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

André S. Pereira[1,2], Eithne Kavanagh[3], Catherine Hobaiter[1], Katie E. Slocombe[3], Adriano R. Lameira[1,4*]

1. School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom.
2. School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom
3. Department of Psychology, University of York, York, United Kingdom
4. Department of Psychology, University of Warwick, Coventry, United Kingdom

*Corresponding author: adriano.lameira@warwick.ac.uk

Abstract

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 ~4 Hz, closing a gap in the evidence for the evolution of speech-rhythm within primate order. We observed sizeable rhythmic variation within and between chimpanzee populations, with differences of over 2Hz 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.

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 last 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 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).

Methods

Study subjects and data collection

We identified lip-smack bouts present in video recordings collected at Edinburgh Zoo (*Pan troglodytes verus* and one hybrid, UK) (Table 1) during August and September 2013 with a

67   Panasonic HDC SDX1; at Leipzig Zoo (*P. t. verus*, Germany) (Table 1) during June and July

68   2017 with a Panasonic HDC-SD90 camcorder with a Sennheiser MKE 400 microphone

69   attached; and in the wild in the Kanyawara community (*P. t. schweinfurthii*, Kibale National

70   Park, Uganda) (Table 1) during December 2014 and August and September 2016 with a

71   Panasonic HDC-SD90 camcorder with a Sennheiser MKE 400 microphone attached; and from

72   the Waibira community (*P. t. schweinfurthii*, Budongo Forest Research, Uganda) (Table 1)

73   community during December 2011, March 2012, December 2014, and August 2017 with a

74   Panasonic SD90. All videos were 25 frames per second. Videos were selected for analysis

75   when the face of the emitter was clearly visible during lip-smack production and this was the

76   sole criteria to include a bout in the analysis. There was no proactive selection of particular

77   individuals. All videos had been collected during opportunistic observation of the subjects'

78   behaviour.

79   Permission to collect video data had been previously obtained from the authors'

80   institutions (either for other projects or routine data collection) and all the relevant bodies

81   responsible for managing research at each population. All procedures followed the Association

82   for the Study of Animal Behaviour/Animal Behavior Society Guidelines for the Use of Animals

83   in Research (Animal Behaviour, 2018, 135, I-X), all institutional guidelines, the legal

84   requirements of the countries in which the work was carried out, and was granted ethical

85   approval by the Biology Animal Welfare Ethical Review Board (AWERB), University of

86   York.

87

88   Table 1. *Lip-smack data used for analysis.*

Population	# individuals	# bouts (# open-close mouth cycles)/individual
Edinburgh	3 (1 female, 2 males)	*Female*: 8(49)
		Males: 16(104), 7(53.5)
Leipzig	3 (1 female, 2 males)	*Female*: 6(24)
		Males: 1(3), 1(9)
Kanyawara	5 (1 female, 4 males)	*Female*: 1(5)
		Males: 2(6), 2(8), 1(5), 1(3)
Waibira	3 (1 female, 2 males)	*Female*: 1(2)
		Males: 2(9), 5(25)

89

90

91   *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 which we considered to be the approximate rate of mouth oscillation across lip-smack bouts (15, 16). To do this, we used the *imtool* 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) (S2 Supplementary Material, Fig. 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) (S1 Supplementary Material). 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) (S2 Supplementary Material). 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 (S2 Supplementary Material).

125   Finally, we used the R package ggplot2 (24) to plot the smoothed out mean ± 95%

126 confidence interval of the standardized power spectrum density of all time-series and used

127 custom R scripts to find the peak of the curve, i.e., the dominant frequency of chimpanzee lip-

128 smacking behaviour. We standardized all power spectrum density curves by standardizing the

129 spectral power variation (Y-axis) from 0 to 100 following the procedure previously described

130 for the standardization of the time-series. This standardization allowed us to account for the

131 relative spectral power at all frequencies of all bouts while avoiding having individual curves

132 contributing differently to the mean curve. To help visualize the data, we used the same

133 procedure to plot the mean ± 95% confidence interval of the power spectrum density of all

134 time-series of each individual in each population, as well as of each pair of populations. All

135 time-series and each time-series' plot and power spectrum density plot can be found in S1

136 Supplementary Material. All code and steps to replicate the analysis described here are

137 available in S2 Supplementary Material.

138   To statistically compare frequency peaks between captivity and the wild, we used the

139 *glmer* function from the R package *lme4* (25) to build a generalized linear mixed model, which

140 we set up with a gamma error structure and inverse link function; the peak of each individual

141 bout was input as the dependent variable; population (Edinburgh, Leipzig, Kanyawara or

142 Waibira) was input as a fixed factor, and the identity of each individual was input as a random

143 factor to control for repeated measures. We confirmed that the distribution of the residuals was

144 normally distributed and that there was no issue of overdispersion. The code for this analysis

145 can be found in S2 Supplementary Material. Because the highest peak of some individual lip-

146 smack bouts reflected the distribution of inter-bout intervals (typically <1Hz) instead of the

147 real peak, which is a regular occurrence in studies of speech rhythmicity (e.g. 14), we assessed

148 all bouts individually and, for such deviant cases, only included the peaks of the dominant

149 frequency plot (S1 Supplementary Material) that corresponded to the true mean of open-mouth

150 cycles per second, as observed from each bout's time-series (S1 Supplementary Material).

151

152

153 Results

154

155 We found that chimpanzee lip-smacks exhibited a mean rhythm per bout of 4.15Hz (Fig. 1).

156 We identified rhythm variation in lip-smack rate production across individuals who exhibited

the behaviour within and across populations (Fig. 2). For each of the populations, individual lip-smack rhythm spanned a frequency range of at least 1Hz, with maximum differences above 2Hz between some individuals in some of the populations (coloured vertical dashed lines, Fig. 2A-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, Fig. 2E-J). The average (arithmetic mean) of the mean rhythm per population was 3.27Hz. The mean rhythm between the two captive populations was nearly equal. Between the two wild populations there was an observed difference of ~1Hz. Any dyad with a captive vs. wild population exhibited a difference between >1 and <2.5Hz in lip-smack rhythm. To investigate the apparent differences in the rhythm of lip-smacks between captive vs. 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 (S2 Supplementary Material). 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).
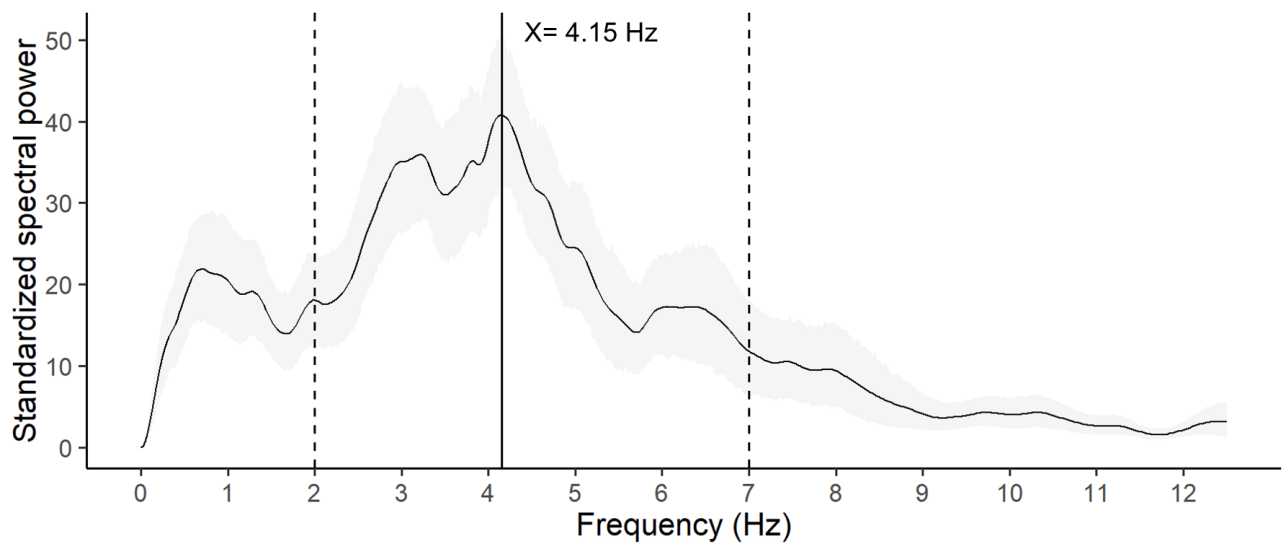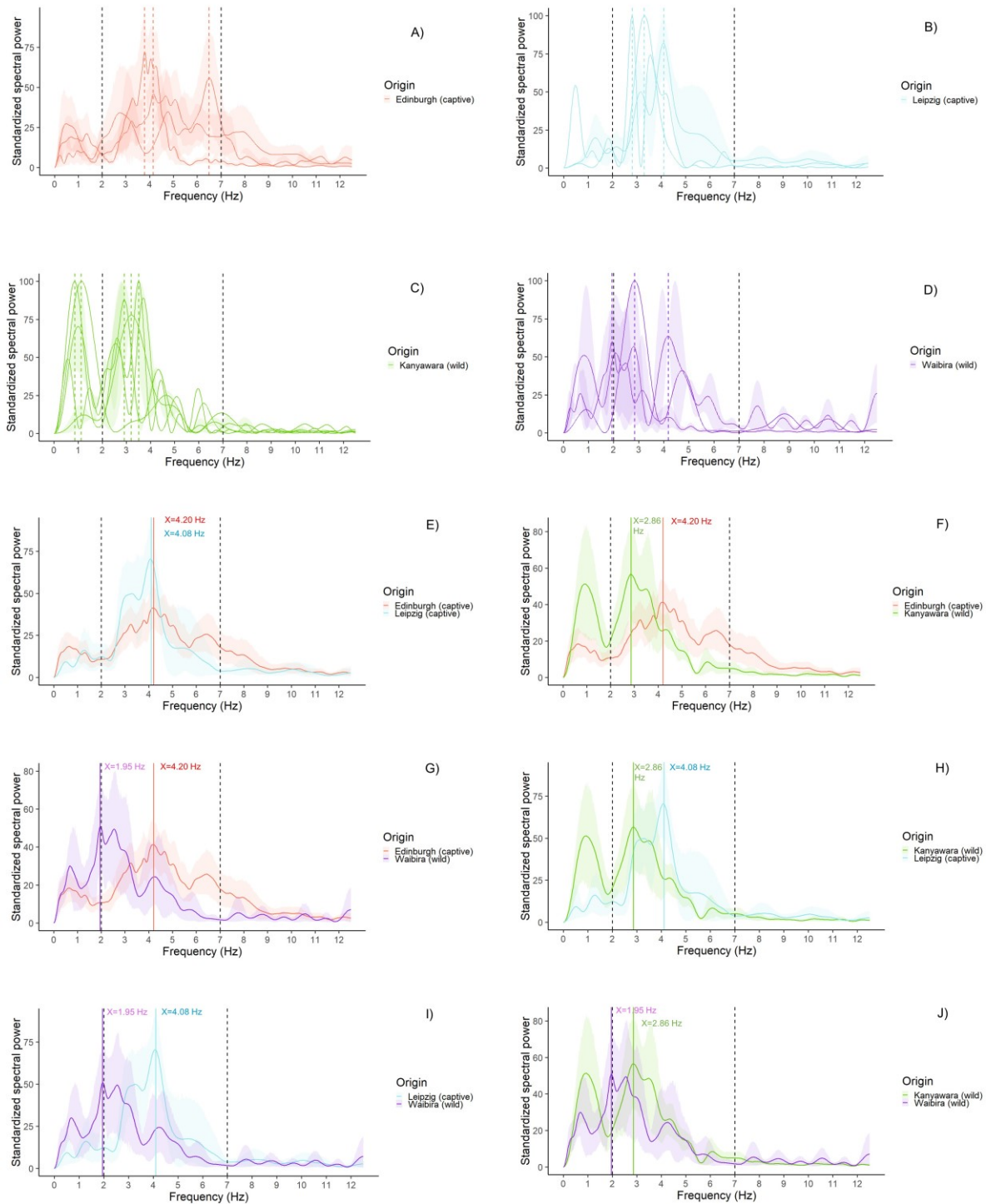
178   Fig. 1. *Mean standardized power spectral density plot of chimpanzee lip-smacking.* The mean
179   ± 95% confidence interval standardized power spectral density plot of all 54 analysed
180   chimpanzee lip-smack bouts peaks at 4.15 Hz, which represents the dominant frequency of
181   chimpanzee lip-smack production rate.

182

183

184

Fig. 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 ± 95% confidence interval standardized power spectral density plot per individual (A-D) and per population (E-J). Coloured dashed vertical lines indicate max frequency peak per individual (A-D) and per population (E-J). Black dashed vertical lines indicate limits of speech-like rhythm frequencies.

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 2Hz 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 by 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 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 >1.5Hz 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, 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 as their group members (Hobaiter, 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 distributions 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.

Acknowledgements

Competing Interests

The authors declare that they have no conflict of interest.

Author contributions

ASP and EK conducted analyses and wrote the paper. CH and KES provided materials and wrote the paper. ARL conceived the study, conducted analyses and wrote the paper.

Funding

Bibliography

1. A. R. Lameira, Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews* **83**, 429–439 (2017).

2. T. J. Bergman, J. C. Beehner, M. C. Painter, M. L. Gustison, The speech-like properties of nonhuman primate vocalizations. *Anim Behav* (2019) https:/doi.org/10.1016/j.anbehav.2019.02.015.

3. A. A. Ghazanfar, D. A. Liao, D. Y. Takahashi, Volition and learning in primate vocal behaviour. *Anim Behav* (2019) https:/doi.org/10.1016/j.anbehav.2019.01.021.

4. L.-J. Boë, *et al.*, Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Sci. Adv.* **5**, eaaw3916 (2019).

5. C. Chandrasekaran, A. Trubanova, S. e bastien Stillittano, A. Caplier, A. A. Ghazanfar, The natural statistics of audiovisual speech. *PLoS computational biology* **5**, e1000436 (2009).

6. P. Ladefoged, I. Maddieson, *The Sounds of the World's Languages* (John Wiley {& Sons, 1996).

7. P. Ladefoged, S. Disner, *Vowels and Consonants*, 3rd Ed. (Wiley-Blackwell, 2012).

8. R. Drullman, J. M. Festen, R. Plomp, Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America* **95**, 2670–2680 (1994).

9. T. M. Elliott, F. E. Theunissen, The Modulation Transfer Function for Speech Intelligibility. *PLoS Comput Biol* **5**, e1000302 (2009).

10. O. Ghitza, S. Greenberg, 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 (2009).

11. A. A. Ghazanfar, D. Y. Takahashi, N. Mathur, T. W. Fitch, Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current biology : CB* **22**, 1176–1182 (2012).

12. A. Toyoda, T. Maruhashi, S. Malaivijitnond, H. Koda, Speech-like orofacial oscillations in stump-tailed macaque (Macaca arctoides) facial and vocal signals. *American journal of physical anthropology* **23**, R268 (2017).

13. T. J. Bergman, Speech-like vocalized lip-smacking in geladas. *Current biology : CB* **23**, R268–R269 (2013).

286   14.   T. A. Terleph, S. Malaivijitnond, U. Reichard, An analysis of white-handed gibbon male
287   song reveals speech-like phrases. *American journal of physical anthropology* **19**, 252
288   (2018).

289   15.   A. R. Lameira, *et al.*, Speech-like rhythm in a voiced and voiceless orangutan call. *PloS*
290   *one* **10**, e116136 (2015).

291   16.   R. J. Morrill, A. Paukner, P. F. Ferrari, A. A. Ghazanfar, Monkey lipsmacking develops
292   like the human speech rhythm. *Developmental science* **15**, 557–568 (2012).

293   17.   S. V. Shepherd, W. A. Freiwald, Functional Networks for Social Communication in the
294   Macaque Monkey. *Neuron* **99**, 413-420.e3 (2018).

295   18.   A. A. Ghazanfar, R. J. Morrill, C. Kayser, Monkeys are perceptually tuned to facial
296   expressions that exhibit a theta-like speech rhythm. *Proceedings of the National*
297   *Academy of Sciences of the United States of America* **110**, 1959–1963 (2013).

298   19.   A. A. Ghazanfar, D. Y. Takahashi, Facial expressions and the evolution of the speech
299   rhythm. *Journal of cognitive neuroscience* **26**, 1196–1207 (2014).

300   20.   A. A. Ghazanfar, D. Y. Takahashi, The evolution of speech: vision, rhythm,
301   cooperation. *Trends in cognitive sciences* **18**, 543–553 (2014).

302   21.   M. PF, The frame/content theory of evolution of speech production. *Behavioral and*
303   *Brain Sciences* **21**, 499–511– discussion 511–46 (1998).

304   22.   P. Fedurek, K. E. Slocombe, J. A. Hartel, K. Zuberbuhler, Chimpanzee lip-smacking
305   facilitates cooperative behaviour. *Scientific reports* **5**, 13460 (2015).

306   23.   D. P. Watts, Production of grooming-associated sounds by chimpanzees (Pan
307   troglodytes) at Ngogo: variation, social learning, and possible functions. *Primates*, 1–12
308   (2015).

309   24.   H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2009).

310   25.   D. Bates, lme4: Mixed-effects modeling with R (2010).

311   26.   A. R. Lameira, J. Call, Time-space–displaced responses in the orangutan vocal system.
312   *Sci Adv* **4**, eaau3401 (2018).

313   27.   P. F. Ferrari, *et al.*, Neonatal imitation in rhesus macaques. *PLoS Biol* **4**, e302 (2006).

314   28.   T. Geissmann, Inheritance of Song Parameters in the Gibbon Song, Analysed in 2
315   Hybrid Gibbons (Hylobates pileatus X H. lar). *Folia Primatologica* **42**, 216–235 (1984).

316