82**, 156–168 (1989).
- [23] Lattenkamp, E. Z., Vernes, S. C. & Wiegrefe, L. Vocal production learning in the pale spear-nosed bat, *Phyllostomus discolor*. *Biology Letters* **16**, 20190928 (2020).
- [24] Balcombe, J. P. Vocal recognition of pups by mother Mexican free-tailed bats, *Tadarida brasiliensis mexicana*. *Animal Behaviour* **39**, 960–966 (1990).
- [25] Knörnschild, M., Feifel, M. & Kalko, E. K. Mother–offspring recognition in the bat *Carollia perspicillata*. *Animal Behaviour* **86**, 941–948 (2013).
- [26] Araya-Salas, M. *et al.* Social group signatures in hummingbird displays provide evidence of co-occurrence of vocal and visual learning. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20190666 (2019).
- [27] Bradbury, J. W. & Balsby, T. J. S. The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology* **70**, 293–312 (2016).
- [28] Wright, T. F. & Dahlin, C. R. Vocal dialects in parrots: patterns and processes of cultural evolution. *Emu - Austral Ornithology* **118**, 50–66 (2018).
- [29] Walcott, C., Mager, J. N. & Piper, W. Changing territories, changing tunes: male loons, *Gavia immer*, change their vocalizations when they change territories. *Animal Behaviour* **71**, 673–683 (2006).
- [30] Loning, H., Verkade, L., Griffith, S. C. & Naguib, M. The social role of song in wild zebra finches. *Current Biology* **33**, 372–380.e3 (2023).
- [31] Podos, J. & Warren, P. S. The evolution of geographic variation in birdsong. In *Advances in the Study of Behavior*, vol. 37, 403–458 (Elsevier, 2007).
- [32] Moran, I. G. *et al.* Vocal convergence and social proximity shape the calls of the most basal Passeriformes, New Zealand Wrens. *Communications Biology* **7**, 575 (2024).
- [33] Zaccaroni, M. *et al.* Group specific vocal signature in free-ranging wolf packs. *Ethology Ecology & Evolution* **24**, 322–331 (2012).
- [34] Deecke, V., Ford, J. & Spong, P. Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Animal Behaviour* **60**, 629–638 (2000).
- [35] Rendell, L. E. & Whitehead, H. Vocal clans in sperm whales (*Physeter macrocephalus*). *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 225–231 (2003).
- [36] Knörnschild, M., Nagy, M., Metz, M., Mayer, F. & von Helversen, O. Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Animal Behaviour* **84**, 761–769 (2012).
- [37] Boughman, J. W. Greater spear-nosed bats give group-distinctive calls. *Behavioral Ecology and Sociobiology* **40**, 61–70 (1997).
- [38] Barker, A. J. *et al.* Cultural transmission of vocal dialect in the naked mole-rat. *Science* **371**, 503–507 (2021).
- [39] Mathevon, N., Keenan, S., Stevens, J. M. G., Zuberbühler, K. & Levréro, F. Vocal accommodation in bonobos. *Animal Behaviour* **219**, 123014 (2025).
- [40] Crockford, C., Herbinger, I., Vigilant, L. & Boesch, C. Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**, 221–243 (2004).
- [41] Marshall, A. J., Wrangham, R. W. & Arcadi, A. C. Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* **58**, 825–830 (1999).
- [42] Masataka, N. & Fujita, K. Vocal learning of Japanese and rhesus monkeys. *Behaviour* **109**, 191–199 (1989).
- [43] Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R. & Bradbury, J. W. Vertical transmission of learned signatures in a wild parrot. *Proceedings of the Royal Society B: Biological Sciences* **279**, 585–591 (2012).

- [44] Wanker, R., Sugama, Y. & Prinage, S. Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. *Animal Behaviour* **70**, 111–118 (2005).
- [45] Trainer, J. M., McDonald, D. B. & Learn, W. A. The development of coordinated singing in cooperatively displaying long-tailed manakins. *Behavioral Ecology* **13**, 65–69 (2002).
- [46] Prat, Y., Taub, M. & Yovel, Y. Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances* **1**, e1500019 (2015).
- [47] Boughman, J. W. Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265**, 227–233 (1998).
- [48] Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S. & Watwood, S. Elephants are capable of vocal learning. *Nature* **434**, 455–456 (2005).
- [49] Stoeger, A. S. *et al.* An Asian elephant imitates human speech. *Current Biology* **22**, 2144–2148 (2012).
- [50] King, S. L. & McGregor, P. K. Vocal matching: the what, the why and the how. *Biology Letters* **12**, 20160666 (2016).
- [51] Sayigh, L. S. *et al.* Bottlenose dolphin mothers modify signature whistles in the presence of their own calves. *Proceedings of the National Academy of Sciences* **120**, e2300262120 (2023).
- [52] Panova, E. M. & Agafonov, A. V. A beluga whale socialized with bottlenose dolphins imitates their whistles. *Animal Cognition* **20**, 1153–1160 (2017).
- [53] Briefer, E. F. & McElligott, A. G. Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Animal Behaviour* **83**, 991–1000 (2012).
- [54] Ruch, H., Zürcher, Y. & Burkart, J. M. The function and mechanism of vocal accommodation in humans and other primates. *Biological Reviews* **93**, 996–1013 (2018).
- [55] Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F. & Hammerschmidt, K. Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20202531 (2020).
- [56] Pepperberg, I. M. Cognitive and communicative abilities of Grey parrots. *Applied Animal Behaviour Science* **100**, 77–86 (2006).
- [57] Wich, S. A. *et al.* Call cultures in orang-utans? *PLOS ONE* **7**, e36180 (2012).