The Ontogeny of Vocal Sequences: Insights from a Newborn Wild Chimpanzee (*Pan troglodytes schweinfurthii*)

Adrian Soldati[1,2,3] · Geresomu Muhumuza[3] · Guillaume Dezecache[3,4] ·
Pawel Fedurek[3,5] · Derry Taylor[2,3] · Josep Call[1] · Klaus Zuberbühler[1,2,3]

Abstract

Observations of early vocal behaviours in non-human primates (hereafter primates) are important for direct comparisons between human and primate vocal development. However, direct observations of births and perinatal behaviour in wild primates are rare, and the initial stages of behavioural ontogeny usually remain undocumented. Here, we report direct observations of the birth of a wild chimpanzee (*Pan troglodytes schweinfurthii*) in Budongo Forest, Uganda, including the behaviour of the mother and other group members. We monitored the newborn's vocal behaviour for approximately 2 hours and recorded 70 calls. We categorised the vocalisations both qualitatively, using conventional call descriptions, and quantitatively, using cluster and discriminant acoustic analyses. We found evidence for acoustically distinct vocal units, produced both in isolation and in combination, including sequences akin to adult pant hoots, a vocal utterance regarded as the most complex vocal signal produced by this species. We concluded that chimpanzees possess the capacity to produce vocal sequences composed of different call types from birth, albeit in rudimentary forms. Our observations are in line with the idea that primate vocal repertoires are largely present from birth, with fine acoustic structures undergoing ontogenetic processes. Our study provides rare and valuable empirical data on perinatal behaviours in wild primates.

Keywords Birth · Call combinations · Pan troglodytes schweinfurthii · Vocal behaviour

---

Josep Call and Klaus Zuberbühler contributed equally to this work.

Handling Editor: Joanna M. Setchell

✉ Adrian Soldati
as391@st-andrews.ac.uk

Extended author information available on the last page of the article

Springer

Introduction

Primates communicate using a limited vocal repertoire, which largely develops in species-specific ways (Seyfarth & Cheney, 1997). The acoustic structure of calls uttered by infants typically resemble the corresponding adult call types, suggesting that vocal structures develop under strong genetic control (Hammerschmidt & Fischer, 2008; Janik & Slater, 2000; Owren *et al.,* 2011), with some room for socially acquired call variants (Ruch *et al.,* 2018; Snowdon, 2009). The acquisition of novel call types is virtually absent in wild primates (Fischer & Hammerschmidt, 2020; Tyack, 2020; but see Lameira, 2017). Overall, primate development of species-typical calls results from a combination of genetic, social, and environmental influences, though the relative role of each is still debated (Fedurek & Slocombe, 2011). Data on very early utterances shortly after birth are critical to assess the departure point in vocal ontogeny, prior to social and environmental influences. However, most primate births occur at night (Dunn, 2012) and are difficult to observe due to the unpredictability of parturition and maternal avoidance of other group members (Nishie & Nakamura, 2018; Otali & Gilchrist, 2006; Ramsay & Teichroeb, 2019), probably as a response to infanticide risk (Palombit, 2012). As a result, primate perinatal behaviours remain poorly understood (Trevathan, 2015), despite their theoretical relevance for developmental research (Nagy, 2011).

Vocal development has usually been analysed at three different levels: (1) production learning — how individuals modify specific acoustic features of calls after exposure to others' calls, (2) usage learning — how individuals give existing calls in new contexts or combine them as part of new vocal sequences, and (3) comprehension learning — how individuals respond appropriately to the vocalisations of others (Janik & Slater, 2000; Seyfarth & Cheney, 1997; Vernes *et al.,* 2021). In primates, production learning is regarded as mostly fixed, while usage and comprehension learning are considered more flexible (Seyfarth & Cheney, 1997; Snowdon, 2009). However, this model is largely based on studies of alarm calls, which are expected to be less flexible than calls with more social functions, which should instead be the focus when making comparisons with human vocal development (Elowson *et al.,* 1992; Snowdon *et al.,* 1997).

The majority of data on primate vocal development stem from studies of monkey vocalisations (Seyfarth & Cheney, 1997; Tomasello & Zuberbühler, 2002). However, monkeys are arguably less directly relevant to studies of language evolution than great apes, who share a more recent last common ancestor with humans (Langergraber *et al.,* 2012) (Fischer & Hage, 2019; Fitch & Zuberbühler, 2013). Great apes often produce sequences of calls (e.g., chimpanzees, *Pan troglodytes verus*: Girard-Buttoz *et al.,* 2022), including combinatorial structures (e.g., bonobos, *Pan paniscus*: Schamberg *et al.,* 2016; gorillas, *Gorilla gorilla beringei* and *Gorilla gorilla gorilla*: Hedwig *et al.,* 2014), with some evidence for socially-learned call variations (e.g., orangutan, *Pongo pygmaeus wurmbii* and *Pongo pygmaeus abelii*: Lameira *et al.,* 2022). While both humans and great apes use a limited set of sounds comparable in size (e.g., McComb & Semple, 2005;

Moran *et al.,* 2012), the ability to combine these sounds hierarchically to form vocal sequences varies greatly between humans and other apes, and sets apart human language from the communication of other animals (Hauser *et al.,* 2002; Townsend *et al.,* 2018). Although vocal learning abilities in great apes are clearly more constrained than in humans, investigating the degree to which great apes vocal sequences are socially learned or hard-wired, and thus present from birth, can inform us about the evolution of more complex vocal structures.

There are only a handful of direct observations of perinatal behaviour in our one of our two closest living relatives, the chimpanzees (Fujisawa *et al.,* 2016; Goodall & Athumani, 1980; Kiwede, 2000; Nishie & Nakamura, 2018; Zamma & Shabani, 2012). As a consequence, very little is known about newborn chimpanzee vocal behaviour, although subsequent stages of vocal development are somewhat better documented (e.g., Dezecache *et al.,* 2019; Laporte & Zuberbühler, 2011; Plooij, 1984; Taylor *et al.,* 2021). Early qualitative descriptions indicate that the first vocalisations of wild chimpanzees are comparable to the corresponding adult call types, such as grunts, whimpers, cries, and screams (Plooij, 1984). Human-reared chimpanzees initially exhibit vocal output that have some similarities to that produced by human infants in the first months of life (Kojima, 2008), although these are often elicited by human caretakers or researchers (Bard, 1998; Kojima, 2008). Human infants are special, however, in producing highly variable and functionally flexible vocal sequences, referred to as babbling — a form of vocal exploration considered a milestone during language acquisition (Oller, 2000; Oller *et al.,* 2021). Typically, babbling starts soon after birth, consists of a subset of the acoustic features characterising the adult repertoire, and does not require a social context or to be communicative (Oller, 2000; ter Haar *et al.,* 2021). In addition to simple vocal practice, one probable function of this peculiar behaviour is to enhance social interactions and bonding with caregivers (Locke, 2006; Oller & Griebel, 2008). However, evidence for babbling-like vocal behaviour is absent in chimpanzees (Oller *et al.,* 2019; ter Haar *et al.,* 2021).

The vocal repertoire of wild chimpanzees consists of a relatively small number of acoustically distinct call types that can grade into each other (Crockford, 2019; Goodall, 1986; Marler & Tenaza, 1977). Different call types often appear in sequences (Crockford & Boesch, 2005; Girard-Buttoz *et al.,* 2022; Leroux & Townsend, 2020), such as pant hoots and food grunts (Leroux *et al.,* 2021) or screams and barks (Fedurek *et al.,* 2015). Whether such call combinations function to convey different information is still unclear and a topic of ongoing research (Engesser & Townsend, 2019; Zuberbühler & Lemasson, 2014). Furthermore, no study to date has investigated how and when the capacity to produce vocal sequences appears during chimpanzee ontogeny.

One vocalization commonly produced by chimpanzees, the pant hoot, comprises smaller vocal components (i.e., phases) produced in an orderly sequence of introduction, build-up, climax, and let-down (Marler & Hobbett, 1975; Marler & Tenaza, 1977). Pant hoot phases, in turn, consist of a varying number of voiced exhalations, which are the smallest units of this vocal sequence and are separated by short periods of silence or panted inhalations (Fedurek *et al.,* 2017). Pant hoots are flexibly produced across many contexts suggesting various functions, including coordinating

fission–fusion dynamics (Fedurek *et al.,* 2014), signalling individual and group identity (Crockford *et al.,* 2004; Mitani *et al.,* 1996), signalling social bonds (Fedurek, Machanda, *et al.,* 2013a; Mitani & Brandt, 1994), or signalling social status (Clark & Wrangham, 1994; Fedurek, Slocombe, *et al.,* 2016a). Some phases can be omitted (Fedurek, Zuberbühler, & Dahl, 2016b; Notman & Rendall, 2005) or produced in isolation (Soldati *et al.,* 2022). In addition, phases are sometimes regarded as equivalent to distinct call types or sub-types within the vocal repertoire (e.g., climax as panted scream: Crockford, 2019; Girard-Buttoz *et al.,* 2022). Although pant hoots are amongst the most common and most studied vocalisation in wild chimpanzees (Marler & Tenaza, 1977), data are mainly from sub-adult and adult males (e.g., Crockford *et al.,* 2004; Fedurek *et al.,* 2014; Mitani & Brandt, 1994), and pant hoots are only rarely uttered by immature individuals from around 2 years of age (Hiraiwa-Hasegawa, 1986), with the rate of production increasing with age (Marler & Tenaza, 1977; Pusey, 1990). Importantly, to our knowledge, newborns have never been reported uttering pant hoots, and no study has systematically investigated the development of this vocal sequence.

In this study, we report on the vocal behaviour of a new-born wild chimpanzee in the Sonso community of Budongo Forest, Uganda. We used two methods of call classification: (1) qualitative spectrographic and auditory categorisation of calls supplemented by auditory categorisation by human experts, and (2) quantitative soft clustering analysis to determine distinct acoustic clusters and discriminant function analyses to investigate pant hoot production across age categories.

Methods

Study Site and Population

We studied the Sonso chimpanzee community in Budongo Forest, western Uganda. Chimpanzees from the Sonso community have been studied and followed daily by field assistants since 1990 (Reynolds, 2005). At the time of the study, the community was composed of 71 individuals, including nine adult males and 31 adult females (Table SI Supplementary Material). The main individuals involved in the study were members of the Kutu family (Table I).

Table I Key Chimpanzee Individuals of the Kutu Family Involved in the Birth of KU7 at Sonso, Budongo Forest, Uganda, in November 2019. Kinship Status Refers to the Relation Between KU and Her Offspring

Name	ID	Age & sex class	Age (years)	Kinship status (birth order)
Kutu	KU	Adult female	40	Mother
N/A	KU7	Newborn male	0	Son (7th)
Kefa	KF	Juvenile male	5	Son (6th)
Kathy	KH	Sub-adult female	11	Daughter (4th)
Kato	KT	Adult male	26	Son (1st)

Data Collection

On the 20th of November 2019 at 10:12 am, the first and second authors observed the birth of KU7. Three additional researchers attended part of the afterbirth period and assisted with data collection and identification of callers. We collected audio recordings of all vocalisations produced by KU7 and collected the vocalisations produced by other individuals in the party opportunistically with a directional microphone, the Sennheiser MKH416 (Sennheiser Electronic GmbH & Co. KG, Wedemark, Germany) with a Marantz PMD661 MkII (Marantz, Kanagawa, Japan) solid-state recorder (sample rate 44.1 kHz, resolution 32 bits, 'wav' format). We defined party composition as all individuals present within a radius of approximately 35 m of the focal individual (Newton-Fisher, 1999). We set the recorder's gain on level 9 to maximise the signal/noise ratio due to the softness of calls and the distance from the subject (approx. 15 m). We maintained this distance due to the delicate nature of the event and to reduce any effects of our presence on the chimpanzees' behaviour. We dictated observations to the microphone or noted them using CyberTracker (ver. 3.496) on a Samsung Xcover 4 portable device (Samsung Group, Seoul, South Korea). We recorded videos using a Panasonic VHC-770 HD (resolution: 1920*1080/50p). We recorded all relevant events and changes in the behaviour of all individuals in the party, and recorded the composition of the party continuously.

Qualitative Acoustic Analysis

We inspected audio recordings to extract vocalisations using spectrograms generated with Praat software (ver. 6.0.42) and Sennheiser HD650 headphones. We transformed calls with the Fourier function using a Hanning window function and 1024 time steps. Four authors independently categorised call types (Table II) based on auditory features and inspection of spectrograms using published chimpanzee vocal repertoires composed of nine call types (Table III). If one of the four authors disagreed with the categorisation, we used group majority to determine the

Table II Terms Used to Describe the Vocalisations Produced by a Newborn Chimpanzee in the Sonso Community, Budongo Forest, Uganda, in November 2019.

	Definition	Reference
Call type	Vocalisation belonging to the species' repertoire.	See Table III
Call unit	Distinct emitted vocal sound separated by a silent gap before and after.	Kershenbaum *et al.,* 2016
Single call	One call unit of any type separated from other calls by more than two seconds of silence.	Kershenbaum *et al.,* 2016
Call sequence	Series of two or more call units belonging to either the same or different call type and separated by maximum two seconds of silence.	Hedwig *et al.,* 2014
Call combination	Call sequence consisting of two or more units of different types.	Crockford & Boesch, 2005

Table III Call Types Produced by Chimpanzees Based on Marler and Tenaza (1977), Slocombe and Zuberbühler (2010), and Taylor *et al.* (2021). The Definition of Pant Hoot is Based on Marler and Hobbett (1975) and Notman and Rendall (2005). Example Spectrograms for Each Call Type Produced by Chimpanzees in the Sonso Community, Budongo Forest, Uganda, are Presented in Slocombe and Zuberbühler (2010)

Call type	Definition
Bark	Short, loud, and noisy call with abrupt onset. Generally low-frequency.
Grunt	Short and soft call that can be either tonal or noisy. Generally low-frequency and produced with variable rhythm.
Hoo	Short and tonal call with highest frequency and amplitude at the start then decreasing over the duration of the call. The call can be either soft (e.g., rest hoo) or loud (e.g., alarm hoo).
Laughter	Short, soft, and noisy sounds produced while inhaling and exhaling. Generally low-frequency and produced with irregular rhythm.
Pant	Short, soft, and low-frequency unvoiced sounds.
Scream	Loud and high-frequency call that can be either noisy or harmonic. Acoustic energy is usually present during exhalation.
Squeak	Short, high-frequency, and clear tonal call.
Whimper	Soft, low-frequency and tonal call similar to "hoo" calls. Units can increase in frequency within a sequence.
Pant hoot	Vocal sequence composed of up to four acoustically distinctive phases produced in this order and composed of:
Introduction	Low-frequency and tonal calls that acoustically resemble "hoo" calls but are longer in duration and alternated with inhaled tonal elements.
Build-up	Short and low-frequency calls produced both during inhalations and exhalations in rapid rhythm. Intensity and frequency typically increase over the phase.
Climax	High-amplitude and high-frequency calls that resemble "scream" calls.
Let-down	Short and low-frequency acoustic calls which acoustically resemble the build-up phase but with decreasing intensity and frequency.

call type. There were no instances where more than one author disagreed with the categorisation.

To provide a more comprehensive and diverse assessment of the call types, we asked seven independent experts in chimpanzee vocal communication (Table SII Supplementary Material), blind to any information about the recordings, to categorise the calls recorded from KU7, estimate the age of the caller, and comment on the vocal structures. We provided an unlabelled audio file in which we collated all the calls produced by KU7 in chronological order, with sequences separated by 1 s of silence (Online Resource 1).

Quantitative Acoustic Analysis

We manually extracted six acoustic features from each call unit using Praat software (ver. 6.0.42): duration of each exhaled unit, fundamental frequencies (F0) at the start, middle, and end of the unit, maximum and minimum F0, and range of the F0. We selected these features based on the acoustic data extractable from the recordings

and on measurements typically considered when determining call types in chimpanzees (e.g., Marler & Tenaza, 1977; Mitani *et al.,* 1999; Mitani & Brandt, 1994; Slocombe & Zuberbühler, 2010). We only considered exhaled vocal units to make our acoustic analyses comparable with previous studies of pant hoots (e.g., Clark & Wrangham, 1993; Desai *et al.,* 2021; Fedurek *et al.,* 2013a, b, Fedurek *et al.,* 2017; Mitani *et al.,* 1992, 1999; Riede *et al.,* 2007; but see Crockford *et al.,* 2004), but also noted the number of inhaled (panted) units produced between exhaled units when these were visible on the spectrogram. Because of the quiet nature of the newborn vocalisations, the presence of environmental background noise, and the distance between the newborn and the microphone, we could not use automated procedures to extract acoustic features.

Clustering Analysis

The general approach to studying how experience mediates vocal development is to catalogue the different call types across developmental stages, using acoustic measurements and classification algorithms (Bradbury & Vehrencamp, 2011; Kershenbaum *et al.,* 2016). A common problem is that vocal repertoires are often graded, making objective classifications particularly challenging (Fischer *et al.,* 2017). However, human vocal behaviour is also highly graded, but receivers still perceive transitions in categorical ways, suggesting that human perceptual judgements can be used to disambiguate gradual transitions (Deecke & Janik, 2006; Janik, 1999). For animal vocal repertoires, data-driven categorisation approaches are preferable, mainly because the degree to which human perceptual bias reflects that of other species remains unclear, and because they allow systematic comparisons across communities (Crockford, 2019; Fischer *et al.,* 2017). Soft clustering methods based on fuzzy-set theory (Zadeh, 2008) are very suitable to describe graded vocal repertoires of primates (e.g., chacma baboons, *Papio ursinus*: Wadewitz *et al.,* 2015), an approach that is also promising for chimpanzees (e.g., immature chimpanzees: Taylor *et al.,* 2021).

We used fuzzy c-means clustering to identify the best fitting model for the number of clusters representing different call types in the newborn vocalisations. The fuzzy c-means algorithm measures the degree to which sounds belong to categories based on their acoustic proprieties without restricting them to a single category, capturing more details than hard clustering methods, including the graded transition between call types. We analysed the stability and reliability of model solutions to evaluate the extent to which the optimal description of the calls depended on a small number of acoustic parameters, and how robust optimal descriptions were to overlap between clusters. We z-transformed the acoustic features prior to analysis to prevent the influence of measurements with different scales (i.e., Hz and s) on cluster solutions. Since fuzzy c-means clustering is based on the individual acoustic features of each call instead of the total number of calls available (Wadewitz *et al.,* 2015), the small number of newborn vocalisations we recorded was not a limiting factor because a minimum number of data points for each call type is not required.

We adjusted two parameters to identify the best cluster solution to describe the newborn calls: the maximum number of clusters extracted (K), and the 'fuzziness parameter' (μ) which limits the degree of overlap between clusters (i.e., lower values allow less overlap between clusters). We ran fuzzy models using the "fanny" implementation in the "cluster" package (ver. 2.1.2, Maechler *et al.,* 2021) varying K values from a minimum of two (required to quantify gradation) and a maximum of seven, which matched the number of call types we assessed qualitatively and was in line with Taylor *et al.* (2021). We varied μ values starting at 1.1 with increments of 0.5 following Taylor *et al.* (2021) and stopped at 3.0 when all membership coefficients were too close to 1/K, which corresponds to the limit of the algorithm to assign cluster membership to calls (Zadeh, 2008). All models considered converged within 500 iterations. We evaluated the fit and confidence of each solution based on the mean silhouette value of all data points combined, which represents how separable the acoustic clusters are. Silhouette values range from −1 to 1, with positive values representing data points that are closer to their primary cluster and indicate some degree of confidence with regard to their cluster membership, while negative values represent datapoints that overlap between clusters and are potentially misclassified (Wadewitz *et al.,* 2015). We assessed the reliability of the model by looking at the range of μ values obtained for any given K value, which provided an indicator of 'gradedness'. Solutions for which low and high μ values can be extracted are regarded as more robust to overlap between clusters (Fischer *et al.,* 2017). Using the "clValid" package (ver. 0.7, Brock *et al.,* 2008) we assessed the stability of the clusters by calculating four measurements that compare the result of the clustering algorithm by systematically removing one variable at a time and measuring how much the clusters are based on a small number of acoustic parameters, which represented how 'generalisable' the cluster separations are. The four measurements we used are: the mean proportion of non-overlap between data points (APN), the mean distance between data points in the same cluster (AD), the mean distance between the cluster's centre and the data points in the same cluster (ADM), and the mean variance of data points in the same cluster (FOM) (Brock *et al.,* 2008). Given that our aim was to categorise calls into clusters, we gave priority to mean silhouette values to identify the best model. We extracted a hard-clustering solution for the best-fitting model and assigned all the calls to their primary cluster membership. We then examined the distribution of qualitatively categorised call types in each cluster.

Comparison with Pant Hoots

We conducted additional acoustic analyses to compare vocal sequences produced by the newborn which contained calls resembling pant hoot phases with pant hoots from infant, juvenile, sub-adult, and adult males of the Sonso community. We selected pant hoots produced during resting or feeding events, because pant hoots vary depending on the behavioural context of production (Fedurek, Zuberbühler, & Dahl, 2016b; Notman & Rendall, 2005), and the newborn vocalised while resting or potentially before/after nursing. Although pant hoot sequences can be composed of repeated vocal units from a single phase (Soldati *et al.,* 2022), we selected calls composed of two or more

phases to be consistent with previous studies (e.g., Fedurek *et al.,* 2014; Mitani *et al.,* 1999; Notman & Rendall, 2005). To control for potential differences between the sexes in the acoustic structure of pant hoots (e.g., Holden, 2017), we only selected male pant hoots. We selected recordings based on their overall quality (lack of background noise or overlap with other callers) and good signal-to-noise ratio. Although these recordings were of higher quality than recordings of the newborn, we extracted the acoustic measurements manually in the same way to avoid introducing a potential bias. We sampled the first and the middle vocal units of the introduction and climax phases for four units from each pant hoot. This allowed us to take into consideration the acoustic gradation that can occur within phases. Where there were an even number of units, we chose the first of the two middle units (as in Desai *et al.,* 2021). In total, we extracted features from 189 vocal units (42 pant hoots) produced by three infants, four juveniles, four sub-adults, and four adults, with a minimum of two pant hoots per individual (Table SVIII).

To determine whether the acoustic structure of Phase 1 and 2 calls produced by the newborn differed from the introduction and climax phases produced by infant, juvenile, sub-adult, and adult individuals from the Sonso community, we used permuted discriminant function analyses (pDFA; Mundry & Sommer, 2007), following previous studies (e.g., Leroux *et al.,* 2021; Soldati *et al.,* 2022). To analyse the introduction phase, we used 19 vocal units from 11 calls produced by four adults, 21 units from 11 calls produced by four sub-adults, six units from four calls produced by three juveniles, and 19 units from ten calls produced by four infants. Together with the Phase 1 calls ($n = 26$), we obtained a total of 91 calls. To analyse the climax phase, we used 17 units from 11 calls produced by four adults, 20 units from 11 calls produced by four sub-adults, 18 units from nine calls produced by four juveniles, and ten units from ten calls produced by three infants. Together with the Phase 2 calls ($n = 18$), we obtained a total of 93 calls. Before analysis, we assessed multicollinearity to avoid including correlated acoustic parameters. We removed at a time the parameter with the highest variance inflation factor (VIF) using the 'performance' R package (version 0.8.0, Lüdecke *et al.,* 2021) until we obtained a set of variables with low correlation. In the final set of four variables (Duration, Start F0, End F0, Range F0), the highest VIF for introduction calls was 2.85 and the highest VIF for climax calls was 1.63. We assessed the distribution of the data, and when variables were not normally distributed and this could be improved, we applied a log or squared-root transformation. We then used nested pDFA with 1000 permutations to test whether the acoustic structure of newborn's Phase 1 & 2 calls differed significantly from the corresponding phases produced by the other age categories (Mundry & Sommer, 2007). In comparison with a conventional DFA, a pDFA allows the inclusion of repeated data points per individual and controls for unbalanced data sets at the same time. We included the 'ID' of the caller as a control factor.

We conducted all statistical analyses in R (version 4.1.2, R Core Team, 2021).

Ethical Note

Data collection was entirely observational, adhering to the ASAB guidelines for the treatment of animals during behavioural studies (Association for the Study of Animal Behaviour, 2018). The study was approved by the Uganda Wildlife Authority

(UWA/COD/96/5) and the Uganda National Council for Science and Technology (NS 637). The research ethics committees of the University of Neuchâtel (38/2019-B) and University of St Andrews (No 171) also approved this project. We evaluated the scope for bias in our study subjects using the STRANGE framework (Webster & Rutz, 2020) (see Supplementary Material). The authors declare that they have no conflict of interest.

Data Availability Data on the newborn vocal and perinatal behaviours generated or analysed during this study are included in this article and its supplementary information files.

Results

A detailed report of the birth is available in the Supplementary Material. Video and audio recordings are available as Online Resources (1, 3–5).

Qualitative Call Classification

We recorded 70 call units from the newborn during 2 hours and 15 minutes of observation (0.5 per minute). These calls were divided into 12 separate vocal occurrences (also referred as 'utterances'; call rate 0.1 per minute), of which three were single calls and nine were call sequences (see Online Resource 2 for the acoustic spectrograms). Vocal sequences contained a mean of 7.4 vocal units (range 2–17). We identified barks ($n = 2$), grunts ($n = 8$), hoos ($n = 6$), squeaks ($n = 2$), whimpers ($n = 8$), and units that we labelled as part of a pant hoot ($n = 44$). We distinguished two variants which we refer to as "pant hoot phase 1" ($n = 26$), hereafter Phase 1 for brevity, and "pant hoot phase 2" ($n = 18$), hereafter Phase 2 (Fig. 1). The four vocal sequences that included Phase 1 or Phase 2 were composed of a mean of 11 vocal units (range 5–15) when excluding other call types. Of all the newborn's Phase 1 and 2 units, 36% ($n = 15$) included panted units between exhaled units (Fig. 1).

Five of seven independent experts agreed with our decision to classify KU7's Phase 1 and Phase 2 calls as resembling adult pant hoots (Table SIV Supplementary Material). One expert classified the calls as either pant hoots or whimpers, and one expert classified the calls as whimpers. The experts did not reliably classify the other call types (barks, grunts, hoos, squeaks, and whimpers) citing challenging conditions (soft signal volume and background noises), but two experts reported the presence of quiet hoos and grunts among these other call types. Three experts classified the caller as a young individual, one as a juvenile, one as an infant, one as immature, and one as either a juvenile or a young adult.

Quantitative Analyses

We extracted six acoustic features from 70 call units and used them to model the best clustering of acoustically similar units using fuzzy analyses (Table IV). Overall, we
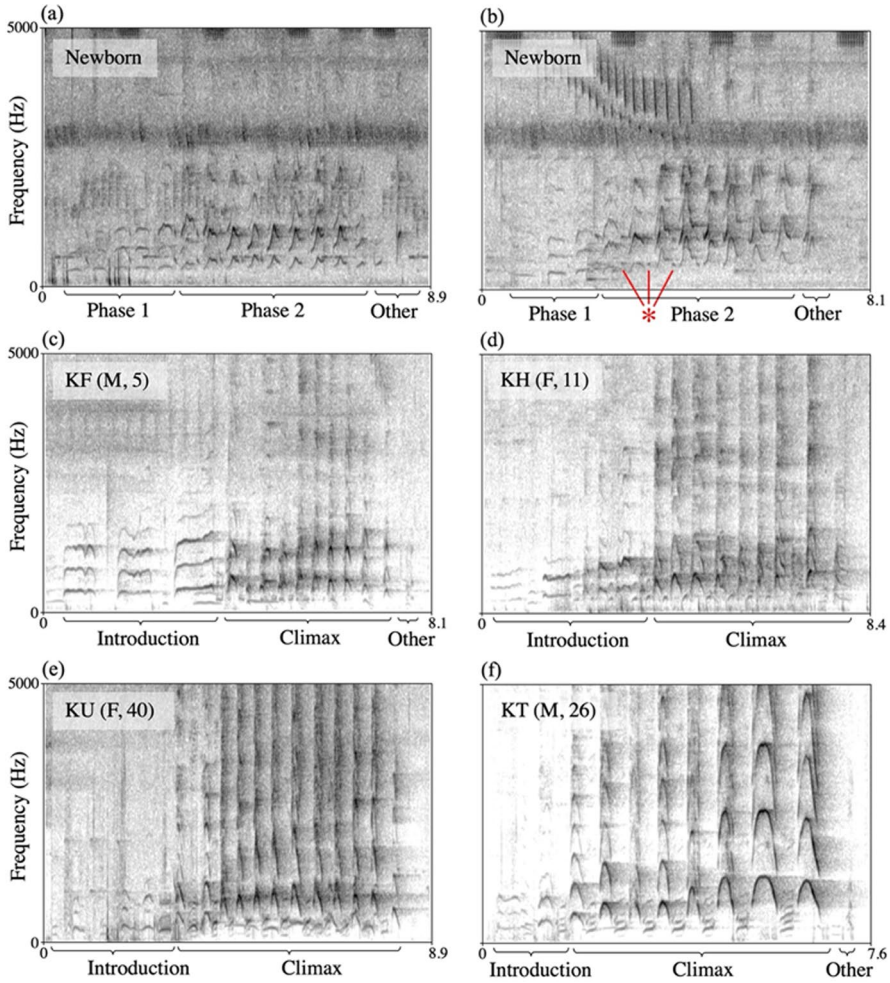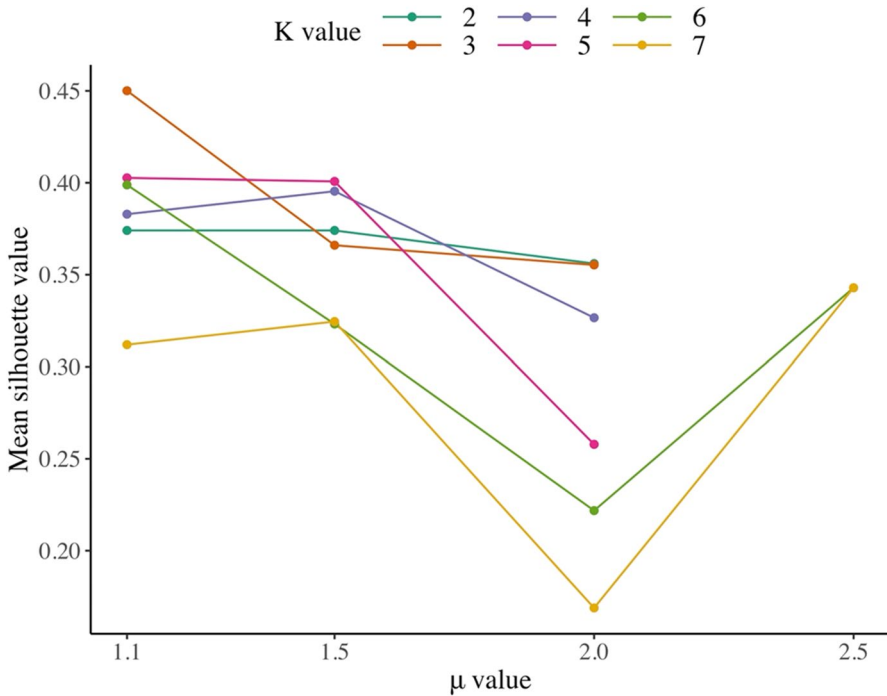
Fig. 1 Spectrographic representations of pant hoot calls produced by a newborn chimpanzee in the Sonso community, Budongo Forest, Uganda on 20 November 2019 with pant hoots produced by members of the mother's family for comparison [sex and age (years) are shown]. For each call, the different phases or types are indicated underneath (*Other* = other call type). The *red asterisk* indicates three examples of panted units. Duration (s) on the *x*-axis. Note the presence of bird songs and cicada sounds above frequencies of approx. 2500 Hz in spectrograms **a** and **b**.

obtained 20 unique models varying between two to seven clusters (K) and fuzziness parameters (μ) of 1.1, 1.5, 2.0, and 2.5 (Fig. 2). Two-, three-, four-, and five-cluster solutions could be calculated up to μ = 2.0, while six- and seven-cluster solutions could be calculated up to μ = 2.5. The model that best fit our data was calculated with three clusters (K = 3) and with a fuzziness parameter of 1.1 (Fig. 2), scoring the highest mean silhouette value of 0.450, which indicates confidence in the overall solution. Although this model was not the most stable, only 6% of call units (*n*

Table IV Mean ± SD of Extracted Parameters for Call Types Produced by a Newborn Chimpanzee in the Sonso Community, Budongo Forest, Uganda on 20 November 2019. Table SIII (Supplementary Material) Contains Extracted Data from Each Call Unit Separately

Type	N	Duration (s)	Start F0 (Hz)	Middle F0 (Hz)	End F0 (Hz)	Max F0 (Hz)	Min F0 (Hz)	Range F0 (Hz)
Bark	2	0.09 ± 0.03	545.9 ± 11	537.0 ± 30	518.4 ± 49	550.1 ± 15	518.4 ± 49	31.7 ± 33
Grunt	8	0.06 ± 0.02	256.1 ± 135	255.4 ± 122	251.7 ± 122	264.8 ± 131	248.5 ± 118	16.3 ± 16
Hoo	6	0.30 ± 0.14	338.2 ± 29	349.7 ± 16	353.1 ± 30	358.6 ± 25	333.9 ± 27	24.7 ± 22
Phase 1	26	0.32 ± 0.06	356.6 ± 59	408.2 ± 83	396.5 ± 84	421.3 ± 86	354.8 ± 60	66.5 ± 40
Phase 2	18	0.29 ± 0.05	386.7 ± 55	624.0 ± 184	504.0 ± 110	646.6 ± 192	386.4 ± 55	260.2 ± 146
Squeak	2	0.15 ± 0.08	408.0 ± 133	849.6 ± 180	936.1 ± 108	1025.6 ± 17	408.0 ± 133	617.7 ± 115
Whimper	8	0.10 ± 0.05	253.7 ± 38	272.3 ± 48	260.7 ± 53	272.3 ± 47	246.7 ± 45	25.6 ± 19

Fig. 2 Mean silhouette values obtained by varying the number of clusters (K = 2 to 7) and fuzziness values (μ = 1.1 to 2.5) using fuzzy c-means clustering. Mean silhouette values measure the confidence of the overall cluster solution of calls produced by a newborn chimpanzee in the Sonso community, Budongo forest, Uganda on 20 November 2019; the higher the silhouette values, the more distinct the acoustic clusters are and the better the model fits the data.

= 4) changed membership when we recalculated the model with one less variable (Table SV and Fig. S1 Supplementary Material). In the most stable model (K = 2, μ = 1.1) 2% of call units (*n* = 1.3) changed membership, but this model had a mean silhouette value of 0.374, which is considerably lower than the best fitting model. Six- and seven-cluster solutions could be calculated for a larger range of fuzziness values, suggesting they might be more reliable. However, these solutions were less consistent in mean silhouette value (range: 0.177 for six clusters, 0.174 for seven) than three-cluster solutions (range: 0.095) (Table SVI Supplementary Material). Furthermore, six and seven-cluster solutions had lower mean silhouette values (0.399 and 0.343 respectively) than the best fitting model with a three-cluster solution.

Cluster Composition

We calculated the percentage of each call type (determined qualitatively) that belonged to each of the three clusters in the best fitting model identified by the quantitative analysis (Fig. S2 Supplementary Material). The first cluster was composed of grunts and whimpers (Table V). The second cluster consisted of barks, squeaks,

Table V Percentage and Number (in Brackets) of Calls per Call Type in Each Cluster Produced by a Newborn Chimpanzee in the Sonso Community, Budongo Forest, Uganda on 20 November 2019. See Table SVII in the Supplementary Material for Within-Cluster Call Type Percentages

	Grunt	Bark	Hoo	Phase 1	Phase 2	Squeak	Whimper	Total
Cluster 1	88 (7)	-	-	-	-	-	100 (8)	15
Cluster 2	12 (1)	100 (2)	-	8 (2)	33 (6)	100 (2)	-	13
Cluster 3	-	-	100 (6)	92 (24)	67 (12)	-	-	42
Total	8	2	6	26	18	2	8	70
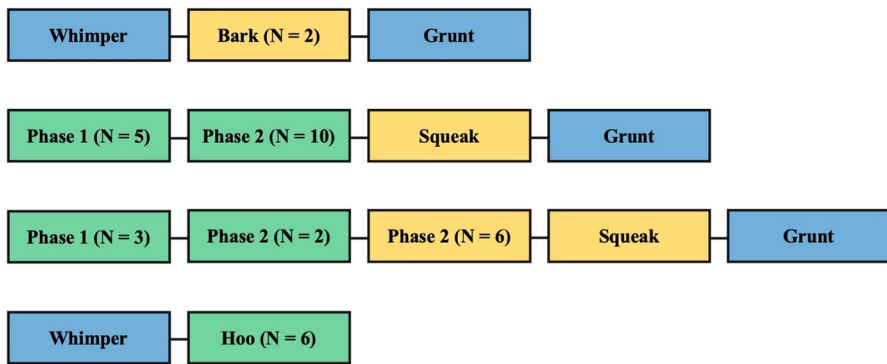
Fig. 3 Four vocal sequences composed of different call types produced by a newborn chimpanzee in the Sonso community, Budongo Forest, Uganda on 20 November 2019. Each series of connected 'blocks' represents one of the four vocal sequences. Each 'block' indicates the call type and the number of repeated call units of the same call type. Colours represent clusters (Cluster 1 = blue; Cluster 2 = yellow; Cluster 3 = green).

Phase 2 calls, Phase 1 calls, and grunts (Table V). The third cluster consisted of hoos, Phase 1 calls, and Phase 2 calls (Table V).

Call Combinations

KU7 produced single calls in three instances (grunts only) and nine different call sequences (range: 2–17 units). Of the nine sequences, four were combinations of units from different call types (Fig. 3). Overall, calls from two or three different clusters were produced in a single combinatorial structure, and two to four different calls were combined in a structure.

Comparison with Pant Hoots

Pant hoots were composed of a mean of 15 vocal units in infants (range 1–18), 12.6 units in juveniles (range 1–12), 10.2 units in sub-adults (range 1–15), and 8.1 units in adults (range 1–15). A panted unit followed 43% of vocal units produced by infants ($n = 21$), 59% of units in juveniles ($n = 20$), 73% of units in sub-adults ($n = 37$), and

94% of units in adults ($n = 44$). Vocal usage rate varied with phase and age category (Table SIX). While introduction, build-up, and climax phases were observed across all age categories, the let-down phase was not observed in infants and juveniles (Table SIX). Furthermore, the let-down phase was rarely produced by sub-adults and adults, and the build-up phase was rarely produced by adults (Table SIX).

The discriminant function could not operate when including the introduction (produced by infant, juvenile, sub-adult, and adult males) and the Phase 1 calls produced by the newborn. This was likely because the within-group variance of the variables was lower than the level accepted by the function, which might indicate that variables are collinear or constant (Venables & Ripley, 2002). Because the variables we considered were not collinear and since pant hoots have mainly been studied in sub-adult and adult males, we repeated the analysis including only individuals from these age categories. The results are not compatible with the idea that calls are acoustically different in newborn vs sub-adult and adult males (expected correctly cross-classified: 42%, $p = 0.257$). For the climax and Phase 2 calls, the results are also not compatible with the idea that calls are acoustically different in newborn vs infant, juvenile, sub-adult, and adult males (expected correctly cross-classified: 26%, $p = 0.164$). We repeated the analysis including only the newborn, sub-adult, and adult individuals, and found similar non-significant results (expected correctly cross-classified: 40%, $p = 0.266$).

Discussion

We qualitatively discriminated seven call types from the 70 units produced by a newborn chimpanzee immediately after birth, in line with what was previously reported for older chimpanzee infants (Plooij, 1984). The majority of units were given as part of sequences. Interestingly, the newborn also produced vocal structures resembling pant hoots. Quantitative analyses revealed three acoustically distinct clusters of calls, with calls from different clusters combined into the same sequence. All call types were also produced in isolation, with the exception of Phase 2 (pant hoot) calls and squeaks, which were only produced in combination with other calls. From these data we concluded that chimpanzees have the capacity to combine some call types into larger structures from birth.

We can suggest four hypotheses explaining what could trigger the newborn's vocal production. First, the newborn produced a series of long vocal sequences to attract the mother's attention, for example due to discomfort or desire to be nursed. However, we did not notice any of the more typical calls for such contexts (i.e., cries and whimpers; Dezecache *et al.,* 2021) and the newborn was always in bodily contact with the mother during our observations, making this explanation unlikely.

Second, the newborn's vocal behaviour may represent a rudimentary form of babbling (Oller, 2000), which serves as vocal practice and to elicit care-giving (ter Haar *et al.,* 2021). However, the newborn produced vocal sequences composed of repeated vocal units that were rhythmically produced and only contained a subset of adult calls. Furthermore, the sequences lacked variable acoustic structures, did not elicit vocal or social responses from the mother, were produced at lower call rates

than human and marmoset infants (Elowson *et al.,* 1998a, 1998b; Oller *et al.,* 2021; Snowdon & Elowson, 2001) and at comparable rates with chimpanzee and bonobo infants (Kojima, 2008; Oller *et al.,* 2019; Taylor, 2020), making the babbling hypothesis an unlikely explanation. In marmosets, one of the few primates where babbling has been reported, parents engage in vocal feedback and exchanges (Takahashi *et al.,* 2015), while chimpanzee mothers rarely direct vocalisations to their offspring (Schick *et al.,* 2022). While language-trained or human-raised chimpanzees have been reported to produce vocalisations similar to babbling when interacting with researchers (Hayes & Hayes, 1951, pp. 106–108), the behaviour is different from human infant babbling in terms of variety, quantity, and duration (Kellogg, 1968; Kojima, 2008). In contrast, the newborn vocalisations presented some similarities with protophone-like sounds produced by young bonobo infants (Oller *et al.,* 2019). These sounds are regarded as akin to the exploratory protophones produced by human infants since birth during low- to moderate arousal contexts, without requiring social stimulation, and prior to babbling (Oller, 2000; Oller *et al.,* 2016). However, the newborn chimpanzee produced sounds at much lower rates than humans, without clear signs of playfulness, and without interacting vocally with the mother, all of which also characterise vocal behaviours in infant bonobos (Oller *et al.,* 2019). One possibility is that the period of vocal exploration in great apes is very reduced and limited to the earliest developmental phase, although further observations are necessary to test this possibility.

Third, the newborn vocal behaviour may have been an artefact resulting from limited vocal control. Interestingly, Phase 1 calls were always followed by Phase 2 calls in sequences, but never the other way around, and terminated with squeaks followed by grunts. This call order is akin to that of pant hoots, in which introduction units are followed by climax units, and akin to how pant hoots tend to be followed by food grunts in call combinations (Leroux *et al.,* 2021). These observations do not support the artefact hypothesis, although there may have been anatomical constraints on vocal production. For instance, in adult pant hoots, the climax is never produced in isolation but is always preceded by an introduction or build-up phase (Soldati *et al.,* 2022), perhaps because producing high-pitch and high-amplitude calls requires more time and effort (Riede *et al.,* 2007).

Finally, intra-uterine auditory exposure to conspecifics' calls may have affected the newborn's vocal development. While we cannot address this hypothesis with our data, it has been documented in marmosets (Narayanan *et al.,* 2022), humans (Gervain, 2018; Varga *et al.,* 2019), and songbirds (Colombelli-Négrel *et al.,* 2021), all of which are regarded as vocal learners (Vernes *et al.,* 2021). Further studies are needed to clarify the effects of pre- and post-natal auditory exposure on the development of great ape vocalisations.

The most puzzling aspect of the newborn's vocal behaviour was the presence of vocal structures that acoustically and visually resembled chimpanzee pant hoots, with clear resemblance to adult as well as infant and juvenile pant hoots. Most experts rated these structures as pant hoot attempts. Although some experts rated them as whimpers, cluster analyses revealed that they did not belong to the whimper cluster. Phase 1 calls closely resemble the introduction phase and Phase 2 calls closely resemble the climax phase, which was supported by the results of the

discriminant analyses. All the hoo calls produced by the newborn belonged to the Phase 1 cluster, in line with the idea that the introduction can be seen as a variant of hoos (Crockford, 2019). The production of panted units, a characteristic feature of pant hoots, followed an incremental pattern from the newborn through all age categories, suggesting that it develops during ontogeny. In addition, we agree with the experts, who pointed out that the overall 'rhythmicity' of the newborn sequences is characteristic of pant hoots. Thus, the newborn produced vocal structures resembling pant hoots, suggesting that such complex structures are part of the innate vocal repertoire of chimpanzees. However, the newborn utterances differed from adult pant hoots in terms of absence of phases that correspond to the build-up and let-down. Alternating phases with a relatively low rate of unit production, such as the introduction and climax, with phases exhibiting a fast-paced and panted unit production, such as the build-up and let-down, is a key feature of pant hoots.

The differences between the arguably rudimentary form of the newborn pant hoot and adult pant hoots suggest that this call type undergoes some ontogenetic processes. Immature individuals might learn to produce certain phases in specific contexts or as part of structurally varying sequences (usage learning; Janik & Slater, 2000; Marshall *et al.,* 1999), as well as pant hoots that resemble the pant hoots of group members or social partners (production learning; Ruch *et al.,* 2018). The latter hypothesis is supported by the presence of community dialects (Crockford *et al.,* 2004; Mitani *et al.,* 1992) and by the stronger call similarity between social partners (Mitani & Brandt, 1994; Mitani & Gros-Louis, 1998), although genetic or ecological factors might also explain community differences (Mitani *et al.,* 1999; Desai et al., 2021). While some developmental changes result from the maturational process (Nishimura *et al.,* 2003), systematic study of the acquisition of vocal capacities, especially at the early ontogenetic stages, is a key missing element in the current debate on vocal learning in primates, particularly in great apes (e.g., Fischer *et al.,* 2015; Watson *et al.,* 2015a, 2015b). Our observations fit with the idea that primate vocal repertoires are largely fixed and present from birth (Fischer & Hammerschmidt, 2020), although they also indicate that fine acoustic structures undergo ontogenetic processes.

The limitations of our study include the short observation period and the small dataset, which reduce our ability to generalise from our study. Having complete access to the context of production typically facilitates call classification, although it can also be misleading with graded and flexible calls (Fischer & Price, 2017; Schamberg *et al.,* 2018). For instance, pant hoots are produced across most contexts and in response to conspecific calls (Goodall, 1986), while food grunts are related to both feeding and agonistic events (Ischer *et al.,* 2020; Marler & Tenaza, 1977). The categorisation of calls based on acoustic features is less subject to biases when the context of production is particularly unclear or flexible, as in pant hoots and immature calls. When applying fuzzy clustering on small datasets, calls can appear more discrete since they are less likely to represent the entire repertoire. In addition, extracting a small number of features can lead to higher spread of values but does not necessarily indicate better separation (Wadewitz *et al.,* 2015). Because we did not observe the production of typical newborn utterances such as cries or screams (Kojima, 2008), it is possible

that their absence affected our quantitative analyses. Specifically, the best solution was the most distinct but was also more influenced by a smaller subset of features. However, in a recent study the repertoire of infant chimpanzees was best described by a two-cluster model, with evidence of a potential third cluster (Taylor *et al.,* 2021), which provides further validity to the model describing the newborn vocalisations. We do not claim that there are only three call types in the repertoire of newborn chimpanzees. To determine its true size, it will be necessary to investigate how receivers react to each vocalisation (Seyfarth & Cheney, 2017), including to graded variants (e.g., Fischer, 1998; Fischer *et al.,* 2001).

The observation of a birth in a wild chimpanzee community provided a rare opportunity to investigate the vocal behaviour of a newborn chimpanzee, the starting point of a long developmental trajectory. The newborn demonstrated the capacity to combine different call types into larger vocal sequences. Some of these combinatorial structures were composed of unique calls that shared several characteristics with pant hoots and were identified by expert human listeners as such. Consequently, our study suggests that acoustically complex structures, akin to adult pant hoots, are part of the chimpanzee vocal repertoire from birth, and that these sequences are subject to ontogenetic processes that shape their acoustic structure. While extensive work has been conducted on adult male calls, and combinatorial capacities have recently gained attention, further work is necessary to elucidate the production, usage, and comprehension of complex vocal sequences in primates from an ontogenetic perspective. Although we remain careful in interpreting observations from a single individual, we believe it provides a valuable contribution to the study of chimpanzee vocal development that will hopefully encourage further research on the ontogeny of great ape vocalisations.

References

Association for the Study of Animal Behaviour (ASAB) (2018). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour, 135*, 1–9. https://doi.org/10.1016/j.anbehav.2017.10.001.

Bard, K. A. (1998). Social-experiential contributions to imitation and emotion in chimpanzees. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 208–227). Cambridge University Press.

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of Animal Communication* (2nd ed.p. 697). Sinauer.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R Package for Cluster Validation. *Journal of Statistical Software, 25*(4), 1–22. https://www.jstatsoft.org/v25/i04/. Accessed 22 Aug 2021

Clark, A. P., & Wrangham, R. W. (1993). Acoustic analysis of wild chimpanzee pant hoots: do Kibale Forest chimpanzees have an acoustically distinct food arrival pant hoot? *American Journal of Primatology, 31*(2), 99–109. https://doi.org/10.1002/ajp.1350310203.

Clark, A. P., & Wrangham, R. W. (1994). Chimpanzee arrival pant-hoots: Do they signify food or status? *International Journal of Primatology, 15*(2), 185–205. https://doi.org/10.1007/BF02735273.

Colombelli-Négrel, D., Hauber, M. E., Evans, C., Katsis, A. C., Brouwer, L., Adreani, N. M., & Kleindorfer, S. (2021). Prenatal auditory learning in avian vocal learners and non-learners. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 376*(1836), 20200247. https://doi.org/10.1098/rstb.2020.0247.

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. (2005). Call combinations in wild chimpanzees. *Behaviour, 142*(4), 397–421. https://doi.org/10.1163/1568539054012047.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology, 110*(3), 221–243. https://doi.org/10.1111/j.1439-0310.2004.00968.x.

Deecke, V. B., & Janik, V. M. (2006). Automated categorization of bioacoustic signals: Avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America, 119*(1), 645–653. https://doi.org/10.1121/1.2139067.

Desai, N. P., Fedurek, P., Slocombe, K. E., & Wilson, M. L. (2021). Chimpanzee pant-hoots encode information about individual but not group differences. *bioRxiv*. https://doi.org/10.1101/2021.03.09.434515.

Dezecache, G., Crockford, C., & Zuberbühler, K. (2019). The development of communication in alarm contexts in wild chimpanzees. *Behavioral Ecology and Sociobiology, 73*(8), 104. https://doi.org/10.1007/s00265-019-2716-6.

Dezecache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2021). A machine learning approach to infant distress calls and maternal behaviour of wild chimpanzees. *Animal Cognition, 24*(3), 443–455. https://doi.org/10.1007/s10071-020-01437-5.

Dunn, R. (2012). Night shift: labor and birth: it's all in the timing. *Natural History, 120*(3), 10–13.

Elowson, A. M., Snowdon, C. T., & Sweet, C. J. (1992). Ontogeny of trill and J-call vocalizations in the pygmy marmoset. *Cebuella pygmaea. Animal Behaviour, 43*(5), 703–715. https://doi.org/10.1016/S0003-3472(05)80195-2.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998a). `Babbling' and social context in infant monkeys: parallels to human infants. *Trends in Cognitive Sciences, 2*(1), 31–37. https://doi.org/10.1016/S1364-6613(97)01115-7.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998b). Infant "babbling" in a nonhuman primate: complex vocal sequences with repeated call types. *Behaviour, 135*(5), 643–664. https://doi.org/10.1163/156853998792897905.

Engesser, S., & Townsend, S. W. (2019). Combinatoriality in the vocal systems of nonhuman animals. *WIREs Cognitive Science, 10*(4), e1493. https://doi.org/10.1002/wcs.1493.

Fedurek, P., & Slocombe, K. E. (2011). Primate vocal communication: a useful tool for understanding human speech and language evolution? *Human Biology, 83*(2), 153–173. https://doi.org/10.3378/027.083.0202.

Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013a). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour, 86*(1), 189–196. https://doi.org/10.1016/j.anbehav.2013.05.010.

Fedurek, P., Schel, A. M., & Slocombe, K. E. (2013b). The acoustic structure of chimpanzee pant-hooting chorusing facilitates chorusing. *Behavioral Ecology and Sociobiology, 67*(11), 1781–1789. https://doi.org/10.1007/s00265-013-1585-7.

Fedurek, P., Donnellan, E., & Slocombe, K. E. (2014). Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behavioral Ecology and Sociobiology, 68*(8), 1345–1355. https://doi.org/10.1007/s00265-014-1745-4.

Fedurek, P., Slocombe, K. E., & Zuberbühler, K. (2015). Chimpanzees communicate to two different audiences during aggressive interactions. *Animal Behaviour, 110*, 21–28. https://doi.org/10.1016/j.anbehav.2015.09.010.

Fedurek, P., Slocombe, K. E., Enigk, D. K., Thompson, M. E., Wrangham, R. W., & Muller, M. N. (2016a). The relationship between testosterone and long-distance calling in wild male chimpanzees. *Behavioral Ecology and Sociobiology, 70*(5), 659–672. https://doi.org/10.1007/s00265-016-2087-1.

Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016b). Sequential information in a great ape utterance. *Scientific Reports, 6*(1), 38226. https://doi.org/10.1038/srep38226.

Fedurek, P., Zuberbühler, K., & Semple, S. (2017). Trade-offs in the production of animal vocal sequences: insights from the structure of wild chimpanzee pant hoots. *Frontiers in Zoology, 14*(1), 50. https://doi.org/10.1186/s12983-017-0235-8.

Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour, 55*(4), 799–807. https://doi.org/10.1006/anbe.1997.0663.

Fischer, J., & Hage, S. R. (2019). Primate vocalization as a model for human speech: scopes and limits. In P. Hagoort (Ed.), *Human language: from genes and brains to behavior* (pp. 639–656). MIT Press.

Fischer, J., & Hammerschmidt, K. (2020). Towards a new taxonomy of primate vocal production learning. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 375*(1789), 20190045. https://doi.org/10.1098/rstb.2019.0045.

Fischer, J., & Price, T. (2017). Meaning, intention, and inference in primate vocal communication. *Neuroscience & Biobehavioral Reviews, 82*, 22–31. https://doi.org/10.1016/j.neubiorev.2016.10.014.

Fischer, J., Metz, M., Cheney, D. L., & Seyfarth, R. M. (2001). Baboon responses to graded bark variants. *Animal Behaviour, 61*(5), 925–931. https://doi.org/10.1006/anbe.2000.1687.

Fischer, J., Wheeler, B. C., & Higham, J. P. (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Current Biology, 25*(21), R1028–R1029. https://doi.org/10.1016/j.cub.2015.09.010.

Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour, 134*, 229–237. https://doi.org/10.1016/j.anbehav.2016.06.012.

Fitch, W. T., & Zuberbühler, K. (2013). Primate precursors to human language: Beyond discontinuity. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of Emotional Communication* (pp. 26–48). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199583560.003.0002.

Fujisawa, M., Hockings, K. J., Soumah, A. G., & Matsuzawa, T. (2016). Placentophagy in wild chimpanzees (*Pan troglodytes verus*) at Bossou, Guinea. *Primates, 57*(2), 175–180. https://doi.org/10.1007/s10329-016-0510-x.

Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences, 21*, 62–67. https://doi.org/10.1016/j.cobeha.2018.02.004.

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), 1–15. https://doi.org/10.1038/s42003-022-03350-8.

Goodall, J. (1986). *The chimpanzees of Gombe: patterns of behavior*. Harvard University Press.

Goodall, J., & Athumani, J. (1980). An observed birth in a free-living chimpanzee (*Pan troglodytes schweinfurthii*) in Gombe National Park, Tanzania. *Primates, 21*(4), 545–549. https://doi.org/10.1007/BF02373843.

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In D. Kimbrough Oller & U. Griebel (Eds.), *Evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication* (pp. 93–119). MIT Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Hayes, K. J., & Hayes, C. (1951). The intellectual development of a home-raised chimpanzee. *Proceedings of the American Philosophical Society, 95*(2), 105–109.

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.

Hiraiwa-Hasegawa, M. (1986). *Development and mother–infant relationship of wild chimpanzees in Mahale National Park*. Doctoral dissertation. Tokyo: University of Tokyo

Holden, E. (2017). *The structure and function of female chimpanzee pant-hoot vocalisations*. University of St Andrews.

Ischer, G., Zuberbühler, K., & Fedurek, P. (2020). The relationship between food calling and agonistic behaviour in wild chimpanzees. *Behavioural Processes, 178*, 104182. https://doi.org/10.1016/j.beproc.2020.104182.

Janik, V. M. (1999). Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. *Animal Behaviour, 57*(1), 133–143. https://doi.org/10.1006/anbe.1998.0923.

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.

Kellogg, W. N. (1968). Communication and language in the home-raised chimpanzee. *Science, 162*(3852), 423–427. https://doi.org/10.1126/science.162.3852.423.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., et al (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus: Acoustic sequences in animals. *Biological Reviews, 91*(1), 13–52. https://doi.org/10.1111/brv.12160.

Kiwede, Z. T. (2000). A Live birth by a primiparous female chimpanzee at the Budongo Forest. *Pan Africa News, 7*(2), 23–25. https://doi.org/10.5134/143554.

Kojima, S. (2008). Early vocal development in a chimpanzee infant. In T. Matsuzawa (Ed.), *Primate Origins of Human Cognition and Behavior* (pp. 190–196). Springer. https://doi.org/10.1007/978-4-431-09423-4_9.

Lameira, A. R. (2017). Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews, 83*, 429–439. https://doi.org/10.1016/j.neubiorev.2017.09.021.

Lameira, A. R., Santamaría-Bonfil, G., Galeone, D., Gamba, M., Hardus, M. E., Knott, C. D., et al (2022). Sociality predicts orangutan vocal phenotype. *Nature Ecology & Evolution, 6*, 644–652. https://doi.org/10.1038/s41559-022-01689-z.

Langergraber, K. E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., Fawcett, K., et al (2012). 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*(39), 15716–15721. https://doi.org/10.1073/pnas.1211740109.

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., & Townsend, S. W. (2020). Call combinations in great apes and the evolution of syntax. *Animal Behavior and Cognition, 7*(2), 131–139. https://doi.org/10.26451/abc.07.02.07.2020.

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.

Locke, J. L. (2006). Parental selection of vocal behavior. *Human Nature, 17*(2), 155–168. https://doi.org/10.1007/s12110-006-1015-x.

Lüdecke, D., Makowski, D., Waggoner, P., & Patil, P. (2021). performance: an R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K.(2021). *cluster: Cluster Analysis Basics and Extensions*. R package version 2.1.2. https://CRAN.R-project.org/package=cluster

Marler, P., & Hobbett, L. (1975). Individuality in a long-range vocalization of wild chimpanzees. *Zeitschrift für Tierpsychologie, 38*(1), 97–109. https://doi.org/10.1111/j.1439-0310.1975.tb01994.x.

Marler, P., & Tenaza, R. (1977). Signalling behaviour of apes with special reference to vocalisations. In P. Marler, R. Tenaza, & T. A. Sebeok (Eds.), *How animals communicate* (pp. 965–1033). Indiana University Press.

Marshall, A. J., Wrangham, R. W., & Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour, 58*(4), 825–830. https://doi.org/10.1006/anbe.1999.1219.

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.

Mitani, J. C., & Brandt, K. L. (1994). Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology, 96*(3), 233–252. https://doi.org/10.1111/j.1439-0310.1994.tb01012.x.

Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour, 135*(8), 1041–1064. https://doi.org/10.1163/156853998792913483.

Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P., & Byrne, R. (1992). Dialects in wild chimpanzees? *American Journal of Primatology, 27*(4), 233–243. https://doi.org/10.1002/ajp.1350270402.

Mitani, J. C., Gros-Louis, J., & Macedonia, J. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology, 17*(4), 569–583. https://doi.org/10.1007/BF02735192.

Mitani, J. C., Hunley, K. L., & Murdoch, M. E. (1999). Geographic variation in the calls of wild chimpanzees: a reassessment. *American Journal of Primatology, 47*(2), 133–151. https://doi.org/10.1002/(SICI)1098-2345(1999)47:2%3C133::AID-AJP4%3E3.0.CO;2-I

Moran, S., McCloy, D., & Wright, R. (2012). Revisiting population size vs. phoneme inventory size. *Language, 88*(4), 877–893.

Mundry, R., & Sommer, C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour, 74*(4), 965–976. https://doi.org/10.1016/j.anbehav.2006.12.028.

Nagy, E. (2011). The newborn infant: a missing stage in developmental psychology. *Infant and Child Development, 20*(1), 3–19. https://doi.org/10.1002/icd.683.

Narayanan, D. Z., Takahashi, D. Y., Kelly, L. M., Hlavaty, S. I., Huang, J., & Ghazanfar, A. A. (2022). Prenatal development of neonatal vocalizations. *eLife, 11, e78485*. https://doi.org/10.7554/eLife.78485

Newton-Fisher, N. E. (1999). Association by male chimpanzees: a social tactic? *Behaviour, 136*(6), 705–730. https://doi.org/10.1163/156853999501531.

Nishie, H., & Nakamura, M. (2018). A newborn infant chimpanzee snatched and cannibalized immediately after birth: implications for "maternity leave" in wild chimpanzee. *American Journal of Physical Anthropology, 165*(1), 194–199. https://doi.org/10.1002/ajpa.23327.

Nishimura, T., Mikami, A., Suzuki, J., & Matsuzawa, T. (2003). Descent of the larynx in chimpanzee infants. *Proceedings of the National Academy of Sciences, 100*(12), 6930–6933. https://doi.org/10.1073/pnas.1231107100.

Notman, H., & Rendall, D. (2005). Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Animal Behaviour, 70*(1), 177–190. https://doi.org/10.1016/j.anbehav.2004.08.024.

Oller, D. K. (2000). *The emergence of the speech capacity*. Psychology Press. https://doi.org/10.4324/9781410602565.

Oller, D. K., & Griebel, U. (2008). Contextual flexibility in infant vocal development and the earliest steps in the evolution of language. In *Evolution of communicative flexibility: Complexity, creativity and adaptability in human and animal communication* (pp. 141–168). MIT Press.

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Topics in Cognitive Science, 8*(2), 382–392. https://doi.org/10.1111/tops.12198.

Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., & Call, J. (2019). Language origins viewed in spontaneous and interactive vocal rates of human and bonobo infants. *Frontiers in Psychology, 10*, 00729 https://www.frontiersin.org/article/10.3389/fpsyg.2019.00729. Accessed 23 May 2022.

Oller, D. K., Ramsay, G., Bene, E., Long, H. L., & Griebel, U. (2021). Protophones, the precursors to speech, dominate the human infant vocal landscape. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 376*(1836), 20200255. https://doi.org/10.1098/rstb.2020.0255.

Otali, E., & Gilchrist, J. S. (2006). Why chimpanzee (*Pan troglodytes schweinfurthii*) mothers are less gregarious than nonmothers and males: the infant safety hypothesis. *Behavioral Ecology and Sociobiology, 59*(4), 561–570. https://doi.org/10.1007/s00265-005-0081-0.

Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology, 73*(6), 530–544. https://doi.org/10.1002/ajp.20913.

Palombit, R. A. (2012). Infanticide: Male Strategies and Female Counterstrategies. In J. C. Mitani, C. Josep, R. A. Palombit, & J. B. Silk (Eds.), *The Evolution of Primate Societies* (pp. 432–468). University of Chicago Press. https://doi.org/10.7208/9780226531731-022.

Plooij, F. X. (1984). *The behavioral development of free-living chimpanzee babies and infants*. Ablex.

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

R Core Team (2021). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing https://www.R-project.org/. Accessed 22 Sept 2020

Ramsay, M. S., & Teichroeb, J. A. (2019). Anecdotes in primatology: temporal trends, anthropocentrism, and hierarchies of knowledge. *American Anthropologist, 121*(3), 680–693. https://doi.org/10.1111/aman.13295.

Reynolds, V. (2005). *The chimpanzees of the Budongo forest: Ecology, behaviour and conservation*. Oxford University Press.

Riede, T., Arcadi, A. C., & Owren, M. J. (2007). Nonlinear acoustics in the pant hoots of common chimpanzees (*Pan troglodytes*): Vocalizing at the edge. *The Journal of the Acoustical Society of America, 121*(3), 1758–1767. https://doi.org/10.1121/1.2427115.

Ruch, H., Zürcher, Y., & Burkart, J. M. (2018). The function and mechanism of vocal accommodation in humans and other primates: vocal accommodation in human and other primates. *Biological Reviews, 93*(2), 996–1013. https://doi.org/10.1111/brv.12382.

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.

Schamberg, I., Wittig, R. M., & Crockford, C. (2018). Call type signals caller goal: a new take on ultimate and proximate influences in vocal production: call type signals caller goal. *Biological Reviews, 93*(4), 2071–2082. https://doi.org/10.1111/brv.12437.

Schick, J., Fryns, C., Wegdell, F., Laporte, M., Zuberbühler, K., Schaik, C. P. van, et al. (2022). The function and evolution of child-directed communication. *PLoS Biology*, 20(5), e3001630. https://doi.org/10.1371/journal.pbio.3001630

Seyfarth, R. M., & Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In C. T. Snowdon & M. Hausberger (Eds.), *Social Influences on Vocal Development* ((1st ed., pp. 249–273). Cambridge University Press. https://doi.org/10.1017/CBO9780511758843.013.

Seyfarth, R. M., & Cheney, D. L. (2017). The origin of meaning in animal signals. *Animal Behaviour, 124*, 339–346. https://doi.org/10.1016/j.anbehav.2016.05.020.

Slocombe, K. E., & Zuberbühler, K. (2010). Vocal communication in chimpanzees. In E. V. Lonsdorf, R. R. Stephen, & T. Matsuzawa (Eds.), *The Mind of the Chimpanzee* (pp. 192–207). University of Chicago Press. https://doi.org/10.7208/9780226492810.

Snowdon, C. T. (2009). Plasticity of communication in nonhuman primates. In M. Naguib, K. Zuberbuhler, N. S. Clayton, & V. M. Janik (Eds.), *Advances in the study of behavior* (Vol. 40, pp. 239–276). Elsevier. https://doi.org/10.1016/S0065-3454(09)40007-X.

Snowdon, C. T., & Elowson, A. M. (2001). "Babbling" in pigmy marmosets: development after infancy. *Behaviour, 138*(10), 1235–1248. https://doi.org/10.1163/15685390152822193.

Snowdon, C. T., Elowson, A. M., & Roush, R. S. (1997). Social influences on vocal development in New World primates. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* ((1st ed., pp. 234–248). Cambridge University Press. https://doi.org/10.1017/CBO9780511758843.012.

Soldati, A., Fedurek, P., Dezecache, G., Call, J., & Zuberbühler, K. (2022). Audience sensitivity in chimpanzee display pant hoots. *Animal Behaviour, 190*, 23–40. https://doi.org/10.1016/j.anbehav.2022.05.010.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science, 349*(6249), 734–738. https://doi.org/10.1126/science.aab1058.

Taylor, D. (2020). *The Infrastructure of Chimpanzee Vocal Ontogeny* (Doctoral dissertation, University of Portsmouth)

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). https://doi.org/10.1002/ajp.23249

ter Haar, S. M., Fernandez, A. A., Gratier, M., Knörnschild, M., Levelt, C., Moore, R. K., et al (2021). Cross-species parallels in babbling: animals and algorithms. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 376*(1836), 20200239. https://doi.org/10.1098/rstb.2020.0239

Tomasello, M., & Zuberbühler, K. (2002). Primate vocal and gestural communication. In *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 293–299). MIT Press. https://doi.org/10.7551/mitpress/1885.001.0001

Townsend, S. W., Engesser, S., Stoll, S., Zuberbühler, K., & Bickel, B. (2018). Compositionality in animals and humans. *PLoS Biology, 16*(8), e2006425. https://doi.org/10.1371/journal.pbio.2006425

Trevathan, W. (2015). Primate pelvic anatomy and implications for birth. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 370*(1663), 20140065. https://doi.org/10.1098/rstb.2014.0065

Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 375*(1789), 20180406. https://doi.org/10.1098/rstb.2018.0406

Varga, Z., Garami, L., Ragó, A., Honbolygó, F., & Csépe, V. (2019). Does intra-uterine language experience modulate word stress processing? An ERP study. *Research in Developmental Disabilities, 90*, 59–71. https://doi.org/10.1016/j.ridd.2019.04.011

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.). Springer ISBN 0-387-95457-0. https://www.stats.ox.ac.uk/pub/MASS4/. Accessed 10 June 2022

Vernes, S. C., Kriengwatana, B. P., Beeck, V. C., Fischer, J., Tyack, P. L., ten Cate, C., & Janik, V. M. (2021). The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 376*(1836), 20200236. https://doi.org/10.1098/rstb.2020.0236

Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015). Characterizing vocal repertoires—hard vs. soft classification approaches. *PLoS One, 10*(4), e0125785. https://doi.org/10.1371/journal.pone.0125785

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., et al (2015a). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology, 25*(4), 495–499. https://doi.org/10.1016/j.cub.2014.12.032

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., et al (2015b). Reply to Fischer et al. *Current Biology, 25*(21), R1030–R1031. https://doi.org/10.1016/j.cub.2015.09.024

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature, 582*(7812), 337–340. https://doi.org/10.1038/d41586-020-01751-5

Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences, 178*(13), 2751–2779. https://doi.org/10.1016/j.ins.2008.02.012

Zamma, K., & Shabani, R. (2012). A wild chimpanzee birth at Mahale. *Pan Africa News, 19*(1), 3–5.

Zuberbühler, K., & Lemasson, A. (2014). Primate communication: meaning from strings of calls. In F. Lowenthal & L. Lefebvre (Eds.), *Language and Recursion* (pp. 115–125). Springer. https://doi.org/10.1007/978-1-4614-9414-0_9

Authors and Affiliations

Adrian Soldati[1,2,3] · **Geresomu Muhumuza[3]** · **Guillaume Dezecache[3,4]** · **Pawel Fedurek[3,5]** · **Derry Taylor[2,3]** · **Josep Call[1]** · **Klaus Zuberbühler[1,2,3]**

[1] School of Psychology and Neuroscience, University of St Andrews, Westburn Lane, St Andrews, Fife, UK

[2] Department of Comparative Cognition, University of Neuchâtel, Neuchâtel, Switzerland

[3] Budongo Conservation Field Station, Masindi, Uganda

[4] Université Clermont Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France

[5] Division of Psychology, Faculty of Natural Sciences, University of Stirling, Stirling, UK