

Research



Cite this article: Pereira AS, Kavanagh E, Hobaiteer C, Slocombe KE, Lameira AR. 2020 Chimpanzee lip-smacks confirm primate continuity for speech-rhythm evolution. *Biol. Lett.* **16**: 20200232.
<http://dx.doi.org/10.1098/rsbl.2020.0232>

Received: 7 April 2020

Accepted: 30 April 2020

Subject Areas:

behaviour, evolution

Keywords:

speech-like rhythm, speech evolution, lip-smacks, great apes, chimpanzees

Author for correspondence:

Adriano R. Lameira

e-mail: adriano.lameira@warwick.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4980008>.

Chimpanzee lip-smacks confirm primate continuity for speech-rhythm evolution

André S. Pereira^{1,2}, Eithne Kavanagh³, Catherine Hobaiteer¹, Katie E. Slocombe³ and Adriano R. Lameira^{1,4}

¹School of Psychology and Neuroscience, University of St Andrews, St Mary's Quad, South Street, St Andrews, KY16 9JP, UK

²School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen, AB24 2TZ, UK

³Department of Psychology, University of York, Heslington, York, YO10 5DD, UK

⁴Department of Psychology, University of Warwick, University Road, Coventry, CV4 7AL, UK

KES, 0000-0002-7310-1887; ARL, 0000-0003-3071-4824

Speech is a human hallmark, but its evolutionary origins continue to defy scientific explanation. Recently, the open–close mouth rhythm of 2–7 Hz (cycles/second) characteristic of all spoken languages has been identified in the orofacial signals of several nonhuman primate genera, including orangutans, but evidence from any of the African apes remained missing. Evolutionary continuity for the emergence of speech is, thus, still inconclusive. To address this empirical gap, we investigated the rhythm of chimpanzee lip-smacks across four populations (two captive and two wild). We found that lip-smacks exhibit a speech-like rhythm at approximately 4 Hz, closing a gap in the evidence for the evolution of speech-rhythm within the primate order. We observed sizeable rhythmic variation within and between chimpanzee populations, with differences of over 2 Hz at each level. This variation did not result, however, in systematic group differences within our sample. To further explore the phylogenetic and evolutionary perspective on this variability, inter-individual and inter-population analyses will be necessary across primate species producing mouth signals at speech-like rhythm. Our findings support the hypothesis that speech recruited ancient primate rhythmic signals and suggest that multi-site studies may still reveal new windows of understanding about these signals' use and production along the evolutionary timeline of speech.

1. Introduction

Throughout history, few traces for the evolution of speech have been found among nonhuman primates (hereafter primates), obscuring the precursors and processes through which our species came to develop a unique and powerful signal system. The past few decades have, however, seen promising new advances [1–4]. A research frontier that has gradually yielded some of the most compelling evidence is the study of the evolutionary origin of speech-rhythm, i.e. the fast open–close mouth cycles characteristic to each and every spoken language in the world [5]. This rhythm is inherent to speech and universal across spoken languages because it expresses the production of syllables, where the opening and closing of the mouth roughly correspond to vowel and consonant production, respectively [6,7]. This rhythm typically exhibits a rate of 2–7 Hz, i.e. 2 to 7 open–close mouth cycles per second [5], and is a visual and acoustic signal of speech that appears to be critical to its intelligibility [8–10].

Speech-like rhythm has been uncovered in a growing number of primate signals: lip-smacks of various macaque species [11,12], stump-tailed macaques' panting calls [12], gelada's wobbles [13], gibbon song [14] and orangutan clicks and faux speech [15]. Further studies have shown that, in macaques, lip-smacks

Table 1. Lip-smack data used for analysis.

population	no. individuals	no. bouts (no. open–close mouth cycles)/individual
Edinburgh	3 (1 female,	female: 8 (49)
	2 males)	males: 16 (104), 7 (53.5)
Leipzig	3 (1 female,	female: 6 (24)
	2 males)	males: 1 (3), 1 (9)
Kanyawara	5 (1 female,	female: 1 (5)
	4 males)	males: 2 (6), 2 (8), 1 (5), 1 (3)
Waibira	3 (1 female,	female: 1 (2)
	2 males)	males: 2 (9), 5 (25)

develop along a similar trajectory to human speech [16] and activate an area homologous to Broca's [17], with individuals being perceptually attuned to lip-smacks' natural frequency [18]. Together, these convergent lines of evidence across fields and taxa indicate, on the basis of homology, that speech-rhythm likely derived from ancient fast-paced mouth signals from deep within the primate lineage [19–21]. The overall validity of this hypothesis for the evolution of speech-rhythm and the assumption of evolutionary continuity across fast-paced mouth movements in primates rest, however, on a last phylogenetic steppingstone for which there is currently no data: the African great apes, the closest extant hominid lineage to humans.

Here, to directly explore this gap in knowledge, we characterize the rhythm of chimpanzee (*Pan troglodytes* spp.) lip-smacks—affiliative signals typically produced by groomers during social grooming [22,23].

2. Methods

(a) Study subjects and data collection

We identified lip-smack bouts present in video recordings collected at Edinburgh Zoo (UK; *Pan troglodytes verus* and one hybrid) (table 1) during August and September 2013 with a Panasonic HDC SDX1; at Leipzig Zoo (Germany; *P. t. verus*) (table 1), during June and July 2017 with a Panasonic HDC-SD90 camcorder with a Sennheiser MKE 400 microphone attached; in the wild in the Kanyawara community (Kibale National Park, Uganda; *P. t. schweinfurthii*) (table 1), during December 2014 and August and September 2016 with a Panasonic HDC-SD90 camcorder with a Sennheiser MKE 400 microphone attached; and from the Waibira community (Budongo Forest Research, Uganda; *P. t. schweinfurthii*) (table 1) during December 2011, March 2012, December 2014 and August 2017 with a Panasonic SD90. All videos were 25 frames per second. Videos were selected for analysis when the face of the emitter was clearly visible during lip-smack production, and this was the sole criterion for including a bout in the analysis. There was no proactive selection of particular individuals. All videos had been collected during opportunistic observation of the subjects' behaviour.

Permission to collect video data had been previously obtained from the authors' institutions (either for other projects or routine data collection) and all the relevant bodies responsible for managing research at each population. All procedures followed the Association for the Study of Animal Behaviour/Animal Behavior Society Guidelines for the Use of Animals in Research (Animal Behaviour, 2018, 135, I-X), all institutional

guidelines, the legal requirements of the countries in which the work was carried out, and was granted ethical approval by the Biology Animal Welfare Ethical Review Board (AWERB), University of York, York, UK, or the Animal Welfare and Ethics Committee, University of St Andrews, UK.

(b) Data analyses

We used Filmora9 (Wondershare Technology Co., Shenzhen) to extract all identified lip-smack bouts from the grooming bout videos. We used the *VideoReader* function to load all lip-smack videos to MATLAB R2018a (MathWorks, Natick, MA) and extracted all frames of each bout.

To investigate whether chimpanzee lip-smacks exhibit a speech-like rhythm, we calculated the dominant frequency of lip-smacking behaviour by extracting the power spectral density, i.e. the quantity of power for each frequency component of a signal, of all lip-smack bouts and then calculating its peak, which reflects the most representative frequency of mouth aperture, and that we considered to be the approximate rate of mouth oscillation across lip-smack bouts [15,16]. To do this, we used the *intool* function to load all frames individually to MATLAB and used the *Measure Distance* tool to measure the distance between a fixed point in the top lip and a fixed point in the bottom lip of the emitter [15,16,18] (electronic supplementary material S1). For open-mouth cycles in which lip movement did not match jaw displacement, we measured the distance between a point in the lower lip and the most fixed and easily identifiable point of the video (e.g. the nasion or the glabella), which allowed us to capture the movements of opening and closing of the jaw [16,18]. For the frames in which the marking points were not clearly visible, we estimated mouth displacement to be the mean of the adjacent frames [15]. This estimation was possible because there was never more than one consecutive frame during which we couldn't identify the marking points.

For each bout, we used the mouth displacement measurements to construct a time-series of mouth displacement [15,16,18] (electronic supplementary material S1). To allow for comparability between bouts, we normalized the amplitude of every time-series so that the mouth displacement measures of each time-series varied between 0 and 100. We did so by subtracting the minimum mouth displacement measurement of each time-series from all its mouth displacement measurements and followed by setting all measurements as a percentage of the maximum mouth displacement measurement of the series [16]. For each time-series, we subtracted the mean of all normalized mouth displacement measurements from each normalized measurement to eliminate the D-C offset (i.e. mean amplitude displacement from zero) and, thus, avoiding getting 0 as the dominant frequency. Subsequently, we used MATLAB's *fft* function to perform a fast Fourier transform (FFT) of each time-series [16] (electronic supplementary material S2). We set the 'NFFT', a parameter that defines the frequency scale of the *fft*, to 1024 for every time-series, a value large enough to allow good resolution of the signal in all series without compromising computational time. We squared the magnitude of each time-series' FFT to obtain the series' power spectrum density (electronic supplementary material S2).

Finally, we used the R package *ggplot2* [24] to plot the smoothed out mean $\pm 95\%$ confidence interval of the standardized power spectrum density of all time-series and used custom R scripts to find the peak of the curve, i.e. the dominant frequency of chimpanzee lip-smacking behaviour. We standardized all power spectrum density curves by standardizing the spectral power variation (Y-axis) from 0 to 100 following the procedure previously described for the standardization of the time-series. This standardization allowed us to account for the relative spectral power at all frequencies of all bouts while avoiding having individual curves contributing differently to the mean

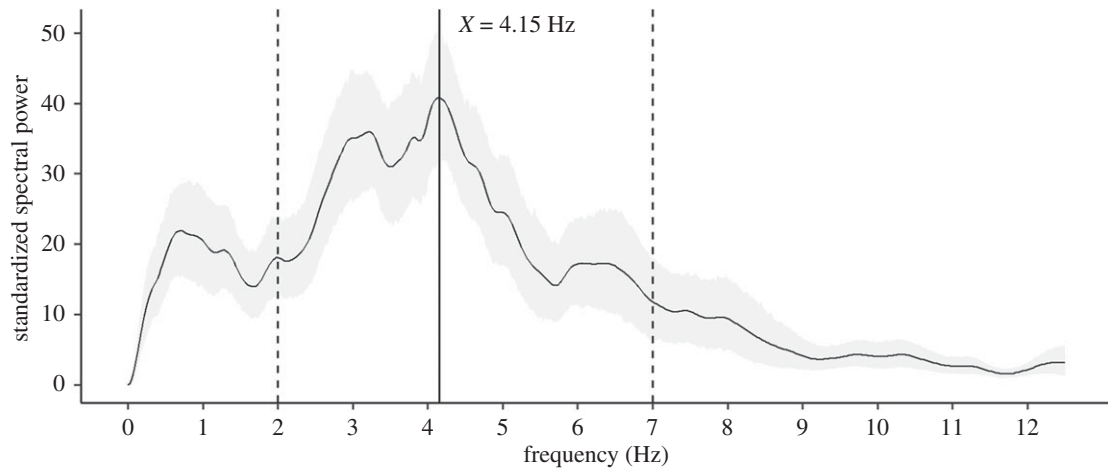


Figure 1. Mean standardized power spectral density plot of chimpanzee lip-smacking. The mean $\pm 95\%$ confidence interval standardized power spectral density plot of all 54 analysed chimpanzee lip-smack bouts peaks at 4.15 Hz, which represents the dominant frequency of chimpanzee lip-smack production rate. Black dashed vertical lines indicate limits of speech-like rhythm frequencies.

curve. To help visualize the data, we used the same procedure to plot the mean $\pm 95\%$ confidence interval of the power spectrum density of all time-series of each individual in each population, as well as of each pair of populations. All time-series and each time-series' plot and power spectrum density plot can be found in electronic supplementary material S1. All code and steps to replicate the analysis described here are available in electronic supplementary material S2.

To statistically compare frequency peaks between captivity and the wild, we used the *glmer* function from the R package *lme4* [25] to build a generalized linear mixed model, which we setup with a gamma error structure and inverse link function; the peak of each individual bout was input as the dependent variable; population (Edinburgh, Leipzig, Kanyawara or Waibira) was input as a fixed factor, and the identity of each individual was input as a random factor to control for repeated measures. We confirmed that the distribution of the residuals was normally distributed and that there was no issue of overdispersion. The code for this analysis can be found in electronic supplementary material S2. Because the highest peak of some individual lip-smack bouts reflected the distribution of interbout intervals (typically less than 1 Hz) instead of the real peak, which is a regular occurrence in studies of speech rhythmicity (for example, see [15]), we assessed all bouts individually and, for such deviant cases, only included the peaks of the dominant frequency plot (electronic supplementary material S1) that corresponded to the true mean of open-mouth cycles per second, as observed from each bout's time-series (electronic supplementary material S1).

3. Results

We found that chimpanzee lip-smacks exhibited a mean rhythm per bout of 4.15 Hz (figure 1). We identified rhythm variation in lip-smack rate production across individuals who exhibited the behaviour within and across populations (figure 2). For each of the populations, individual lip-smack rhythm spanned a frequency range of at least 1 Hz, with maximum differences above 2 Hz between some individuals in some of the populations (coloured vertical dashed lines, figure 2*a–d*). Per population, chimpanzees produced lip-smacks with a mean rhythm of 4.20 Hz at Edinburgh (*P. t. verus* or hybrid, captive), 4.08 Hz at Leipzig (*P. t. verus*, captive), 2.86 Hz at Kanyawara (*P. t. schweinfurthii*, wild) and 1.95 Hz at Waibira (*P. t. schweinfurthii*, wild)

(coloured vertical lines, figure 2*e–j*). The average (arithmetic mean) of the mean rhythm per population was 3.27 Hz. The mean rhythm between the two captive populations was nearly equal. Between the two wild populations there was an observed difference of approximately 1 Hz. Any dyad with a captive versus wild population exhibited a difference between greater than 1 and less than 2.5 Hz in lip-smack rhythm. To investigate the apparent differences in the rhythm of lip-smacks between captive versus wild populations, we ran a generalized linear mixed model with contrasts between the weighted means of the two captive populations and the two wild populations (electronic supplementary material S2). The mean average (standard deviation) rhythm peak in captivity was 4.69 Hz (± 1.32 Hz) and in the wild was 3.07 Hz (± 0.79 Hz) (corresponding arithmetic average, that is, sum of each population average divided by number of populations, was 4.37 Hz in captivity and 3.09 Hz in the wild); however, we found no difference between groups ($p = 0.0866$).

4. Discussion

We found that chimpanzees produce lip-smacks at an average speech-like rhythm of 4.15 Hz. These results close the gap between available data on primate fast-paced rhythmic mouth signals and human speech, offering clear support for the hypothesis that speech-rhythm has deep origins within the primate lineage [3,19,20] and was built upon existing signal systems (e.g. [26]).

Our multi-population analyses revealed a level of variation in chimpanzee lip-smack rhythmic production that to our knowledge has not been so far reported in any primate species with similar signals. Differences between individuals and populations reached more than 2 Hz at times. Considering that, in great apes, the fastest oscillatory vocal signals do not surpass mouth rhythms of 1 Hz [15], the observed variability span in lip-smack production may suggest that these are not hard-wired or stereotypical signals, and/or that socio-ecological factors differently affect lip-smack rhythm in chimpanzees at the level of individuals and/or populations. Despite having pooled for the first-time data across four populations for the analyses of primate fast-paced mouth signals, current sample sizes did not offer adequate

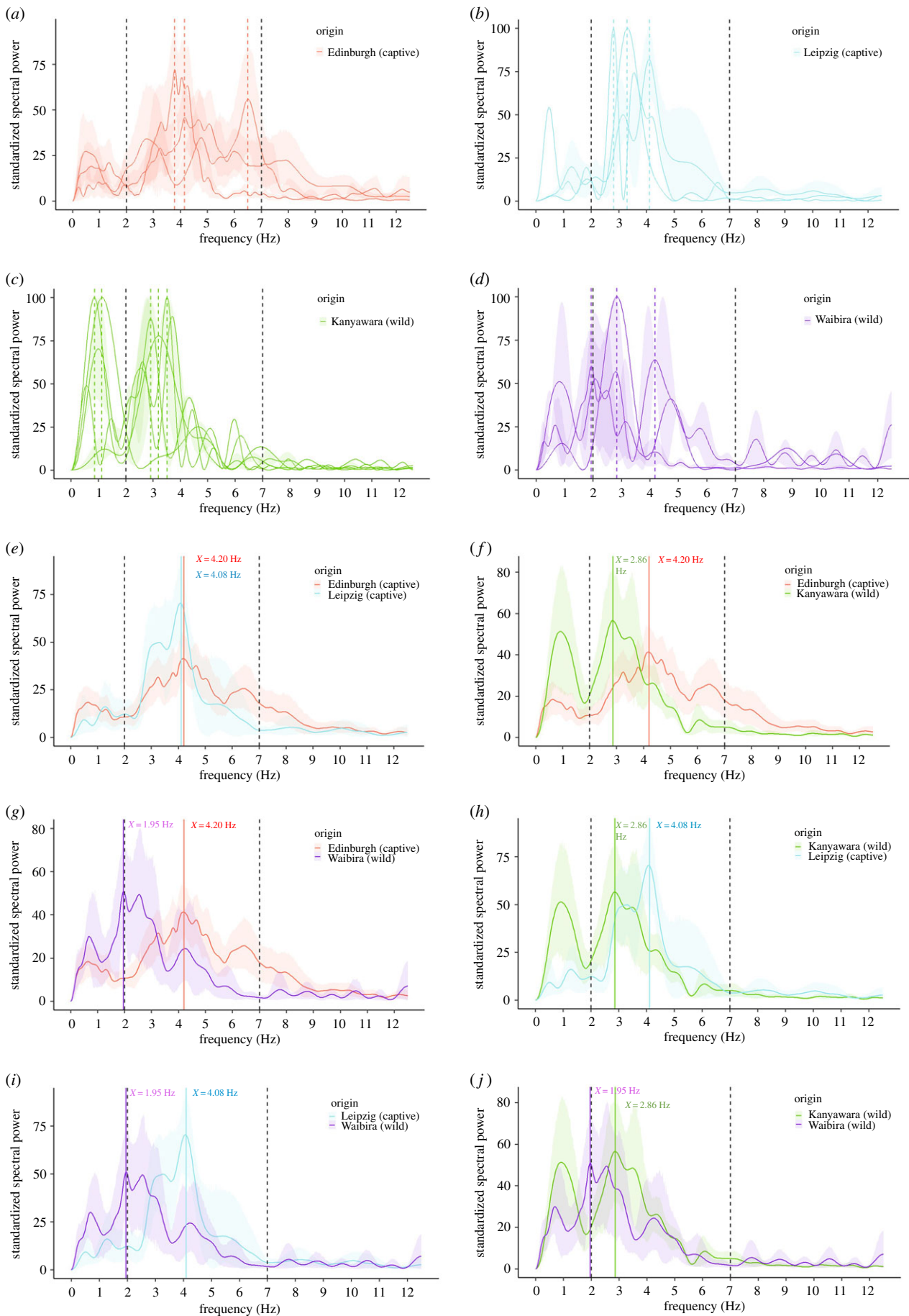


Figure 2. Mean standardized power spectral density plot of each individual's lip-smack production rate in each population (*a–d*) and mean standardized power spectral density plot of each pair of populations (*e–j*). Shaded areas represent the mean $\pm 95\%$ confidence interval standardized power spectral density plot per individual (*a–d*) and per population (*e–j*). Coloured dashed vertical lines indicate maximum frequency peak per individual (*a–d*) and per population (*e–j*). Black dashed vertical lines indicate limits of speech-like rhythm frequencies.

statistical power to identify significant differences with confidence or help identify possible correlates. Comparison between captive and wild populations was possible; despite rhythmic differences of greater than 1.5 Hz between the two types of populations, we found no systematic difference, likely as the result of striking within-population variability and substantial overlap in the range of rhythms present.

Alas, despite several primate species being known to exhibit mouth signals at speech-like rhythm, few of the respective studies have disclosed or analysed the levels of variation found between individuals. Although measures of variation in cycle durations (e.g. *SD*) are available (e.g. [12]), it is impossible to deduce whether this variation is attributable to intra-individual variation, context or inter-individual variation. Moreover, the lack of multi-site analyses in any of these species prevents a comparison with our results and an interpretation of evidence from a wider phylogenetic or evolutionary angle. Data on variation between individuals and sites would be particularly valuable for gaining new insight into the natural history of primate signals with speech-like rhythm. For example, signals exhibiting speech-like rhythm in macaques and gibbons are generally thought to be innate [27,28], but orangutan speech-rhythm has been identified in idiosyncratic, species-atypical, individual-specific calls presumed to be learned [15]. In our own analyses, there seemed to be variation in the frequency with which individual chimpanzees produced lip-smacks, with some never or only very rarely observed to produce lip-smacks despite similar observation hours to their group members (Hobaiter 2020, unpublished data). Together with the observed degree of variation in lip-smack rhythm across chimpanzee individuals and populations, available great ape data could hint at the intriguing possibility of a fixed-to-flexible transition in the ontogeny of the primate speech-like rhythmic phenotype at the base of the hominid lineage. However, this possibility remains tentative until new, more detailed data become available from both non-hominid and hominid primates. Future research across primate species employing a similar inter-individual and inter-population

approach and focusing on prevalence and rhythm variation is critical to discerning the evolutionary trajectory of fast-paced facial movements along the primate lineage, movements that ultimately culminated in the 2–7 Hz rhythm of speech in our species.

Ethics. Permission to collect video data had been previously obtained from the authors' institutions (either for other projects or routine data collection) and all the relevant bodies responsible for managing research at each population. All procedures followed the Association for the Study of Animal Behaviour/Animal Behavior Society Guidelines for the Use of Animals in Research (Animal Behaviour, 2018, 135, I-X), all institutional guidelines, the legal requirements of the countries in which the work was carried out, and was granted ethical approval by either the Biology Animal Welfare Ethical Review Board (AWERB), University of York, or the Animal Welfare and Ethics Committee, University of St Andrews.

Data accessibility. All data needed to evaluate the conclusions in the paper are present in the paper and in the electronic supplementary material.

Authors' contributions. A.S.P. and E.K. conducted analyses and wrote the paper. C.H. and K.E.S. provided recording materials for video analyses and wrote the paper. A.R.L. conceived the study, conducted analyses and wrote the paper. All authors are accountable for the content and approved the final version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Funding. This research was supported by the Research Incentive Grant of The Carnegie Trust for the Universities of Scotland (grant no. RIG008132) attributed to A.R.L.

Acknowledgements. We thank Inês Rebelo, Vasilis Louca and Sol Milne for helpful discussion about our methods. We are grateful to our handling editor and to three anonymous reviewers for important suggestions. We thank the keepers at Budongo Trail chimpanzee facility at Edinburgh Zoo, Royal Zoological Society of Scotland and the Wolfgang Kohler Primate Centre, Leipzig Zoo for their support with data collection. We are grateful to the field assistants and directors of Kibale Chimpanzee Project and all of the staff of the Budongo Conservation Field Station for assistance with data collection in Uganda. Thanks to the Ugandan National Council for Science and Technology, the Uganda Wildlife Authority and the President's office for permission to collect data in Uganda and to the Royal Zoological Society of Scotland for providing core funding for Budongo Conservation Field Station.

References

- Lameira AR. 2017 Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neurosci. Biobehav. Rev.* **83**, 429–439. (doi:10.1016/j.neubiorev.2017.09.021)
- Bergman TJ, Beehner JC, Painter MC, Gustison ML. 2019 The speech-like properties of nonhuman primate vocalizations. *Anim. Behav.* **151**, 229–237. (doi:10.1016/j.anbehav.2019.02.015)
- Ghazanfar AA, Liao DA, Takahashi DY. 2019 Volition and learning in primate vocal behaviour. *Anim. Behav.* **151**, 239–247. (doi:10.1016/j.anbehav.2019.01.021)
- Boë L-J, Sawallis TR, Fagot J, Badin P, Barbier G, Captier G, Ménard L, Heim JL, Schwartz JL. 2019 Which way to the dawn of speech? Reanalyzing half a century of debates and data in light of speech science. *Sci. Adv.* **5**, eaaw3916. (doi:10.1126/sciadv.aaw3916)
- Chandrasekaran C, Trubanova A, Stillitano S, Caplier A, Ghazanfar AA. 2009 The natural statistics of audiovisual speech. *PLoS Comput. Biol.* **5**, e1000436. (doi:10.1371/journal.pcbi.1000436)
- Ladefoged P, Maddieson I. 1996 *The sounds of the world's languages*. Oxford, UK: John Wiley & Sons.
- Ladefoged P, Disner S. 2012 *Vowels and consonants*, 3rd edn. Oxford, UK: Wiley-Blackwell.
- Drullman R, Festen JM, Plomp R. 1994 Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* **95**, 2670–2680. (doi:10.1121/1.409836)
- Elliott TM, Theunissen FE. 2009 The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.* **5**, e1000302. (doi:10.1371/journal.pcbi.1000302)
- Ghitza O, Greenberg S. 2009 On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica* **66**, 113–126. (doi:10.1159/000208934)
- Ghazanfar AA, Takahashi DY, Mathur N, Fitch TW. 2012 Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Curr. Biol.* **22**, 1176–1182. (doi:10.1016/j.cub.2012.04.055)
- Toyoda A, Maruhashi T, Malaivijitnond S, Koda H. 2017 Speech-like orofacial oscillations in stump-tailed macaque (*Macaca arctoides*) facial and vocal signals. *Am. J. Phys. Anthropol.* **23**, R268. (doi:10.1002/ajpa.23276)
- Bergman TJ. 2013 Speech-like vocalized lip-smacking in geladas. *Curr. Biol.* **23**, R268–R269. (doi:10.1016/j.cub.2013.02.038)
- Terleph TA, Malaivijitnond S, Reichard U. 2018 An analysis of white-handed gibbon male song reveals speech-like phrases. *Am. J. Phys. Anthropol.* **19**, 252. (doi:10.1002/ajpa.23451)

15. Lameira AR, Hardus ME, Bartlett AM, Shumaker RW, Wich SA, Menken SB. 2015 Speech-like rhythm in a voiced and voiceless orangutan call. *PLoS ONE* **10**, e116136. (doi:10.1371/journal.pone.0116136)
16. Morrill RJ, Paukner A, Ferrari PF, Ghazanfar AA. 2012 Monkey lipsmacking develops like the human speech rhythm. *Dev. Sci.* **15**, 557–568. (doi:10.1111/j.1467-7687.2012.01149.x)
17. Shepherd SV, Freiwald WA. 2018 Functional networks for social communication in the macaque monkey. *Neuron* **99**, 413–420.e3. (doi:10.1016/j.neuron.2018.06.027)
18. Ghazanfar AA, Morrill RJ, Kayser C. 2013 Monkeys are perceptually tuned to facial expressions that exhibit a theta-like speech rhythm. *Proc. Natl Acad. Sci. USA* **110**, 1959–1963. (doi:10.1073/pnas.1214956110)
19. Ghazanfar AA, Takahashi DY. 2014 Facial expressions and the evolution of the speech rhythm. *J. Cogn. Neurosci.* **26**, 1196–1207. (doi:10.1162/jocn_a_00575)
20. Ghazanfar AA, Takahashi DY. 2014 The evolution of speech: vision, rhythm, cooperation. *Trends Cogn. Sci.* **18**, 543–553. (doi:10.1016/j.tics.2014.06.004)
21. MacNeilage PF. 1998 The frame/content theory of evolution of speech production. *Behav. Brain Sci.* **21**, 499–511; discussion 511–546. (doi:10.1017/S0140525X98001265)
22. Fedurek P, Slocombe KE, Hartel JA, Zuberbuhler K. 2015 Chimpanzee lip-smacking facilitates cooperative behaviour. *Sci. Rep.* **5**, 13460. (doi:10.1038/srep13460)
23. Watts DP. 2015 Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: variation, social learning, and possible functions. *Primates* **57**, 61–72. (doi:10.1007/s10329-015-0497-8)
24. Wickham H. 2009 *Ggplot2: elegant graphics for data Analysis*. Berlin, Germany: Springer.
25. Bates D. 2010 lme4: Mixed-effects modeling with R. (Available at <http://lme4.r-forge.r-project.org/IMMwR/lrgprt.pdf>.)
26. Lameira AR, Call J. 2018 Time-space-displaced responses in the orangutan vocal system. *Sci. Adv.* **4**, eaau3401. (doi:10.1126/sciadv.aau3401)
27. Ferrari PF, Visalberghi E, Paukner A, Fogassi L, Ruggiero A, Suomi SJ. 2006 Neonatal imitation in rhesus macaques. *PLoS Biol.* **4**, e302. (doi:10.1371/journal.pbio.0040302)
28. Geissmann T. 1984 Inheritance of song parameters in the Gibbon song, analysed in 2 hybrid Gibbons (*Hylobates pileatus* × *H. lar*). *Folia Primatol.* **42**, 216–235. (doi:10.1159/000156165)