**Developmental Science** WILEY

# Slow development of vocal sequences through ontogeny in wild chimpanzees (*Pan troglodytes verus*)

Tatiana Bortolato[1,2,3] | Roger Mundry[4,5,6] | Roman M. Wittig[1,2,3] | Cédric Girard-Buttoz[1,2,3] | Catherine Crockford[1,2,3]

[1]The Ape Social Mind Lab, Institut des Sciences Cognitives, CNRS, Bron, France

[2]Department of Human Behaviour, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[3]Taï Chimpanzee Project, Centre Swiss des Recherches Scientifiques, Abidjan, Ivory Coast

[4]Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany

[5]Department for Primate Cognition, Georg-August-University, Göttingen, Germany

[6]Leibniz Science Campus Primate Cognition, Göttingen, Germany

**Correspondence**
Tatiana Bortolato and Catherine Crockford, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.
Email: tatiana_bortolato@eva.mpg.de and crockford@eva.mpg.de

**Abstract**

The development of the unique, hierarchical, and endless combinatorial capacity in a human language requires neural maturation and learning through childhood. Compared with most non-human primates, where combinatorial capacity seems limited, chimpanzees present a complex vocal system comprising hundreds of vocal sequences. We investigated how such a complex vocal system develops and the processes involved. We recorded 10,929 vocal utterances of 98 wild chimpanzees aged 0–55 years, from Taï National Park, Ivory Coast. We developed customized Generalized non-Linear Models to estimate the ontogenetic trajectory of four structural components of vocal complexity: utterance length, diversity, probability of panting (requiring phonation across inhalation and exhalation), and probability of producing two adjacent panted units. We found chimpanzees need 10 years to reach adult levels of vocal complexity. In three variables, the steepest increase coincided with the age of first non-kin social interactions (2–5 years), and plateaued in sub-adults (8–10 years), as individuals integrate into adult social life. Producing two adjacent panted units may require more neuromuscular coordination of the articulators, as its emergence and steepest increase appear later in development. These results suggest prolonged maturational processes beyond those hitherto thought likely in species that do not learn their vocal repertoire. Our results suggest that multifaceted ontogenetic processes drive increases in vocal structural complexity in chimpanzees, particularly increases in social complexity and neuro-muscular maturation. As humans live in a complex social world, empirical support for the "social complexity hypothesis" may have relevance for theories of language evolution.

KEYWORDS
communication, language, non-linear methods, primates, social complexity, vocal complexity

---

Cédric Girard-Buttoz and Catherine Crockford contributed equally to this work.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**Research Highlights**

- Chimpanzees need around 10 years to develop the vocal structural complexity present in the adult repertoire, way beyond the age of emergence of every single vocal unit.
- Multifaceted ontogenetic processes may drive increases in vocal structural complexity in chimpanzees, particularly increases in social complexity and neuro-muscular maturation.
- Non-linear increases in vocal complexity coincide with social developmental milestones.
- Vocal sequences requiring rapid articulatory change emerge later than other vocal sequences, suggesting neuro-muscular maturational processes continue through the juvenile years.

## 1 | INTRODUCTION

The uniqueness of human language has captured the attention of scientists since the time of Darwin (Darwin, 1871). Despite copious research, its evolutionary origin remains poorly understood (Friederici, 2017; Zuberbühler, 2019). A promising research avenue compares human and non-human animals (hereafter "animals") communication systems (Christiansen & Kirby, 2003; Russell & Townsend, 2017; Zuberbühler, 2019). A central aspect of human language is its virtually unlimited and hierarchical combinatorial capacity using a limited set of sounds. Language emerges slowly across protracted development, in parallel with neural maturation and learning processes (Locke & Snow, 1997; Skeide & Friederici, 2016). Children utter their first words around 10–12 months of age (Schneider et al., 2015). Initially, their vocabulary expands slowly, but expansion accelerates in the second year as the first word combinations appear. However, it is not until the school years that children process complex syntax in an adult-like form (Skeide et al., 2016).

Many animals also combine distinct vocal units into sequences (hyraxes, Kershenbaum et al., 2012; whales, Payne, 2000; birds, Catchpole & Slater, 2003; bats, Bohn et al., 2013; non-human primates (hereafter "primates"), Clarke et al., 2006; Girard-Buttoz, Zaccarella et al., 2022; Ham et al., 2016; Hedwig et al., 2014; Marler & Mitani, 1989; Ouattara et al., 2009; Schamberg et al., 2016; Suzuki & Zuberbühler, 2019). However, little is known about the emergence of vocal sequences in social animals with protracted development, such as whales, primates, or elephants (Connor et al., 1998).

Chimpanzees are a good model species to study vocal development, as they are one of our closest living relatives and adult chimpanzees emit hundreds of different vocal sequences containing up to 10 vocal units and comprising most of the single vocal units present in the vocal repertoire (Girard-Buttoz, Zaccarella et al., 2022). This contrasts with the patterns observed in many animals, including other primates, where sequences often contain only two units and are confined to limited parts of the vocal repertoire, such as song, or alarm calls (reviewed in Girard-Buttoz, Zaccarella et al., 2022). In terms of meaning, previous research provides observational and experimental evidence that most vocal units in the chimpanzee repertoire are highly context-specific across diverse contexts, including food, greeting, alarm, and hunt (reviewed in Crockford, 2019). However, with only 12 distinct vocal units, the diversity of meanings communicated using these single units is rather limited. In contrast, combining single units (potential meaning-bearing units) into vocal sequences, many of which are highly structured and follow adjacency rules in call ordering (Girard-Buttoz, Zaccarella et al., 2022), may expand meaning generation potential well beyond the scope of single vocal units.

As such, examining the structural complexity of a vocal repertoire is an important first step in understanding meaning generation potential. Thus, vocal structural complexity in chimpanzees, if accompanied by functional complexity, suggests the presence of a mechanism that can potentially defeat the constraints imposed by small vocal repertoires, moving toward a more open-ended repertoire, as has occurred in humans. Given that extensive combinatorial capacity is one of the key features of human language, examining the emergence, structure, and usage of sequences, through ontogeny in animals, will shed light on the processes involved in the acquisition of vocal complexity in animals, which may be of relevance for understanding the evolution of the developmental processes present in human language.

The aim of this study is to analyze the development of vocal structural complexity in chimpanzees and investigate which processes are involved. We propose four hypotheses whereby we contrast a (1) *limited maturation process* with three non-mutually exclusive maturational processes, (2) *sexual selection*, (3) *neuro-muscular maturation*, and/or (4) *social-complexity* (below we elaborate on these processes further). These four different processes would lead to some differences in ontogenetic trajectories. Specifically, the age at which chimpanzees show the fastest development of vocal complexity (i.e., the inflection point in Figure 1a) and the age at which they reach the adult levels of complexity (i.e., the asymptote in Figure 1a) will differ depending on which of the four mechanisms is in place (Figure 1b-e).
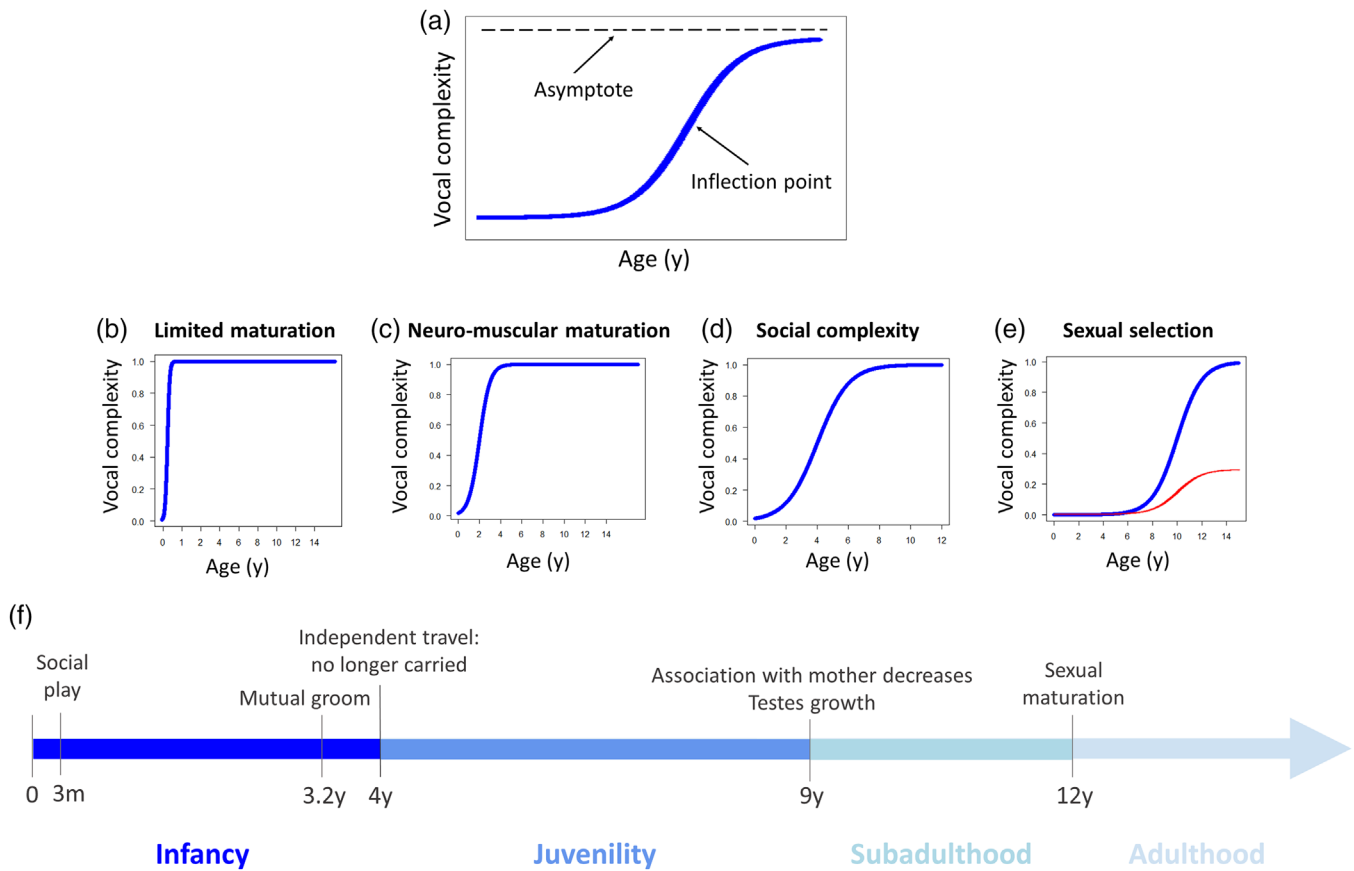
**FIGURE 1** Illustration of the predictions (emergence, inflection point and asymptote) according to each hypothesis regarding the development of vocal complexity in chimpanzees. The developmental periods from birth to adulthood are differentiated by a color gradient and indicate milestones in line with our four hypotheses.

The first possibility is that only *limited maturational processes* are required to shape vocal complexity. This is considered to be the case for some primate vocal repertoires that consist largely of single units, such that most call types are emitted soon after birth (Ey et al., 2007; Hammerschmidt et al., 2001). If maturational processes also play little role in the emission of vocal sequences, we expect little difference in the age of emergence of single vocal units and vocal sequences. Here, maximal vocal structural complexity would be reached at a similar time as the emergence of all single units (Figure 1b).

The second possibility is that vocal complexity is a *sexually selected trait* (Andersson & Iwasa, 1996), whereby the sex competing for the most over access to mates should present the most complex vocal system, and use such a system to attract mates and warn male rivals in territorial defense. This is the case, for instance, in humpback whale songs (*Megaptera novaeangliae*) (Janik, 2014), where only adult males perform the full song. In passerine birds, adult males commonly present more complex songs and larger vocal repertoires than females, in order to attract mates (e.g., sedge warblers, *Acrocephalus schoenobaenus;* red-winged blackbirds, *Agelaius phoeniceus;* Catchpole, 1987; Catchpole & Slater, 2003; but see; Baptista et al., 1993; Langmore, 2000). Within primates, sexual selection has been suggested as explanatory in the song of gibbons (Hylobates sp.; Cowlishaw, 1996; Geissmann & Orgeldinger, 2000), indris (*Indri indri;* De Gregorio et al., 2019; Gamba

et al., 2016), and calls of gelada baboons (*Theropithecus gelada*; Gustison & Bergman, 2016). The latter study shows that females exhibit a stronger response to vocal sequences of males when the sequences include derived complex calls. If sexual selection is the main driver of vocal complexity in chimpanzees, we expect vocal complexity to emerge, as in singing species, only with the onset of reproductive maturation, and to be more pronounced in males, the more competitive sex. We expect vocal complexity to increase during the beginning of the sub-adult period when testosterone levels are increasing (Behringer et al., 2014) and secondary sexual characteristics emerge such as testes (9–10 years, Goodall, 1986; Pusey, 1990) and permanent canines (10–13 years, Zihlman et al., 2007), reaching adult levels of complexity at sexual maturity (12 years) (Figure 1e-f).

A third possibility is that *neuro-muscular maturational processes* may only allow production of articulatory demanding vocal sequences later in ontogeny. Studies in mammals suggest a strong link between brain development and motor development (Garwicz et al., 2009). In macaques, neurological changes improve muscular control over the vocal tract, increasing the coordination of articulators and the respiratory system (Fitch et al., 2016). Therefore, the production of certain complex utterances in chimpanzees might only emerge after the maturation of the required neurological and muscular processes related to the vocal tract. For instance, in addition to combining single units

into vocal sequences, chimpanzees emit panted units which involve phonation and articulatory changes of lips and jaw (Grawunder et al., 2022) across rapid and alternating inhalations and exhalations (Crockford & Boesch, 2005), occurring at rates of 10–15 times faster than their resting breathing rates (Hewitt et al., 2002). This requires coordination of the respiratory system (sublaryngeal) and supralaryngeal muscles and articulators, such as the lips and the jaw (Figure S4), while maintaining vocal cord adduction. An additional neuro-muscular challenge may involve producing sequences of different adjacent panted units, for example, the production of a sequence combining panted hoos, which require lip protrusion, followed by panted screams which require lip retraction (Grawunder et al., 2022), as occurs in the common pant-hoot sequence (Arcadi, 1996). In sum, vocal emission may be limited by the maturation of the neuro-motor control of the muscles that move the articulators involved in vocal production. Thereby, we expect utterances requiring greater coordination of sublaryngeal, laryngeal, and/or supralaryngeal systems will emerge later in development, as such single units should emerge first (e.g., hoos or screams), followed by panted units (e.g., panted hoos or panted screams), followed by adjacent panted units (e.g., panted hoos + panted screams). Precisely, we expect adult levels of vocal structural complexity by weaning age (4 years, Lonsdorf et al., 2020) when chimpanzees reach 95%–99% of their adult cranial capacity (Zihlman et al., 2007) and coordination of articulations and muscles for other activities such as feeding or travelling are in place (Figure 1c-f).

Finally, a fourth possible process is that vocal complexity is related to changes in *social complexity*, whereby species or populations living in more complex social groups should require a more complex signalling system (e.g., a greater number of unique vocalization types) (Blumstein & Armitage, 1997; Freeberg, 2006; Freeberg et al., 2012; Manser et al., 2014; McComb & Semple, 2005). This mechanism, which is thought to operate across species or populations (Blumstein & Armitage, 1997; Freeberg, 2006; Manser et al., 2014; McComb & Semple, 2005), could also operate at a developmental level. More specifically, for species which are largely constrained and inflexible in the single calls they can produce, like most primates, including chimpanzees (Hammerschmidt & Fischer, 2008), combining existing vocal units into sequences could increase the diversity of the repertoire, thereby expanding the messaging potential, as the social world becomes more complex. In fact, while the repertoire of single call types is relatively universal across chimpanzee populations (Crockford, 2019), recent evidence suggest flexibility in the order in which single calls are combined in vocal sequences (Girard-Buttoz, Bortolato et al., 2022).

There are two key ages in chimpanzee development where social context expands (Figure 1f). The first one is at weaning age where chimpanzees are no longer being carried by the mothers (Van Lawick-Goodall, 1973). This likely results in a dramatic social shift requiring more independent engagement with adults of the community, such as vocalizing during greeting, aggression, feeding and alarm contexts, each with context-specific vocalisations (Crockford & Boesch, 2003b; Dezecache et al., 2019; Fedurek et al., 2021; Laporte & Zuberbühler, 2011; Slocombe & Zuberbühler, 2005a, 2005b). The second relevant developmental phase is at sub-adulthood (9 years, Pusey, 1990), where association time with the mother starts to drop (Pusey, 1983; Reddy &

Sandel, 2020) as offspring become independent socially active members (Pusey, 1990) of their community. For instance, social grunts of juveniles but not of subadults differ from those of adults (Laporte & Zuberbühler, 2011). Accordingly, we predict that, if social complexity drives vocal complexity through development, vocal repertoire expansion will coincide with major social developmental milestones. Particularly, we expect vocal complexity to be low at birth, increase during infancy along with the first social interactions, show the steepest increase at around weaning age accompanying the expansion of social opportunities, to reach the adult levels during the sub-adult period when they are finally integrating into the adult community (Figure 1d).

Whilst each process generates somewhat different predictions, it is possible that vocal complexity requires a combination of different processes, as occurs in humans (Kuhl, 2010; Skeide & Friederici, 2016), or even that two or more of these processes influence each other. Thus, vocal complexity might not be commensurate with any one of the just described patterns but may show a mix of several.

To assess the developmental trajectories of the vocal structural complexity, we used customized Generalized non-Linear Models, allowing us to identify the age of the inflection point and the age when adult capacities are reached. We focus on two key structural characteristics of the vocal combinatorial complexity of chimpanzees which both contribute to the potential to generate new meanings: (a) the length of the utterances (i.e., the number of different vocal units within an utterance, such as hoos + grunts + panted barks = 3), and (b) the diversity of the utterances (or "repertoire size", i.e., number of unique utterances) produced per individual. In addition, we examine two variables that capture articulatory complexity and hence assess neuro-muscular processes: (c) the probability of an utterance to comprise a panted unit (i.e., utterances with at least one panted unit, such as panted hoos), and (d) the probability of an utterance to comprise two different and adjacent panted units (such as panted hoos followed by panted grunts).

## 2 | METHODS

### 2.1 | Study group and data collection

For this study, TB collected 1553.2 h (1044.3 focal hours + 508.9 ad libitum hours; Altmann, 1974) of vocal data from 98 fully habituated wild chimpanzees (Mean ± SD = 14.82 ± 7.99 focal hours per individual, Figure S9) from three communities (East, North, and South) living in the Taï National Park, Côte d'Ivoire (5°45′N, 7°07′W) within the study area of the Taï Chimpanzee Project (Wittig, 2018). TB collected vocal data during two study periods: January to May 2019 and December 2019 to March 2020. We determined the age of each individual at the time of sampling based on the long-term database (see Supplementary material). TB sampled each individual over a 2-month period each season. The individuals sampled for this study ranged in age from newborn to 55 years old (see Supplementary material). TB followed the chimpanzees from dawn to dusk. Each observation day, TB conducted a focal follow (Altmann, 1974) of two different individuals for ca. 6 h each. During each focal follow, TB audio-recorded all occurrences of

vocalizations emitted by the focal and ad-libitum vocalizations emitted by identified individuals around the focal (Altmann, 1974) using a Sennheiser ME67 directional microphone connected to a Tascam DR-40X digital recorder (digitized at a 48 kHz sampling rate and 24-bit sampling depth). TB also collected social behavioral data on submissive greetings to determine the dominance rank of each individual using a smartphone and Cybertracker software (Steventon, 2002).

## 2.2 | Processing and analysis of recordings

TB recorded 11,164 vocal utterances across all ages. We examined each recording using PRAAT software (Boersma & Weenink, 2009), which shows spectrograms with the frequency distribution of the vocalization over time. Chimpanzee vocalizations exhibit distinctive acoustic and visual features, which allow differentiating them based on the spectrogram (Table S1 and Appendix show a repertoire classification used to train blind coders, including a detailed description, spectrograms, and sound files of each vocal unit). We divided the repertoire into 13 vocal units: six simple forms—barks "BK," grunts "GR," hoos "HO," non-vocal sounds "NV," screams "SC," and whimpers "WH"—(Figure S1) and seven panted forms—pants "PN," panted barks "PB," panted grunts "PG," panted hoos "PH," panted roars "PR," panted screams "PS," and panted whimpers "PW" (Figure S2). Panted whimpers were present only in the repertoire of immature individuals, thus, the vocal repertoire of adults consists of 12 vocal units. Vocal utterances can be classified as single units when comprising only one type of these 13 vocal units. Particularly, we defined a single unit to be either a vocal unit (or call type) produced individually or a repetition of the same vocal unit, emitted with less than 2 s pause in between. Different vocal units can also be combined in sequences of different length (Girard-Buttoz, Zaccarella et al., 2022). We defined a sequence as the combination of at least two different vocal units emitted within less than 1-s of each other (Figure S3). The chimpanzee repertoire is highly graded. For instance, a vocalization can start with grunts and gradually turn into barks. In those cases, the intermediate units in between the clearly defined grunts and barks were not coded as a different unit. We acknowledge that this might under-represent the variation present in the repertoire. A description of the total number of utterances recorded and the proportion of single units versus sequences and panted versus non-panted utterances per age category is shown in Table S2. From a randomly chosen sample of 314 sequences (10% of all sequences recorded), we measured the interval between adjacent units. We found that the average interval duration was $0.23 \pm 0.04$ s (mean $\pm$ SE), being considerably below the 1-s rule we used. This suggests chimpanzees are capable of producing different combined sounds in rapid succession, as humans do. We determined inter-rater reliability for 301 random utterances between TB, CC, and a third blind coder; and found an agreement of 94.6% across all the utterances. We did not differentiate between variants of the same vocal unit (e.g., "rest" or "alert" hoos in Crockford et al., 2018). We included in the analysis only recordings of good quality, with all the vocal units identified which were recorded from the beginning to the end of the utterance and

with caller ID clearly defined. We did not include in our analysis utterances with unclear units (213 utterances). This deletion did not affect our analysis since the unidentified sequences were equally distributed across all ages. Our final dataset comprised 10,929 utterances (73.9 average utterances per individual and season, with a range of 5–265 utterances).

## 2.3 | Statistical analyses

Most studies analyzing developmental traits use linear models despite the inherent non-linear nature of ontogenetic trajectories. In this study, we rather used non-linear models to study developmental traits such as vocal structural complexity. To this aim, we developed customized Generalized non-Linear Models for each of the four variables of interest in this study. Each model comprised a sigmoidal effect of age and all models had a similar structure. We used the average age per individual and season. In each of the analyses we modelled how sex influenced the developmental trajectory of chimpanzees in order to estimate the potential contribution of sexual selection. In all the models we also included dominance rank as a predictor since dominance was shown to influence vocal production in chimpanzees (Crockford et al., 2015; Fedurek & Slocombe, 2013; Gruber & Zuberbühler, 2013). We determined the dominance rank for each individual using a modified version of the Elo rating method (Neumann et al., 2011) (see Supplementary material for details). Prior to fitting the models, we z-transformed rank to a mean of zero and a standard deviation of one.

### 2.3.1 | Maximum utterance length (model 1)

In our first analysis, we modelled the developmental trajectory of the maximum length of the utterances produced per individual and study period. We estimated the utterance length as the number of different successive vocal units (or "call types") produced within a vocal utterance, which were each different from the preceding unit (in keeping with Girard-Buttoz, Zaccarella et al., 2022). Utterance length ranged from 1 to 10 vocal units in our study. Our definition diverges from some other studies which used "length" to include the number of repetitions of the same elements or vocal units (i.e., the number of "dee" elements in a row, Templeton et al., 2005). However, in chimpanzees studies show that the acoustic parameters of the single vocal units rather than their repetition make them context-specific (Crockford & Boesch, 2003a; Dezecache et al., 2019; Fedurek et al., 2021; Laporte & Zuberbühler, 2011; Leroux et al., 2021; Slocombe & Zuberbühler, 2005a, 2005b).

The principal expectation we had was that the maximum utterance length would be initially low, then, increase with age, and finally level off at an asymptote. Such a sigmoidal function of age can be parameterized, as:

$$Max.\ length = c_1 + c_2 \cdot \frac{1}{1 + e^{\frac{age - c_3}{c_4}}} \qquad (1)$$
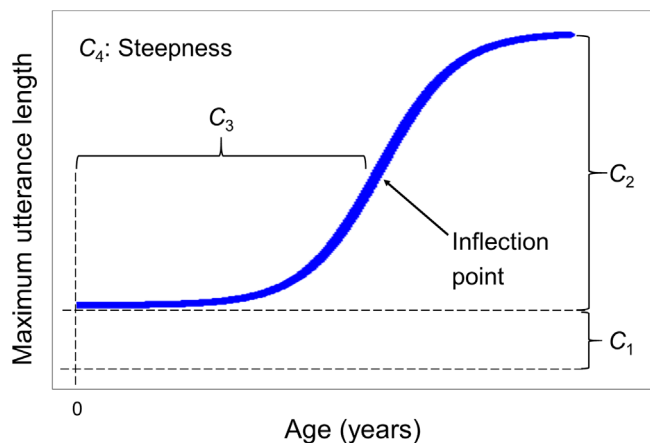
**FIGURE 2** Illustration of our hypotheses describing the ontogenetic trajectory of maximum utterance length as a function of age and interpretation of the parameters of a sigmoidal function.

For the interpretation of the respective parameters to be fitted ($c_1$, $c_2$, $c_3$, and $c_4$) see Figure 2.

We expected the location of the inflection point ($c_3$; i.e., the age of the steepest increase) and the steepness of the change ($c_4$) to both depend on the sex and rank of the individuals. Furthermore, we expected the final asymptote ($c_2$) to potentially differ between sexes. Besides, we expected $c_3$ and $c_4$ to depend on the dominance rank. Assuming sex to be represented by a dummy variable ($DM$), coding females as 0 and males as 1, this can be achieved by letting the model formula being:

$$Max.\ length = c_1 + (c_2 + c_{2M} \cdot DM) \cdot \cfrac{1}{1 + e^{\left(\frac{age - c_3 + c_{3M} \cdot DM + c_{3R} \cdot rank}{c_4 - c_{4M} \cdot DM + c_{4R} \cdot rank}\right)}} \quad (2)$$

In this equation, $c_2$ represents how much the estimated final asymptote for females (with a rank being zero, corresponding to the average rank as rank was z-transformed) differs from the initial asymptote ($c_1$), and $c_{2M}$ estimates how much the final asymptote in males (with rank being zero) differs from that of females. The parameters $c_3$, $c_{3M}$ and $c_{3R}$ estimate the location of the inflection point and how much it varies with sex and rank. Correspondingly, $c_4$, $c_{4M}$, and $c_{4R}$ estimate the steepness of the increase and how it varies between sexes and depends on rank (note in several equations we inserted non-essential brackets to improve clarity).

A further consideration with this model is that all its parameters and the response are constrained. The response variable is constrained to be >0, and the parameters in Equation (1) need to be constrained to be >0 for $c_1$, $c_2$, and $c_3$, and <0 for $c_4$. In case the parameters are constrained as just described, the response is automatically constrained appropriately. One way of achieving this is to let the estimation take place in an unconstrained parameter space but transform the parameters such that the actual parameters determining the fitted model are constrained (Bolker, 2008). This can be achieved by exponentiation,

thus Equation (2) turns into:

$$Max.\ length = e^{c_1} + e^{(c_2 + c_{2M} \cdot DM)} \cdot \cfrac{1}{1 + e^{\left(\frac{age - e^{(c_3 + c_{3M} \cdot DM + c_{3R} \cdot rank)}}{-e^{(c_4 + c_{4M} \cdot DM + c_{4R} \cdot rank)}}\right)}} \quad (3)$$

where $e$ is Euler's number.

In the following, we shall refer to the right-hand side of Equation (3) as "sigmoidal."

An issue with the model as laid out in Equations (1) to (3) is that applies only in case the number of recorded utterances is large and identical for all individuals in both seasons. However, the total number of recorded utterances differed between individuals and seasons, and it is obvious that the maximum utterance length will be the lower, the lower the number of recorded utterances (and obviously zero for zero recordings). On the other hand, for large numbers of recorded utterances, varying numbers of recorded utterances will have less of an effect on the maximum utterance length. Hence, assuming anything else to be identical, the maximum utterance length will reach an asymptote as the total number of recorded utterances approaches infinity (Figure S5a). Such an exponential curve can be parameterized as:

$$y = c - a \cdot b^x \quad (4)$$

Where $x$ and $y$ are the number of recorded utterances and the value of the function, respectively, $c$ is the asymptote of $y$ for $x = \infty$, $b$ is a steepness parameter and $a$ determines the value of $y$ for $x = 0$ (since $y_{(x=0)} = c + a^* b^0 = c + a^* 1 = c + a$). For a large number of recorded utterances, the estimated ontogeny will be dominated by the sigmoidal function (Equation 3) but the smaller the number of recorded utterances the more it will be simple function of the number of recorded utterances (Equation 4). This can be achieved by first setting the asymptote of Equation (4) to 1 and its value at $x = 0$ to 0, in which case Equation (4) simplifies to

$$y = 1 - 1 \cdot b^x \quad (5)$$

This leaves $b$ the only parameter to be estimated (Figure S5b; the right-hand side in Equation 5 we thereafter refer to as "saturation"). In a second step, we can then model maximum utterance length as follows:

$$Max.length = sigmoidal \cdot saturation \quad (6)$$

An asymptote of 1 for the saturation part in Equation (6) has the consequence that the parameters in *sigmoidal* (Equation 3) reveal the estimated maximum utterance length for an infinite number of utterances (Figure S5c).

Such an asymptote for the saturation part in Equation (5) can be achieved by bounding $b$ between 0 and 1. However, it might be desirable to model $b$ in an unconstrained parameter space. This can be achieved by means of the inverse logit transformation; that is,

$$b_{constrained} = \frac{e^{b_{unconstrained}}}{1 + e^{b_{unconstrained}}} \quad (7)$$

Hence, the right-hand side of Equation (7) replaced $b$ in Equation (5).

As starting values of the fitting process of the sigmoidal model we chose the logarithm of the minimum utterance length for $c_1$, the logarithm of half the difference between the maximum and the minimum of the maximum utterance length for $c_2$, the logarithm of seven for $c_3$, and zero for all others.

## 2.3.2 | Diversity of utterances (model 2)

With the second model, we investigated the diversity of utterances, namely the proportion of unique utterances in a set of recorded utterances, produced by each individual. At a first glance, one might think that a standard logistic model could be used. However, as this is linear in link space it necessarily asymptotes at values of 0 and 1. While an asymptote of zero for very young ages makes intuitive sense, and assuming an asymptote of 1 for old individuals is inappropriate. We hence fitted a function which was defined as:

$$P_{unique} = \frac{e^{(c_0+c_A\cdot age+c_R\cdot rank+c_S\cdot sex)}}{1+e^{(c_0+c_A\cdot age+c_R\cdot rank+c_S\cdot sex)}} \cdot \frac{e^{(c_{asym}+c_{asymM}\cdot DM)}}{1+e^{(c_{asym}+c_{asymM}\cdot DM)}}$$
$$\cdot \left( \left(\frac{e^b}{1+e^b}\right)^{(N-1)} \right) \tag{8}$$

The first part of this function (the first fraction) is identical to a standard logistic regression model and has asymptotes of 0 and 1, respectively. Also, the parameters to be estimated in this part ($c_0$, $c_A$, $c_R$, $c_M$) have the same interpretation as in a standard logistic model and allow to estimate the age dependent development of the probability of an utterance to be panted ($c_0$) and also how this probability varied with sex and rank ($c_M$, $c_R$; as above, rank is z-transformed and $M$ is a dummy variable). These parameters can be modelled in an unbound space.

The second fraction, is included to allow for an asymptote smaller than 1. The asymptote needs to be bound between 0 and 1, and we again, achieved this by modelling $c_{asym}$ and $c_{asymM}$ in an unconstrained parameter space and transforming their sum (after multiplying $c_{asymM}$ with $M$) by means of the inverse logit transformation. The term $c_{asymM}$ estimates how much the asymptote in males (with rank being zero) differs from that of females.

The third fraction of the function, is included to control for the number of recorded utterances. Since one can reasonably assume that utterances are composed out of a limited set of building blocks and that utterances are limited in their maximum length; the proportion of unique utterances among a set of recorded utterances, will be smaller the larger the number of recorded utterances. This can again be achieved by means of an exponential function (Figure S6). By modifying Equation (4) such that $c$ is fixed at a value of 0, $a$ at a value of 1, and exponentiating $b$ with the number utterances $-1$, it can be ensured that the function has a value of 1 when the number of recorded utterances is one and asymptotically approaches a value of 0 when the number of recorded utterances increases toward infinity. Hence, $b$ is the only parameter that needs to be estimated. As above, $b$ needs to be constrained, this time such that $0 < b < 1$, and we again achieved this by

estimating it in an unconstrained parameter space and then transforming it by means of the inverse logit transformation. In the Equation (8), $N$ is the number of recorded utterances. The starting values for $c_{asym}$ and $c_{asymM}$ was set to 0 (corresponding to an asymptote of 0.5), and $b$ got a starting value of 5 which was chosen such that the exponential function roughly parallelized the decrease of the proportion of unique utterances among all utterances when plotting it against the number of recorded utterances. For the other parameters, we used the estimated coefficients of a standard logistic model fitted to the data as starting values. The response variable for this model was a two columns matrix with the number of unique and repeated utterances, respectively (see Baayen, 2008). Hence, we essentially modelled the proportion of panted utterances as a function of age, sex and, rank.

## 2.3.3 | Probability of panted utterances (model 3)

With the third model, we investigated how the probability of an utterance to have a panted unit developed with age and, as previously, whether the age related trajectory varied between sexes. In this model, we also controlled for dominance rank. We defined an utterance as panted if it comprised at least one panted unit (i.e., PN, PB, PG, PH, PR, PS, or PW). We essentially modelled the proportion of panted utterances as a function of age, sex, and rank. The response can again be a matrix with two columns being the number of panted and non-panted utterances, respectively. Here, again, an asymptote of zero for very young ages makes intuitive sense, and assuming an asymptote of 1 for old individuals is inappropriate. In fact, chimpanzee use six non-panted vocal units in their repertoire and these units are often produced singly or in sequences comprising other non-panted units (Girard-Buttoz, Zaccarella et al., 2022). Therefore, the probability of panted utterances must clearly be below 1. Hence, the function fitted for model 3 was identical to that fitted for model 2 (see Equation 8), with the exception that the last parameter controlling for the total number of utterances recorded was not needed:

$$P_{panted} = \frac{e^{(c_0+c_A\cdot age+c_R\cdot rank+c_S\cdot sex)}}{1+e^{(c_0+c_A\cdot age+c_R\cdot rank+c_S\cdot sex)}} \cdot \frac{e^{(c_{asym}+c_{asymM}\cdot DM)}}{1+e^{(c_{asym}+c_{asymM}\cdot DM)}} \tag{9}$$

Starting parameters for the fitting process were set as described for model 2.

## 2.3.4 | Probability of two adjacent pants (model 4)

With the last model, we investigated how the probability of an utterance to have two adjacent panted units (e.g., PH-PS) developed with age, and whether the age-related trajectory varied between sexes, controlling for dominance rank. We model this in the exact same way as model 3, with the response variable being whether an utterance comprised at least two adjacent panted units instead of the probability of an utterance to comprise one panted unit.

## 2.4 | Implementation

We fitted all models in R (version 4.0.3; R Core Team, 2020) using the function optim (with the optimizer "BFGS"). As the value evaluated to determine model fit (given a set of parameter values) we used the log-likelihood, that is the sum of the logarithms of the probabilities to observe the individual instances of the response, given the fitted model. Hence, the models were fitted using maximum likelihood. After an initial call of the function optim, we called it repeatedly, each time using as starting values the estimated parameters of the previous call of optim until the difference between the log-likelihoods of the current and the previous solution was <0.0001. The function optim indicated all models to have converged (convergence value = 0). The functions we used for determining the probabilities of the individual observations of the response were dztpois of the R package actuar (version 3.1-1; Dutang et al., 2008) in the case of model 1 for which we used a zero-truncated Poisson error distribution and dbinom in case of models 2, 3, and 4 for which we used a binomial error distribution.

One issue was that we had two observations for a fraction of the individuals in the data set (46 out of 96). To avoid the consequences of pseudo-replication, we drew inferences by means of a non-parametric bootstrap (N = 1000 bootstraps). To this end, we randomly sampled the individuals (with replacement). We used these bootstraps to obtain 95% confidence limits of the estimated coefficients, fitted values, and parameters of the fitted models such as the locations of the inflection points or the final asymptotes. For model 1 we determined these confidence intervals of the fitted model assuming an infinite number of utterances and for model 2 we determined them for a number of utterances equalling its average across the entire data set (the fitted probabilities of model 3 and 4 are unconditional on the number of recordings). A confidence interval not comprising the value 0, means that the estimate is different from 0% with 95% confidence. We assessed model fit by visually judging how close the fitted model matched the observations of the response and how wide the confidence intervals of the fitted models are. We did not assess model fit more formally due to having two data points for a fraction of the individuals in our data. Hence, formal tests would be pseudo-replicated and thus biased (mixed models were computationally not feasible).

We estimated model stability by dropping individuals from the data, one at a time, fitting the models to the respective subsets, and finally determining the range of the derived estimates. This revealed all models to be of good stability (Table 1). In model 1, the response was quite underdispersed (dispersion parameter: 0.39). In model 2, the response was only slightly overdispersed (dispersion parameter: 1.13). In model 3 the response was clearly overdispersed (dispersion parameter: 1.81), and in model 4 it was heavily overdispersed (dispersion parameter: 2.36). Hence, the confidence intervals of these models are likely too narrow. The samples analyzed comprised 144 data points obtained from 98 individuals.

Since age and sex did not have a consistent effect in our main model, we fitted, a-posteriori, a simplified version of each of the four models, including only age as a predictor. We conducted this additional analysis to confirm that the developmental trajectory found hold in the simpler models removing variables with weak effects. We found similar results with respect to the effect of age compared to the original models in which we also included sex and rank as additional predictors (Figure S8, Tables S8 and S9).

## 3 | RESULTS

### 3.1 | Maximum utterance length (model 1)

The maximum utterance length (model 1) clearly increased with age, and the age-related trajectory did not strongly depend on sex or dominance rank (Figure 3a; Figure S7a; Table 1). Translating the fitted model parameters into the parameters of the sigmoidal function revealed that the steepest increase of maximum utterance length happened at an age of 3.8 and 4.2 years in females and males, respectively. The age of the steepest increase varied with dominance rank, but these differences were also associated with large uncertainty (wide confidence intervals in Table S4). The maximum utterance length was estimated to reach values of 6.18 and 5.89 different vocal units within a sequence in adult females and males, respectively (confidence interval (CI), females: 5.75–7.36; males: 5.17–7.44; Table S4). A visual inspection of the fitted model (Figure 3a) revealed that the adult levels of utterance length (i.e., the asymptote) were reached at around 8 years of age.

Although individuals under 1 year produced utterances with sequences of 2 and 3 vocal units, those cases were rare, with most of their utterances composed of single vocal units only (Figure 4). When we determined the proportion of utterances emitted as sequences per age-year, an increase in the production of sequences through the first 15 years of life was obvious, with a more consistent use of utterances with six or more vocal units from around 8–9 years of life (Figure 4).

### 3.2 | Diversity of utterances (model 2)

The proportion of unique utterances in a set of utterances (model 2) clearly increased with age and the age-related increase did not strongly depend on sex or dominance rank (Figure 3b; Figure S7b; Table 1). The model indicated that the steepest increase took place at ages of 2.8 and 2.2 in females and males, respectively (Table S5). The uncertainty with regard to the dependency of the age of steepest increase on sex and rank was considerable, and the model has to be considered inconclusive regarding this aspect. Finally, the proportion of unique utterances was estimated to reach values of 0.42 and 0.37 in adult females and males, respectively (CI, females: 0.34–0.45; males: 0.27–0.43; Table S5), whereby the average number of utterances per individual was 74 (the average number of utterances per chimpanzee in the data set). In other words, for an adult female chimpanzee from which we recorded 74 utterances, we expect to find on average about 31.08 unique different utterances and 42.92 utterances that will be repetitions of those unique ones (Figure S7b). A visual inspection of the fitted model (Figure 3b) revealed that the adult levels of the proportion of unique utterances (i.e., the asymptote) were reached at around 10 years of age.
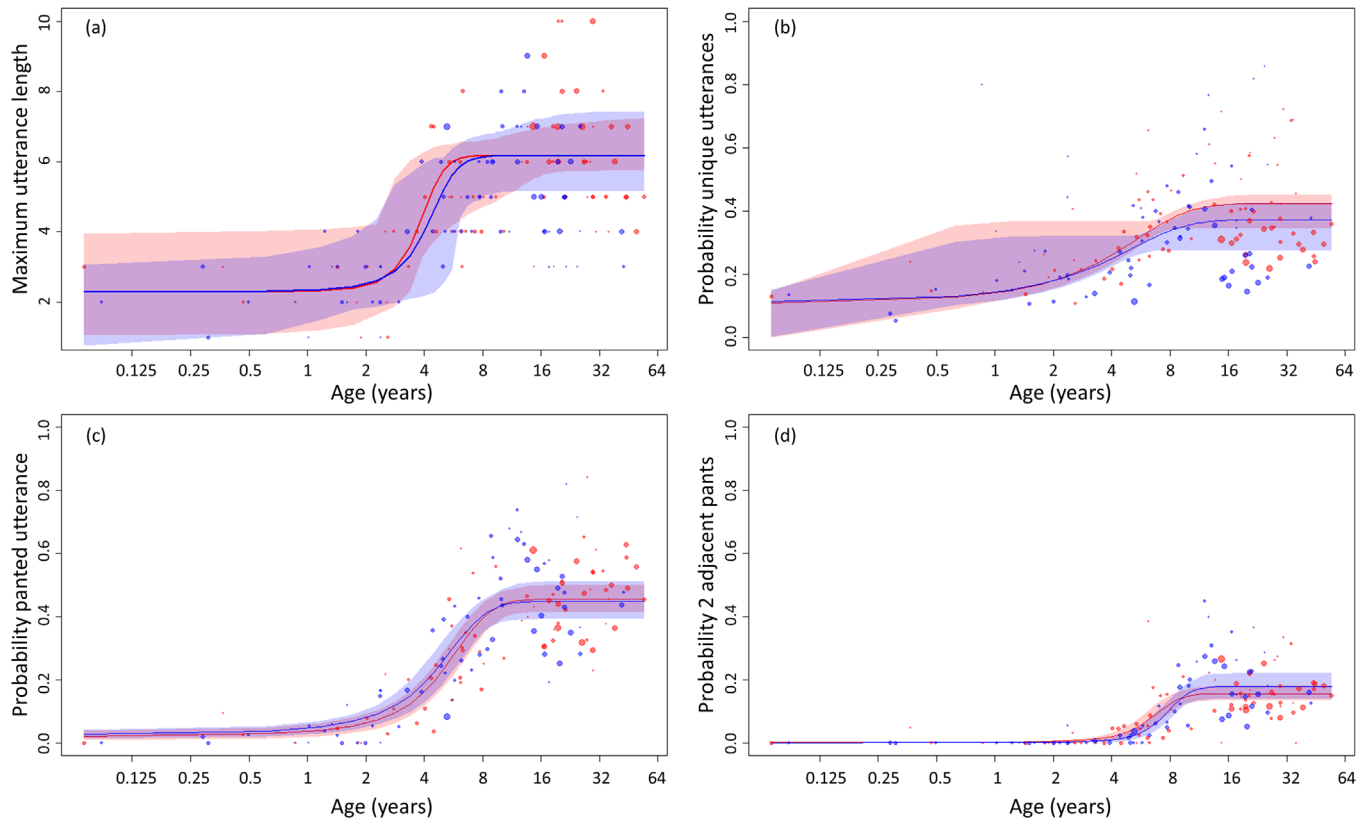
**FIGURE 3** Changes of chimpanzee vocal complexity through development in terms of maximum utterance length (a; model 1), proportion of unique utterances in a set of recorded utterances (b; model 2), probability of panted utterance (c; model 3), and probability of two adjacent pants (d; model 4), as a function of age (on a log scale). Red depicts females and blue males. The lines and polygons depict the fitted model and its confidence limits for rank being centred. The area of the dots is proportionate to the total number of utterances per chimpanzee and study period (range: 5–265). The fitted models in (a) and (b) show the estimated response for a number of utterances equalling their average across the entire data set.
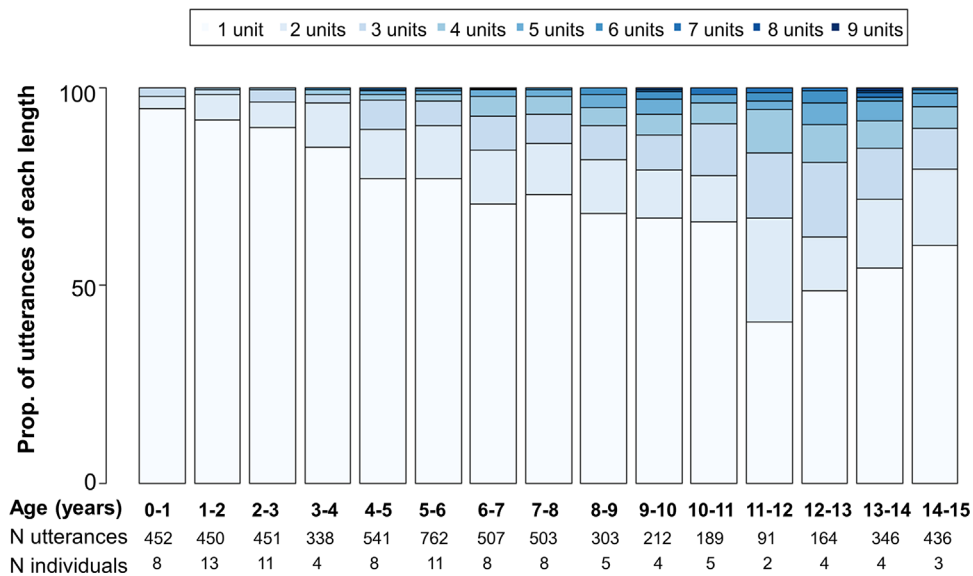


**FIGURE 4** Proportion of utterances of different lengths through chimpanzee ontogeny up to 15 years of age. N utterances represents the total number of utterances (single units and sequences) produced by all individuals of a given age range. The length of the utterance refers to the number of different successive vocal units produced within a vocal utterance.

**TABLE 1** Results of the generalized non-linear models for each of our four variables of vocal structural complexity[a]

|  | Term | Estimate | SE | Lower CI | Upper CI | Min | Max |
|---|---|---|---|---|---|---|---|
| Model 1 | $c_1$ | 0.821 | 0.179 | −8.716 | 1.014 | 0.748 | 0.880 |
|  | **$c_2$** | **1.363** | **0.140** | **1.228** | **1.926** | **1.324** | **1.401** |
|  | $c_{2M}$ | -0.075 | 0.179 | -0.402 | 0.207 | -0.128 | 0.005 |
|  | **$c_3$** | **1.327** | **0.186** | **0.491** | **2.025** | **1.268** | **1.436** |
|  | $c_{3M}$ | 0.099 | 0.322 | -0.910 | 0.601 | -0.024 | 0.300 |
|  | $c_{3R}$ | 0.095 | 0.227 | -0.712 | 0.472 | -0.055 | 0.164 |
|  | $c_4$ | -0.516 | 1.020 | -3.266 | 2.255 | -1.054 | 0.177 |
|  | $c_{4M}$ | 0.282 | 1.288 | -5.735 | 3.486 | -0.547 | 0.829 |
|  | $c_{4R}$ | 1.168 | 0.583 | -0.568 | 3.870 | 1.065 | 1.708 |
|  | **b** | **2.535** | **0.287** | **2.265** | **2.854** | **2.499** | **2.607** |
| Model 2 | $c_0$ | -1.097 | 0.175 | -8.255 | -0.663 | -1.158 | -1.024 |
|  | **$c_A$** | **0.391** | **0.036** | **0.326** | **17.130** | **0.379** | **0.401** |
|  | $c_S$ | 0.227 | 0.187 | -2.593 | 2.915 | 0.159 | 0.335 |
|  | $c_R$ | 0.209 | 0.123 | -4.876 | 1.752 | 0.142 | 0.256 |
|  | $c_{asym}$ | 0.347 | 0.092 | -0.373 | 0.652 | 0.293 | 0.454 |
|  | $c_{asymM}$ | -0.283 | 0.078 | -0.596 | 0.054 | -0.371 | -0.241 |
|  | **$b_N$** | **5.415** | **0.063** | **5.168** | **6.128** | **5.291** | **5.448** |
| Model 3 | $c_0$ | -3.075 | 0.209 | -3.941 | -2.426 | -3.319 | -2.973 |
|  | **$c_A$** | **0.620** | **0.041** | **0.479** | **0.813** | **0.595** | **0.668** |
|  | $c_S$ | 0.287 | 0.165 | -0.466 | 1.076 | 0.152 | 0.509 |
|  | $c_R$ | 0.618 | 0.117 | -0.002 | 1.171 | 0.469 | 0.744 |
|  | $c_{asym}$ | -0.180 | 0.032 | -0.346 | -0.005 | -0.224 | -0.145 |
|  | $c_{asymM}$ | -0.028 | 0.049 | -0.311 | 0.260 | -0.078 | 0.020 |
| Model 4 | $c_0$ | -5.232 | 0.496 | -6.928 | -4.431 | -5.502 | -5.112 |
|  | **$c_A$** | **0.823** | **0.083** | **0.689** | **1.147** | **0.797** | **0.886** |
|  | $c_S$ | -0.949 | 0.316 | -2.231 | 0.140 | -1.205 | -0.688 |
|  | $c_R$ | 0.304 | 0.217 | -0.392 | 1.163 | 0.089 | 0.551 |
|  | **$c_{asym}$** | **-1.687** | **0.044** | **-1.866** | **-1.515** | **-1.747** | **-1.660** |
|  | $c_{asymM}$ | 0.159 | 0.068 | -0.188 | 0.485 | 0.094 | 0.219 |

[a]Model 1: maximum utterance length, dispersion parameter 0.39. Model 2: diversity of utterances, dispersion parameter 1.13. Model 3: probability of panted utterance, dispersion parameter 1.81. Model 4: probability of two adjacent pants, dispersion parameter 2.36. In all models the rank was z-transformed, the mean and SD of the original rank are 0.447 and 0.249, respectively. Shown are estimated coefficients (in unconstrained parameter space) of the fitted models, their confidence limits, and the minimum and maximum of estimates obtained when excluding individuals one at a time. Terms in bold depict estimates different from 0% with 95% of confidence.

## 3.3 | Probability of panted utterances (model 3)

The probability of an utterance to have a panted unit (model 3) clearly increased with age, whereby the age-related trajectory did not strongly depend on sex or dominance rank (Figure 3c; Figure S7c; Table 1). Translating the model estimates into the fitted curve revealed that the steepest increase happened at ages of 4.9 and 4.5 in females and males, respectively. However, given that each of these ages fell well within the wide confidence interval of the respective other (Table S6), these cannot be considered as a consistent difference. The probability of panted utterances was estimated to reach values of 0.46 and 0.45 in adult female and males, respectively (CI, females: 0.41–0.49; males:

0.39–0.51; Table S6). A visual inspection of the fitted model (Figure 3c) revealed that the adult levels of the probability of an utterance to be panted (i.e., the asymptote) were reached at around 10 years of age.

## 3.4 | Probability of two adjacent pants (model 4)

The probability of an utterance to comprise two adjacent panted units (model 4) increased with age, whereby the age-related trajectory did not strongly depend on sex or dominance rank (Figure 3d; Figure S7d; Table 1). Translating the model estimates into the fitted curve revealed that the steepest increase happened at ages of 6.3 and 7.5 in females

### (a) Proportion of each vocal unit in single units

### (b) Proportion of each vocal unit in sequences
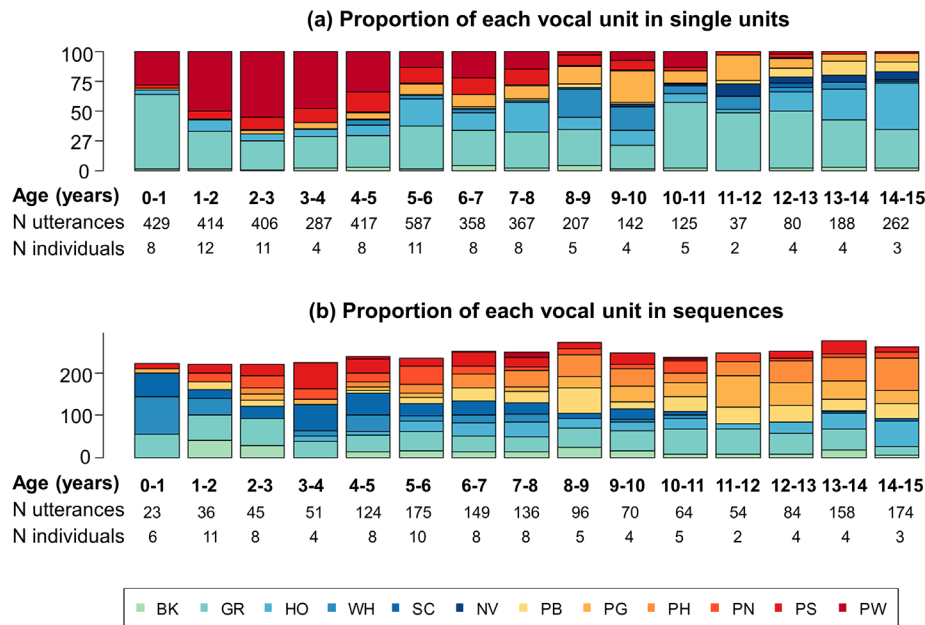


**FIGURE 5** Proportion of each vocal unit produced in (a) single units and (b) sequences throughout chimpanzee ontogeny up to 15 years of age. Gradient of colors: blue corresponds to non-panted units, orange to panted units. N utterances represents the total number of utterances (single units or sequences) produced by all individuals in a given age category. (b) In sequences with three or more units, the same vocal unit can be repeated (in non-adjacent positions) within the utterance, leading to a y-axis with proportions larger than 100%.

and males, respectively. However, given the wide confidence interval (Table S7), these cannot be considered as a consistent difference. The probability of two adjacent pants was estimated to reach values of 0.16 and 0.18 in adult female and males, respectively (CI, females: 0.13–0.18; males: 0.14–0.22; Table S7). A visual inspection of the fitted model (Figure 3d) revealed that the probability to produce two adjacent pants began to clearly increase around 4 years and adult levels (i.e., the asymptote) were reached at around 10 years of age. The emergence and steepest increase of the probability to produce two adjacent panted units (Figure 3d) seemed to occur later in ontogeny that that of producing one panted unit (Figure 3c).

To further understand the structural changes seen in our variables of vocal complexity, we visually examined the proportion of each vocal unit produced within single units and sequences (Figure 5). Chimpanzees seemed to increase their repertoire diversity through ontogeny mostly by the production of panted utterances (orange gradient in Figure 5) in both single units and sequences. Besides this, it revealed that the increase in the probability to produce panted utterances seen in model 2 was not driven by one or two panted units, but by all of the panted units of the repertoire. Furthermore, immature chimpanzees by the age of 4 years were able to produce all the vocal units present in the adultt's repertoire, except for PR, which we recorded only once in this dataset (Table S3).

## 4 | DISCUSSION

Most studies on the ontogeny of vocal communication in primates focus on song, acoustic variation or vocal usage (i.e., the context in which the call is being used; Janik & Slater, 2000) of specific vocali-

sations in a species' repertoire (Dezecache et al., 2020; Gouzoules & Gouzoules, 1989; Hammerschmidt et al., 2000; Hauser, 1989; Koda et al., 2013; Laporte & Zuberbühler, 2011; Levréro & Mathevon, 2013; Lieblich et al., 1980; Omedes, 1985; Seyfarth & Cheney, 1986, 2010). In contrast, we here investigated the development of structural complexity across the whole vocal repertoire and in particular the production of vocal sequences. Assessing the latter, especially, is critical to gain understanding of how and why flexible combinatorial capacities evolved in human language.

Our analysis revealed that the four structural components of vocal complexity that we analyzed followed similar developmental trajectories, namely utterance length, diversity, probability of an utterance to comprise a panted unit, and probability of an utterance to comprise two adjacent panted units. Chimpanzees produced a low diversity of short and non-panted utterances in the first year of life. There was a clear increase in the four variables of vocal complexity throughout ontogeny: utterance diversity presented the steepest increase just before weaning age (~2.5 years), utterance length and probability to produce panted utterances at around the weaning age (4 years) and probability to produce two adjacent pants at around 7 years. Vocal structural complexity reached asymptote at the beginning of the sub-adult life (8–10 years), indicating the four variables studied had reached adult levels. This is much later in development than the age at which chimpanzees were able to produce all the single vocal units present in the adult repertoire (4 years).

These results suggest that somewhat complex and slow paced developmental maturational process are likely to be involved in the development of vocal structural complexity, which extends well beyond the age when the full range of sounds can be produced. This finding differs from the pattern of vocal development presumed in other

primates, where the vocal repertoire is thought to be inflexible and present soon after birth (but see Gultekin et al., 2021). However, most of these primate studies are based only on the production of the single units of the primatet's repertoires. Inferring which precise maturational processes might affect the development of vocal complexity is not trivial, as more than one mechanism can operate concurrently. This is the case not only in humans but also in other primates such as marmosets, where changes in vocal behavior correlate with both motor and social milestones through development (Gultekin et al., 2021). Nonetheless, we can discount our first hypothesis, that of the involvement of limited maturational processes in vocal complexity development. We discuss the results with respect to each of the remaining three hypotheses we examined.

## 4.1 | Sexual selection hypothesis

Our results do not support the hypothesis that vocal structural complexity is primarily a sexually selected trait. First, the age of the steepest increase for all four variables appeared before the onset of sexual maturation, for both sexes. Second, the ontogenic trajectories of the four structural characteristics of vocal complexity did not obviously vary between sexes. This is interesting given that specific vocalizations of the chimpanzee repertoire show sexual dimorphism in rates of production. For instance, adult males produce higher rates of the loud pant-hoot sequence than females (Kalan, 2019; Pusey, 1990), and only females emit copulation screams (Townsend et al., 2008). Sexual selection may thus operate on call rates of certain vocal units rather than on the overall complexity of the vocal repertoire.

## 4.2 | Neuro-muscular maturation hypothesis

Some, but not all, of our specific predictions regarding the neuro-muscular hypothesis were fulfilled, suggesting some influence of neuro-muscular processes in the development of vocal structural complexity. Indeed, producing utterances with two adjacent panted units, which might require a higher level of sublaryngeal, laryngeal, and supralaryngeal coordination, seems to emerge and increase later in ontogeny (Figure 3d) than the other three variables we investigated (Figure 3a–c), and later than we predicted. This suggests that either the neuro-muscular development of fine movements such as supralaryngeal articulators is not completed by weaning age, or the neuro-muscular maturation is not the main factor limiting the development of vocal complexity at the juvenile stage, as not only utterances with adjacent panted units increased with ontogeny but all types or sequences.

## 4.3 | Social complexity hypothesis

Our findings regarding three of the four structural characteristics studied (utterance length, diversity and probability of panted utterance)

support the model that vocal complexity might be influenced by social complexity, with the steepest increase of vocal complexity occurring from 2 to 5 years, and a plateau reached at around the beginning of sub-adult life. First social interactions outside the family unit (i.e., genital inspection or aggression; ∼ 2 years; Bründl et al., 2021) and independent travel (no longer being carried by the mother which coincides with weaning; 4 years) should enhance the social exposure of young chimpanzees. It is therefore likely that the steep increase in vocal structural complexity, between 2 and 5 years, is driven, at least in part, by an expansion of the social world. This is also supported by our finding that chimpanzees reach adult levels of production of the four variables at 8–10 years of age. This is the age at which young chimpanzees begin to associate with other group members (Pusey, 1990), independently from their mother, and become integrated in the community, such that interaction partners begin to include all community members in a wider range of contexts. In a comparative study across primates species, McComb and Semple (2005) suggested that a greater number of unique vocalizations (repertoire size or diversity) might be needed for animals to navigate more complex networks of social relations. Our results expand the suggestions of McComb and Semple (2005), demonstrating its applicability within species in a developmental context, and opening up the concept not only to single vocal units within a vocal repertoire but also in terms of vocal sequences. Moreover, our findings complement a recent study showing that chimpanzees also increase the acoustic gradation within vocal units through ontogeny (Taylor et al., 2021). Overall, both acoustic gradation of vocal units and the ability to combine vocal units into longer and more diverse sequences could enable an expansion of the range of information conveyed throughout ontogeny, as the social and ecological environment becomes more complex.

Although social complexity and vocal complexity milestones co-occur across development, because of the confound of age, we cannot rule out that these simply co-occur because of general maturational processes rather than because one causes the other. However, the concurrent expansion of both is of interest as maturational processes of different cognitive processes are not necessarily temporally linked (Bründl et al., 2021). For example, with growing independence from the mother, diversity of activities is not necessarily temporally linked to increases in social complexity: "non-social" activities such as independent foraging may occur substantially earlier in ontogeny than related social activities, such as food sharing or cooperative hunting. In some cases, vocal complexity requires complex cognition (such as human language). However, the need for cognition with respect to the variables in this study is unclear, thus, we do not impute cognition here.

Regarding the statistical models we have used, we acknowledge that the use of non-linear models requires several decisions, for instance, about the particular shape of the ontogenetic trajectories assumed (i.e., their parameterization). With regular linear models, one gets standard solutions to standard problems, but, for non-linear models, such simple standards are not a given. Therefore, the particular parameterizations we chose in this study may leave room for debate, as other researchers may hypothesize different ontogenetic trajectories and hence parameterizations. For instance, all our models imply an ontogeny that is

symmetric, that is, the trajectory before the inflection point is rotationally symmetric to the one after it. However, such an assumption cannot be taken for granted. Furthermore, the existence of senescence (i.e., a decrease of the feature in very old individuals) seems a plausible hypothesis for the four variables we studied. When visually inspecting the figures, particularly the maximum utterance length (Figure 3a) seems to suggest that such senescence might indeed exist. We did not pursue this idea further as our focus was on early development. Additionally, we have limited data for individuals older than 35 years of age, which made it unlikely that such attempts could be successfully implemented, specifically given that they would have led to a considerable increase in the complexity of our models. For the same reasons, we did not include the community (East, North and South) as another factor in our models. Despite these limitations, we consider that our models fitted the data quite well and revealed something meaningful about the ontogeny of the structural characteristics of vocal complexity that we investigated. Nevertheless, with a larger amount of data, particularly for relatively old individuals it might be worth attempting to include potential senescence.

## 5 | CONCLUSION

Our study provides some valuable insight into the development of vocal structural complexity in a long-lived mammalian species with a protracted development, the chimpanzee. Chimpanzees need about 10 years to develop the vocal structural complexity present in the adult repertoire, way beyond the age of emergence of all the single vocal units. This contradicts current thinking that the full vocal repertoire in primates emerges rapidly after birth (Ey et al., 2007; Hammerschmidt et al., 2001). However, comparison with other primates is difficult as there is a lack of comparable studies of ontogenetic development of vocal sequences across the whole vocal repertoire. Our results suggest that multifaceted processes drive increases in vocal structural complexity, namely social complexity and neuro-muscular maturation. From a behavioral perspective, a reasonable developmental hypothesis is that as exposure to greater social complexity increases through ontogeny, chimpanzees may develop the ability to combine vocal units into complex sequences to overcome the constraints of their limited number of vocal units, thereby enhancing the range of information that can be encoded. Our results support the idea that social complexity might promote the evolution of complex vocal repertoires, as an ultimate explanation for human language evolution. However, we need further studies to investigate the extent to which the increase in vocal structural complexity observed through development enhances meaning generation, to pinpoint the processes involved and how these are driven by the expansion of social complexity in chimpanzees. To confirm this hypothesis, future studies should assess whether vocal complexity of different chimpanzee populations varies with varying social and ecological exposure. In addition, further studies are needed with a similar developmental approach in other social mammals that vary in the complexity of their social systems. Also relevant is to assess whether changes in communication-related neural pathways occur during chimpanzee development. To date, neural pathways relevant for language and articulation, such as the arcuate fasciculus, have been studied only in adult primate brains (Rilling et al., 2008, 2012), thus how these pathways develop through ontogeny remains an open question.

## CONFLICT OF INTEREST

We have no conflict of interest to declare.

## DATA AVAILABILITY STATEMENT

The data analyzed for this study are openly available in "figshare" ahttps://doi.org/10.6084/m9.figshare.19336853. The code used for the analysis can be obtained by request to Roger Mundry.

## ETHICS STATEMENT

The study used only observational, non-invasive procedures. The Ministère de l'Enseignement Supérieur et de la Recherche Scientifique of Cote d'Ivoire (008/MESRS/DGRI) and the Office Ivoirien des Parcs et Réserves approved our study. International, national and institutional ethical guidelines were followed.

## REFERENCES

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3–4), 227–266. https://doi.org/10.1163/156853974İ00534

Andersson, M., & Iwasa, Y. (1996). Sexual selection. *Trends in Ecology & Evolution*, 11(2), 53–58. https://doi.org/10.1016/0169-5347(96)81042-1

Arcadi, A. C. (1996). Phrase structure of wild chimpanzee pant hoots: Patterns of production and interpopulation variability. *American Journal of Primatology*, 39(3), 159–178. https://doi.org/10.1002/(SICI)1098-2345(1996)39:3<159::AID-AJP2>3.0.CO;2-Y

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge, UK.

Baptista, L. F., Trail, P. W., DeWolfe, B. B., & Morton, M. L. (1993). Singing and its functions in female white-crowned sparrows. *Animal Behaviour*, 46(3), 511–524. https://doi.org/10.1006/anbe.1993.1219

Behringer, V., Deschner, T., Deimel, C., Stevens, J. M. G., & Hohmann, G. (2014). Age-related changes in urinary testosterone levels suggest differences in puberty onset and divergent life history strategies in bonobos and chimpanzees. *Hormones and Behavior*, 66(3), 525–533. https://doi.org/10.1016/j.yhbeh.2014.07.011

Blumstein, D. T., & Armitage, K. B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *The American Naturalist*, 150(2), 179–200. https://doi.org/10.1086/286062

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer (version 5.1. 05). Retrieved May 1, 2009.*

Bohn, K. M., Smarsh, G. C., & Smotherman, M. (2013). Social context evokes rapid changes in bat song syntax. *Animal Behaviour*, 85(6), 1485–1491. https://doi.org/10.1016/j.anbehav.2013.04.002

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton University Press.

Bründl, A. C., Tkaczynski, P. J., Nohon Kohou, G., Boesch, C., Wittig, R. M., & Crockford, C. (2021). Systematic mapping of developmental milestones in wild chimpanzees. *Developmental Science*, 24(1), e12988. https://doi.org/10.1111/desc.12988

Catchpole, C. K. (1987). Bird song, sexual selection and female choice. *Trends in Ecology & Evolution*, 2(4), 94–97. https://doi.org/10.1016/0169-5347(87)90165-0

Catchpole, C. K., & Slater, P. J. B. (2003). *Bird song: biological themes and variations*. Cambridge University Press.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, 7(7), 300–307. https://doi.org/10.1016/S1364-6613(03)00136-0

Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS ONE*, 1(1), e73. https://doi.org/10.1371/journal.pone.0000073

Connor, R. C., Mann, J., Tyack, P. L., & Whitehead, H. (1998). Social evolution in toothed whales. *Trends in Ecology & Evolution*, 13(6), 228–232. https://doi.org/10.1016/S0169-5347(98)01326-3

Cowlishaw, G. (1996). Sexual selection and information content in gibbon song bouts. *Ethology*, 102(2), 272–284. https://doi.org/10.1111/j.1439-0310.1996.tb01125.x

Crockford, C. (2019). Why does the chimpanzee vocal repertoire remain poorly understood and what can be done about it? In C. Boesch, R. Wittig, C. Crockford, L. Vigilant, T. Deschner, & F. Leendertz, (Eds.), *The chimpanzees of the taï forest* (1st ed., pp. 394–409). Cambridge University Press. https://doi.org/10.1017/9781108674218.025

Crockford, C., & Boesch, C. (2003a). Context-specific calls in wild chimpanzees, pan troglodytes verus: Analysis of barks. *Animal Behaviour*, 66(1), 115–125. https://doi.org/10.1006/anbe.2003.2166

Crockford, C., & Boesch, C. (2003b). Context-specific calls in wild chimpanzees, pan troglodytes verus: Analysis of barks. *Animal Behaviour*, 66(1), 115–125. https://doi.org/10.1006/anbe.2003.2166

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, 142(4), 397–421.

Crockford, C., Gruber, T., & Zuberbühler, K. (2018). Chimpanzee quiet hoo variants differ according to context. *Royal Society Open Science*, 5(5), 172066. https://doi.org/10.1098/rsos.172066

Crockford, C., Wittig, R. M., & Zuberbühler, K. (2015). An intentional vocalization draws others' attention: A playback experiment with wild chimpanzees. *Animal Cognition*, 18(3), 581–591. https://doi.org/10.1007/s10071-014-0827-z

Darwin, C. (1871). *The descent of man and selection in relation to sex*. John Murray.

De Gregorio, C., Valente, D., Torti, V., Zanoli, A., Colaci Cirillo, D., Bazzano, G., Sugliano, S., Randrianarison, R. M., Giacoma, C., & Gamba, M. (2019). *Song structure and sex specific features in the indris*. The Journal of the Acoustical Society of America, 145, 1919. https://doi.org/10.1121/2.0001021

Dezecache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2019). Machine learning reveals adaptive maternal responses to infant distress calls in wild chimpanzees [Preprint]. *Animal Behavior and Cognition*, https://doi.org/10.1101/835827 bioRxiv, 835827.

Dezecache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2020). Flexibility in wild infant chimpanzee vocal behavior. *Journal of Language Evolution*, 6(1), 37–53. https://doi.org/10.1093/jole/lzaa009

Dutang, C., Goulet, V., & Pigeon, M. (2008). actuar: An r package for actuarial science. *Journal of Statistical Software*, 25(7), 1–37.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., & Fischer, J. (2007). Age- and sex-related variations in clear calls of papio ursinus. *International Journal of Primatology*, 28(4), 947–960. https://doi.org/10.1007/s10764-007-9139-3

Fedurek, P., & Slocombe, K. E. (2013). The social function of food-associated calls in male chimpanzees: Chimpanzee food-associated calls. *American Journal of Primatology*, 75(7), 726–739. https://doi.org/10.1002/ajp.22122

Fedurek, P., Tkaczynski, P. J., Hobaiter, C., Zuberbühler, K., Wittig, R. M., & Crockford, C. (2021). The function of chimpanzee greeting calls is modulated by their acoustic variation. *Animal Behaviour*, 174, 279–289. https://doi.org/10.1016/j.anbehav.2021.02.002

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2(12), e1600723. https://doi.org/10.1126/sciadv.1600723

Freeberg, T. M. (2006). Social complexity can drive vocal complexity: Group size influences vocal information in carolina chickadees. *Psychological Science*, 17(7), 557–561. https://doi.org/10.1111/j.1467-9280.2006.01743.x

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1785–1801. https://doi.org/10.1098/rstb.2011.0213

Friederici, A. D. (2017). Evolution of the neural language network. *Psychonomic Bulletin & Review*, 24(1), 41–47. https://doi.org/10.3758/s13423-016-1090-x

Gamba, M., Torti, V., Estienne, V., Randrianarison, R. M., Valente, D., Rovara, P., Bonadonna, G., Friard, O., & Giacoma, C. (2016). The indris have got rhythm! timing and pitch variation of a primate song examined between sexes and age classes. *Frontiers in Neuroscience*, 10, 249. https://doi.org/10.3389/fnins.2016.00249

Garwicz, M., Christensson, M., & Psouni, E. (2009). A unifying model for timing of walking onset in humans and other mammals. *Proceedings of the National Academy of Sciences*, 106(51), 21889–21893. https://doi.org/10.1073/pnas.0905777106

Geissmann, T., & Orgeldinger, M. (2000). The relationship between duet songs and pair bonds in siamangs, hylobates syndactylus. *Animal Behaviour*, 60(6), 805–809. https://doi.org/10.1006/anbe.2000.1540

Girard-Buttoz, C., Bortolato, T., Laporte, M., Grampp, M., Zuberbühler, K., Wittig, R. M., & Crockford, C. (2022). Population-specific call order in chimpanzee greeting vocal sequences. *IScience*, 25(9), 104851. https://doi.org/10.1016/j.isci.2022.104851

Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M., & Crockford, C. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology*, 5(1), 410. https://doi.org/10.1038/s42003-022-03350-8

Goodall, J. (1986). *The chimpanzees of gombe: Patterns of behavior*. Cambridge Mass.

Gouzoules, H., & Gouzoules, S. (1989). Sex differences in the acquisition of communicative competence by pigtail macaques (Macaca nemestrina). *American Journal of Primatology*, 19(3), 163–174. https://doi.org/10.1002/ajp.1350190304

Grawunder, S., Uomini, N., Samuni, L., Bortolato, T., Girard-Buttoz, C., Wittig, R. M., & Crockford, C. (2022). Chimpanzee vowel-like sounds and voice quality suggest formant space expansion through the hominoid lineage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1841), 20200455. https://doi.org/10.1098/rstb.2020.0455

Gruber, T., & Zuberbühler, K. (2013). Vocal recruitment for joint travel in wild chimpanzees. *PLoS ONE*, 8(9), e76073. https://doi.org/10.1371/journal.pone.0076073

Gultekin, Y. B., Hildebrand, D. G. C., Hammerschmidt, K., & Hage, S. R. (2021). High plasticity in marmoset monkey vocal development from infancy to adulthood. *Science Advances*, 7(27), eabf2938. https://doi.org/10.1126/sciadv.abf2938

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, 6(1), 19680. https://doi.org/10.1038/srep19680

Ham, S., Hedwig, D., Lappan, S., & Choe, J. C. (2016). Song functions in non-duetting gibbons: Evidence from playback experiments on javan gibbons (Hylobates moloch). *International Journal of Primatology*, 37(2), 225–240. https://doi.org/10.1007/s10764-016-9897-x

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. *The evolution of communicative creativity: From fixed signals to contextual flexibility* ( pp. 93–119). MIT Press.

Hammerschmidt, K., Jürgens, U., & Freudenstein, T. (2001). Vocal development in squirrel monkeys. *Behaviour*, 138(9), 1179–1204. https://doi.org/10.1163/156853901753287190

Hammerschmidt, K., Newman, J. D., Champoux, M., & Suomi, S. J. (2000). Changes in rhesus macaque "coo" vocalizations during early development. *Ethology*, 106(10), 873–886. https://doi.org/10.1046/j.1439-0310.2000.00611.x

Hauser, M. D. (1989). Ontogenetic changes in the comprehension and production of vervet monkey (Cercopithecus aethiops) vocalizations. *Journal of Comparative Psychology*, 103(2), 149–158. https://doi.org/10.1037/0735-7036.103.2.149

Hedwig, D., Hammerschmidt, K., Mundry, R., Robbins, M. M., & Boesch, C. (2014). Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour*, 151(8), 1091–1120. https://doi.org/10.1163/1568539X-00003175

Hewitt, G., MacLarnon, A., & Jones, K. E. (2002). The functions of laryngeal air sacs in primates: A new hypothesis. *Folia Primatologica*, 73(2–3), 70–94. https://doi.org/10.1159/000064786

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, 28, 60–65. https://doi.org/10.1016/j.conb.2014.06.010

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), 1–11. https://doi.org/10.1006/anbe.2000.1410

Kalan, A. K. (2019). Evidence for sexual dimorphism in chimpanzee vocalisations: A comparison of male and female call production and acoustic parameters. In *The chimpanzees of the taï forest: 40 years of research* (pp. 410–421). Cambridge University Press. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_3034776

Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1740), 2974–2981. https://doi.org/10.1098/rspb.2012.0322

Koda, H., Lemasson, A., Oyakawa, C., Pamungkas, J., & Masataka, N. (2013). Possible role of mother-daughter vocal interactions on the development of species-specific song in gibbons. *PLoS ONE*, 8(8), 10. https://doi.org/10.1371/journal.pone.0071432

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, 67(5), 713–727. https://doi.org/10.1016/j.neuron.2010.08.038

Langmore, N. E. (2000). Why female birds sing. In Y. Espmark, T. Amundsen, & G. Rosenqvist (Eds.). *Animal signals. In signalling and signal design in animal communication.* ( pp. 317–327). Tapir Academic Press.

Laporte, M. N. C., & Zuberbühler, K. (2011). The development of a greeting signal in wild chimpanzees: Development of greeting signal in chimpanzees. *Developmental Science*, 14(5), 1220–1234. https://doi.org/10.1111/j.1467-7687.2011.01069.x

Leroux, M., Bosshard, A. B., Chandia, B., Manser, A., Zuberbühler, K., & Townsend, S. W. (2021). Chimpanzees combine pant hoots with food calls into larger structures. *Animal Behaviour*, 179, 41–50. https://doi.org/10.1016/j.anbehav.2021.06.026

Levréro, F., & Mathevon, N. (2013). Vocal signature in wild infant chimpanzees: Vocal signature in baby chimpanzees. *American Journal of Primatology*, 75(4), 324–332. https://doi.org/10.1002/ajp.22108

Lieblich, A. K., Symmes, D., Newman, J. D., & Shapiro, M. (1980). Development of the isolation peep in laboratory-bred squirrel monkeys. *Animal Behaviour*, 28(1), 1–9. https://doi.org/10.1016/S0003-3472(80)80002-9

Locke, J. L., & Snow, C. (1997). Social influences on vocal learning in human and nonhuman primates. *Social influences on vocal development* (pp. 274–292). Cambridge University Press. https://doi.org/10.1017/CBO9780511758843.014

Lonsdorf, E. V., Stanton, M. A., Pusey, A. E., & Murray, C. M. (2020). Sources of variation in weaned age among wild chimpanzees in Gombe National Park, Tanzania. *American Journal of Physical Anthropology*, 171(3), 419–429. https://doi.org/10.1002/ajpa.23986

Manser, M. B., Jansen, D. A. W. A. M., Graw, B., Hollén, L. I., Bousquet, C. A. H., Furrer, R. D., & le Roux, A. (2014). Vocal complexity in meerkats and other mongoose species. In *Advances in the study of behavior* (Vol. 46, pp. 281–310). Elsevier. https://doi.org/10.1016/B978-0-12-800286-5.00006-7

Marler, P., & Mitani, J. C. (1989). A phonological analysis of male gibbon singing behavior. *Behaviour*, 109(1–2), 20–45. https://doi.org/10.1163/156853989İ00141

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, 1(4), 381–385. https://doi.org/10.1098/rsbl.2005.0366

Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., Widdig, A., & Engelhardt, A. (2011). Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour*, 82(4), 911–921. https://doi.org/10.1016/j.anbehav.2011.07.016

Omedes, A. (1985). Infantile calls of silvery marmosets (Callithrix argentata melanura) during the first ten weeks. *Miscel·Lània Zoològica*, 9, 413–418.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, 106(51), 22026–22031. https://doi.org/10.1073/pnas.0908118106

Payne, K. (2000). The progressively changing songs of humpback whales: A window on the creative process in a wild animal. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music*, (pp. 135–150). Cambridge, Mass: The MIT Press.

Pusey, A. E. (1983). Mother-offspring relationships in chimpanzees after weaning. *Animal Behaviour*, 31(2), 363–377. https://doi.org/10.1016/S0003-3472(83)80055-4

Pusey, A. E. (1990). Behavioural changes at adolescence in chimpanzees. *Behaviour*, 115(3/4), 203–246. https://doi.org/10.1163/156853990X00581

Reddy, R. B., & Sandel, A. A. (2020). Social relationships between chimpanzee sons and mothers endure but change during adolescence and adulthood. *Behavioral Ecology and Sociobiology*, 74(12), 150. https://doi.org/10.1007/s00265-020-02937-7

Rilling, J. K., Glasser, M. F., Jbabdi, S., Andersson, J., & Preuss, T. M. (2012). Continuity, divergence, and the evolution of brain language pathways.

*Frontiers in Evolutionary Neuroscience*, 3, 3–11. https://doi.org/10.3389/fnevo.2011.00011

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11(4), 426–428. https://doi.org/10.1038/nn2072

Russell, A. F., & Townsend, S. W. (2017). Communication: Animal steps on the road to syntax? *Current Biology*, 27(15), R753–R755. https://doi.org/10.1016/j.cub.2017.06.066

Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, 122, 109–116. https://doi.org/10.1016/j.anbehav.2016.10.003

Schneider, R. M., Yurovsky, D., & Frank, M. C. (2015). *Large-scale investigations of variability in children's first words*. Cognitive Science, 2110–2115.

Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34(6), 1640–1658. https://doi.org/10.1016/S0003-3472(86)80252-4

Seyfarth, R. M., & Cheney, D. L. (2010). The ontogeny of vervet monkey alarm calling behavior: A preliminary report. *Zeitschrift Für Tierpsychologie*, 54(1), 37–56. https://doi.org/10.1111/j.1439-0310.1980.tb01062.x

Skeide, M. A., Brauer, J., & Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cerebral Cortex*, 26(5), 2127–2139. https://doi.org/10.1093/cercor/bhv042

Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nature Reviews Neuroscience*, 17(5), 323–332. https://doi.org/10.1038/nrn.2016.23

Slocombe, K. E., & Zuberbühler, K. (2005a). Agonistic screams in wild chimpanzees (pan troglodytes schweinfurthii) vary as a function of social role. *Journal of Comparative Psychology*, 119(1), 67–77. https://doi.org/10.1037/0735-7036.119.1.67

Slocombe, K. E., & Zuberbühler, K. (2005b). Functionally referential communication in a chimpanzee. *Current Biology*, 15(19), 1779–1784. https://doi.org/10.1016/j.cub.2005.08.068

Steventon, J. (2002). Cybertracker v. 3.284. *Cybertracker Software*.

Suzuki, T. N., & Zuberbühler, K. (2019). Animal syntax. *Current Biology*, 29(14), R669–R671. https://doi.org/10.1016/j.cub.2019.05.045

Taylor, D., Dezecache, G., & Davila-Ross, M. (2021). Filling in the gaps: Acoustic gradation increases in the vocal ontogeny of chimpanzees (Pan troglodytes). *American Journal of Primatology*, 83(5), e23249. https://doi.org/10.1002/ajp.23249

Team, R. C. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Version 4.0. 3*.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science*, 308(5730), 1934–1937. https://doi.org/10.1126/science.1108841

Townsend, S. W., Deschner, T., & Zuberbühler, K. (2008). Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS ONE*, 3(6), e2431. https://doi.org/10.1371/journal.pone.0002431

Van Lawick-Goodall, J. (1973). The behavior of chimpanzees in their natural habitat. *American Journal of Psychiatry*, 130(1), 1–12. https://doi.org/10.1176/ajp.130.1.1

Wittig, R. M. (2018). 40 years of research at the Taï chimpanzee project. *Pan Africa News*, 25(2), 16–18. https://doi.org/10.5134/236289

Zihlman, A. L., Bolter, D. R., & Boesch, C. (2007). Skeletal and dental growth and development in chimpanzees of the Taï National Park, Côte D'Ivoire. *Journal of Zoology*, 273(1), 63–73. https://doi.org/10.1111/j.1469-7998.2007.00301.x

Zuberbühler, K. (2019). Evolutionary roads to syntax. *Animal Behaviour*, 151, 259–265. https://doi.org/10.1016/j.anbehav.2019.03.006

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---

**How to cite this article:** Bortolato, T., Mundry, R., Wittig, R. M., Girard-Buttoz, C., & Crockford, C. (2022). Slow development of vocal sequences through ontogeny in wild chimpanzees (*Pan troglodytes verus*). *Developmental Science*, e13350. https://doi.org/10.1111/desc.13350