

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/158212>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

1 **Orangutan information broadcast via consonant-like and vowel-**
2 **like calls breaches mathematical models of linguistic evolution**

3

4 **Adriano R. Lameira^{1,2*}, António Alexandre³, Marco Gamba⁴, Matthew G. Nowak^{5,6}, Raquel**
5 **Vicente³, Serge Wich^{7,8}**

6

7 ¹Department of Psychology, University of Warwick, Coventry, UK

8 ²School of Psychology and Neuroscience, University of St Andrews, Scotland, UK

9 ³Independent researcher

10 ⁴University of Turin, Department of Life Sciences and Systems Biology, Turin, Italy

11 ⁵Sumatran Orangutan Research Program (PanEco-YEL), North Sumatra, Indonesia

12 ⁶Department of Anthropology, Southern Illinois University, Illinois, USA

13 ⁷School of Natural Sciences and Psychology, Liverpool John Moores University, UK

14 ⁸Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, the Netherlands

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

16

17 **Keywords:** Language origin; language evolution; proto-consonants; proto-vowels; great apes;
18 orangutans (*Pongo ssp*)

19

20 **Abstract**

21

22 **The origin of language is one of the most significant evolutionary milestones of life on Earth,**
23 **but one of the most persevering scientific unknowns. Two decades ago, game theorists and**
24 **mathematicians predicted that the first words and grammar emerged as a response to**
25 **transmission errors and information loss in language’s precursor system, however,**
26 **empirical proof is lacking. Here, we assessed information loss in proto-consonants and**
27 **proto-vowels in human pre-linguistic ancestors as proxied by orangutan consonant-like and**
28 **vowel-like calls that compose syllable-like combinations. We played-back and re-recorded**
29 **calls at increasing distances across a structurally complex habitat (i.e. adverse to sound**
30 **transmission). Consonant-like and vowel-like calls degraded acoustically over distance, but**
31 **no information loss was detected regarding three distinct classes of information (viz.**
32 **individual ID, context and population ID). Our results refute prevailing mathematical**
33 **predictions and herald a turning point in language evolution theory and heuristics. Namely,**
34 **explaining how the vocal-verbal continuum was crossed in the hominid family will benefit**
35 **from future mathematical and computational models that, in order to enjoy empirical validity**
36 **and superior explanatory power, will be informed by great ape behavior and repertoire.**

37

38 **Introduction**

39

40 Communication in natural (e.g. human language) and artificial systems (e.g. computer language)
41 rests on three vertices: the encoder, the decoder, and the communication channel linking the two¹.
42 With regards to language origin – the last major evolutionary transition of life on Earth² – much
43 attention has been dedicated to the role of the encoder (its anatomical^{3–6} and motoric attributes^{7–}
44 ¹¹), the receiver (its anatomical^{12,13} and perceptual attributes^{14–18}) and the interactions between the
45 two¹⁹. Surprisingly, however, the role of the channel¹ – the interval between encoder and decoder
46 that a signal must traverse – in the emergence of language has remained virtually ignored²⁰.

47 This knowledge gap is particularly problematical in light of game theory and mathematical
48 models of language evolution^{21–23}. Notably, these models have predicted that the first words and
49 grammatical rules emerged to minimize error and information loss in language’s precursor channel.
50 Regarding word origin, this argument asserts that the lengthier a signal combination, the lower the
51 probability of mistaking signals for each other. Regarding syntax origin, it asserts that the more
52 varied a sequence of signal combinations, the lower the probability of mistaking the events being
53 referred to, with words and syntax having, thus, developed in the human lineage to decrease
54 transmission errors. Without basic knowledge about the communication channel used by our

55 ancestors to broadcast information and its “error limit”^{21–23}, it is impossible, however, to validate
56 these models or their proposed evolutionary scenario.

57 Human evolution unfolded in parallel with acute climate and ecological changes in the
58 African continent²⁴, however, it is unclear when and where the first forms of language manifested
59 among human ancestors. Regardless of whether proto-language originated in the rainforest,
60 woodland or savannah, the hypothesis that the first linguistic structures emerged to avert error can
61 be best tested in forested habitats, which pose the most adverse conditions to sound transmission,
62 and thus, where signal and information limits can be assessed.

63 To implement an the empirical proof of the currently prevailing mathematical models of
64 linguistic evolution, we assessed information loss in wild orangutan voiceless consonant-like and
65 voiced vowel-like calls⁷. These calls exhibit articulatory homology with their human counterparts,
66 and therefore, represent living proxies of spoken language’s putative pre-linguistic units^{25–27}.
67 Namely, we played-back consonant-like kiss-squeaks and vowel-like grumphs²⁸ and re-recorded
68 these calls at increasing distances. Critically, bar humans, orangutans are the only known great
69 ape to produce consonant-like and vowel-like calls combined into syllable-like combinations²⁹,
70 therefore, presenting a privileged hominid model for this study³⁰.

71

72 **Materials and Methods**

73

74 *In brief*

75 Calls were originally recorded from wild orangutan individuals across contexts and populations of
76 Sumatran (*Pongo abelii*) and Bornean orangutans (*P. pygmaeus*). Only consonant- and vowel-like
77 calls that were prevent from the same syllable-like combination were used for playback. We
78 extracted four acoustic parameters over distance. We used individual, contextual and geographical
79 acoustic signatures²⁵ to assess information loss. This setup mimicked the putative proto-
80 combinatoric conditions at the moment of language origin. Methodologically, this allowed to control
81 for biasing factors between consonant- and vowel-like calls (e.g. individuals, context, recording
82 settings).

83

84 *Study site*

85 Playback experiments were conducted at the Sikundur Research Station (3°55′48.07″;
86 98°2′31.17″), Leuser Ecosystem, North Sumatra, Indonesia. The Sikundur forest is located on the
87 eastern forest margin of the Alas river dividing the Leuser Ecosystem along its North-South axis
88 and constituting a major dispersal barrier for orangutans at this altitude³¹. Presently, the
89 forest is a dipterocarp tropical rainforest, comprised of disturbed primary forest and
90 secondary/regrowth forest that was the target of previous logging operations (between 1970 and

91 1980, and later during 1990s³²). Research at the station is managed by the Sumatran Orangutan
92 Conservation Programme (SOCP)–PanEco Foundation. The study was performed in agreement
93 with regulations and permissions from the relevant Indonesian authorities. No animal observation,
94 handling, contact or interaction took place during this study.

95

96 *Data collection*

97 Recordings for the playback playlist were previously collected at three research stations: Tuanan
98 and Gunung Palung (Central and West Kalimantan, respectively, Indonesian Borneo) and Sampan
99 Getek (North Sumatra, Indonesia). The playback playlist included 120, 118 and 249 calls to assess
100 individual ID, context and population ID information, respectively (see more in ESM). Orangutan
101 kiss squeaks²⁸ were used as living proxies of voiceless proto-consonants, orangutan grumphs²⁸ as
102 living proxies of voiced proto-vowels.

103 All kiss-squeaks and grumphs were selected from call combinations composed of the two
104 calls, specifically kiss-squeak+grumph (see *Data Analyses* and ESM). All recordings were set to
105 the same peak amplitude prior to playback using Raven interactive sound analysis (version 1.2.1,
106 Cornell Lab of Ornithology, Ithaca, New York). No further signal transformations were conducted.

107 Playbacks were conducted using a Marantz Digital Recorder PMD-660 (D&M Holdings,
108 Kawasaki, Japan) connected to a Nagra DSM speaker (Audio Technology Switzerland S.A.,
109 Romanel, Switzerland). The speaker was set at 1 - 1,5m from the ground. Because Sikundur is
110 partially a regrowth/secondary forest, with abundant undergrowth below the understory, this height
111 offered a suitable means to explore the effects of complex habitat structure on broadcast
112 performance. Playback volume was set at ~100dB SPL at 1-meter distance to facilitate assessment
113 of sound degradation over distance and was not meant to emulate orangutan natural vocal
114 loudness. Playbacks were conducted between 5h30 and 6h30 local time in absence of wind and
115 with no rain during the previous 48 hours. This time was elected for playbacks because in this
116 habitat, early mornings presented the time of day with least biotic noise. We made no presumptions
117 on whether early human ancestors communicated predominantly at this time. All recordings along
118 the same transect were conducted in the same morning.

119 Playbacks were conducted twice at two locations (i.e. along two transects), once at each
120 location. Re-recordings were conducted every 25m along the two transects across the forest up
121 until 100m away, at which point playbacks became too faint to be analysed. Transects started
122 within 10m from each other and advanced forward in oblique direction one from other. Using
123 different transects allowed to assess the impact of particular phonological features (e.g. larger tree
124 trunks, leaf density) on broadcast performance. Transects were straight, flat and included no
125 obvious canopy openings or clearings. Playbacks were re-recorded using ZOOM H4next Handy
126 Recorder (ZOOM Corporation, Tokyo, Japan) connected to a RØDE NTG-2 directional microphone

127 (RØDE LLC, Sydney, Australia). Audio data were recorded using the WAVE PCM format at 16-bit.
128 The microphone was set at 1 - 1,5m from the ground. Data for distance zero were extracted from
129 the original playback recordings. In total, 7826 calls (incl. original at 0m and re-recordings up to
130 100m) were collected (see ESM for sample breakdown). For each transect, three playbacks
131 sessions were conducted, one for each information type: one playlist comprised recordings varying
132 in individual subjects, the other in context and the other in population.

133

134 *Data measurements*

135 We manually measured four acoustic parameters from all calls using Raven interactive sound
136 analysis (version 1.2.1, Cornell Lab of Ornithology, Ithaca, New York) using the spectrogram
137 window (window type: Hann; 3-dB filter bandwidth: 124 Hz; grid frequency resolution: 2.69 Hz; grid
138 time resolution: 256 samples): duration (s), maximum frequency (Hz), maximum power
139 (uncalibrated dB), and maximum time. Duration was the time difference between call offset and
140 onset. Maximum frequency was the frequency with maximum energy (i.e. power, dB) in a call.
141 Maximum power was the power of the maximum frequency. Maximum time was the moment when
142 the maximum power occurred proportionally to the total duration of a call (e.g. max time=0.5 means
143 it occurred half way the call's duration). These parameters have been found to be strong descriptors
144 of orangutan calls and their informational content^{25,28,33}. Critically, they were extractable from both
145 consonant- and vowel-calls, enabling direct comparison between acoustic and information
146 broadcast performance between the two call categories.

147

148 *Data analyses – Acoustic performance*

149 To assess acoustic broadcast performance during transmission, Linear Mixed Models (model type:
150 III sum of squares; test model terms: Satterthwaite, using restricted maximum likelihood) were
151 conducted using JASP³⁴ (version 0.14.1). One model was generated per acoustic parameter (x4)
152 per call type (x2), with a total of 8 models. Per model, the acoustic parameter was inserted as
153 dependent variable (N=3560 per call type). Distance (treated as ordinal: 0, 25, 50, 75, 100m),
154 transect (2 levels), context (3 levels: towards human observers, tiger-patterned predator-model,
155 plain-white predator-model)²⁹ and population (3 levels: Tuanan, Gunung Palung, Sampan Getek)
156 were inserted as fixed effect variables. Individual (20 levels) and call number (N=249 per call type)
157 were inserted as random effect, since some calls were re-used for different playbacks and from the
158 same individual. Random slopes for distance and transect were allowed to vary per individual. No
159 explicit indication of nested variables (e.g. individual within population) was provided since this is
160 automatically identified by the model (see²⁵ and ESM).

161

162 *Data analyses – Information performance*

163 To assess information broadcast performance, we conducted discriminant function analyses (DFA)
 164 per distance³³. All analyses were based on the four measured acoustic parameters simultaneously.
 165 Six analyses were conducted to test information content (x3; individual ID, context, population ID)
 166 for each call type (x2). LMM results indicated that “transect” had a significant effect acoustic
 167 performance over distance, hence, all (p)DFA analyses were conducted using one transect only.
 168 We conducted DFA with leave-one-out procedure using SPSS (IBM SPSS Statistics, version 27;
 169 ESM) to assess information content about individual identify (same context used across
 170 individuals). To assess information content about context and population, we performed permuted
 171 DFAs (pDFA) with cross-classification³⁵: crossed pDFA for context (to control for individual
 172 variation) and nested pDFA for population (individual variation nested within population; ESM).
 173 pDFA was conducted in R³⁶ with MASS³⁷ and using a function provided by R. Mundry³⁵. Because
 174 crossed pDFAs do not tolerate null data, only three individuals with calls in all contexts were
 175 included. Figures were prepared using ggplot2³⁸ and gridExtra³⁹. A script example was:
 176 pdfa.res=pDFA.crossed(test.fac="Context", contr.fac="Individual", variables=c("Duration", "Max
 177 frequency", "Max time", "Max power"), n.to.sel=NULL, n.sel=100, n.perm=1000,
 178 pdfa.data=test.data).

179

180 Results

181 *Acoustic performance over distance*

182

183 Consonant-like and vowel-like call acoustic parameters changed significantly during transmission
 184 (Table 1, Fig. 1, ESM). This was expected since different parameters interact differentially with the
 185 environment (e.g. max power declines over distance following the general inverse square law of
 186 sound attenuation). Several significant differences were found between transects (ESM),
 187 confirming that acoustic performance was (partly) dictated by the physical structure of the
 188 transmission channel. Context had a significant effect on the acoustic performance of some
 189 parameters (ESM). Given that both call types are known to exhibit marked contextual variation²⁵,
 190 this shows that the acoustic features of different contextual sub-types affect how their transmission
 191 plays out. For both consonant-like and vowel-like calls, population had a significant effect on some
 192 acoustic parameters (ESM), suggesting that geographic accents²⁵ may endow calls with better
 193 transmission properties. Given that forest structure is no longer pristine across virtually all
 194 orangutan sites, it is unclear whether these gains can be attributed to adaptive selection in some
 195 populations.

196

Consonant-like calls (Kiss-squeaks)			Vowel-like calls (Grumps)		
df	F	p	df	F	p

Duration (s)	4, 16.81	14.492	<0.001	4, 20.35	51.298	<0.001
Max frequency (Hz)	4, 19.22	8.453	<0.001	4, 14.11	17.600	<0.001
Max power (dB ^{uncalibrated})	4, 21.34	1825.322	<0.001	4, 23.79	1140.558	<0.001
Max time	4, 14.29	28.214	<0.001	4, 19.25	9.693	<0.001

197 Table 1. Acoustic performance over distance: LMM ANOVA Summary

198 [approximate position of Fig. 1]

199

200 *Information performance over distance*

201

202 Despite poor acoustic performance, informational performance of consonant- and vowel-like calls
 203 was not affected during transmission (Fig. 2). Both call categories allowed correct assessment of
 204 information about individual identity, context and population well above chance levels (Fig. 2).
 205 Information loss was only observed for individual identity when transmitted by vowel-like calls,
 206 however, this effect was only observed when computing a leave-one-out DFA procedure (a more
 207 stringent model) and information performance remained overall above chance (Table. 2; ESM).
 208 Information performance was equivalent between consonant- and vowel-like calls; their trend lines
 209 remained relatively parallel over distance (Fig. 2). Consonant-like calls tended to exhibit higher
 210 percentage of correct assignments, suggesting heavier information load (Fig. 2).

211

212

	Consonant-like calls (Kiss-squeaks)						Vowel-like calls (Grumphs)					
	Individual		Context		Population		Individual		Context		Population	
	norm	L1out	selec	cross	selec	cross	norm	L1out	selec	cross	selec	cross
Spearman's rho	-0.9	-0.8	0.6	-0.5	0.9	-0.6	-0.7	-1	-0.3	-0.8	0.8	-0.5
p	0.083	0.133	0.35	0.45	0.083	0.35	0.233	0.017	0.683	0.133	0.133	0.45

213 Table 2. Information performance over distance: Spearman's Correlation Summary (n=5)

214 norm: correlation based on % correctly classified selected cases using DFA

215 L1out: correlation based on % correctly cross classified using DFA with leave-one-out procedure

216 selec: correlation based on % correctly classified selected cases using pDFA

217 cross: correlation based on % correctly cross classified cases using pDFA

218

219 [approximate position of Fig. 2]

220

221

222 **Discussion**

223

224 We found no evidence for information loss in the only nonhuman living hominid that combines
225 consonant-like and vowel-like calls to produce syllable-like combinations. Information content
226 remained uncompromised until either call type become inaudible, indicating that homologous proto-
227 linguistic units would have remained functionally discriminable as long as they could be heard.
228 Results refute, therefore, mathematical predictions for linguistic evolution.

229 Orangutan consonant-like calls exhibited extreme spectral differences compared with their
230 vowel-like counterparts (i.e. frequency centered at ~4000 vs. 250Hz, respectively, Fig. 1A, 1D).
231 However, both can be information-dense²⁵ and their information performance was equivalent. This
232 suggests that similar results would have been likely when other nonhuman hominid consonant- and
233 vowel-like calls had been selected. Our analyses covered a wide frequency band wherein the actual
234 (but now extinct) proto-linguistic units of language have probably laid.

235 Information loss was assessed by measuring calls' biometric information content (i.e. about
236 individual ID, context and population ID). There is no evidence that other types of informational
237 content (e.g. culturally conventionalized arbitrary information, such as, a word's meaning) transmit
238 differently via the same acoustic signals. Some orangutan consonant-like calls exhibit arbitrary
239 function⁴⁰ and other great ape consonant-like and vowel-like calls are transmitted culturally<sup>7,10,11,41-
240 46</sup>. Thus, these calls are not unescapably limited to the transmission of biometric information, even
241 though this was the information used for our empirical validation.

242 Findings offer three insights into language origin and linguistic evolution. First, proto-
243 consonants and -vowels encoded ample information²⁵ and were resilient against information loss
244 up to 100m distance across channels adverse to signal transmission.

245 Second, the structural complexity of our first linguistic ancestors' habitat was an unlikely
246 source of transmission error and information loss. Paleo-climate change across African habitats
247 brought about major habitat structural changes, and with then, new soundscapes. Open habitats
248 offer few physical obstructions to signal transmission (e.g. savannah), thus, ecological changes
249 happening across Africa are predicted to have diminished channel noise in language's precursor
250 system, not the opposite. Systematic assessment will be required for conclusive resolution.

251 Third, mathematical and computational approaches to language evolution have not, thus
252 far, explicitly or implicitly modeled hominid behavior. Theoretically, current models could apply to
253 any communication system transitioning to a combinatorial state, not necessarily within the hominid
254 family. The fact that language transpired in the human clade, but none other, implies, thus, that
255 "being a hominid" cannot be discounted from theoretical incursions that might stand a chance to
256 enlighten how linguistic evolution ensued from the repertoire of an ape-like ancestor⁴⁷. While
257 current models assuredly encapsulate a possible evolutionary scenario, this was not the one to
258 have likely catalyzed language. The most beneficial future theoretical models will be those that

259 conform with, and factor in, the (consonant-vowel-based) combinatorics shared between great
260 apes and humans.
261

262 **Acknowledgments**

263

264 We thank the Indonesian RISTEK, PHKA, TNGL, BPKEL, Universitas Sumatran Utara and SOCP–
265 PanEco Foundation. We thank Graham Usher for technical support and two anonymous reviews
266 for insightful comments. A.R.L. was supported by the European Union’s Horizon 2020 Research
267 and Innovation Program under the Marie Skłodowska-Curie grant agreement no. 702137 and the
268 UK Research and Innovation’s Future Leaders Fellowship grant agreement no. MR/T04229X/1.

269

270 **References**

271

- 272 1. Shannon, C. A mathematical theory of communication, Part I, Part II. *Bell Syst. Tech. J.* **27**, 623–656
273 (1948).
- 274 2. Szathmáry, E. & Smith, J. M. The major evolutionary transitions. *Nature* **374**, 227–232 (1995).
- 275 3. Lieberman, P., Klatt, D. & Wilson, W. Vocal tract limitations on the vowel repertoires of rhesus
276 monkey and other nonhuman primates. *Science* (1969).
- 277 4. Boë. Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech
278 Precursors. *PLOS ONE* **12**, e0169321 (2017).
- 279 5. Boë, L.-J. *et al.* Which way to the dawn of speech?: Reanalyzing half a century of debates and data in
280 light of speech science. *Sci. Adv.* **5**, eaaw3916 (2019).
- 281 6. Fitch, T. W., Boer, B., Mathur, N. & Ghazanfar, A. A. Monkey vocal tracts are speech-ready. *Science*
282 *Advances* **2**, e1600723–e1600723 (2016).
- 283 7. Lameira, A. R. Bidding evidence for primate vocal learning and the cultural substrates for speech
284 evolution. *Neuroscience & Biobehavioral Reviews* **83**, 429–439 (2017).
- 285 8. Lameira, A. R. & Shumaker, R. W. Orangutans show active voicing through a membranophone. *Sci*
286 *Rep* **9**, 12289 (2019).
- 287 9. Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A. & Shumaker, R. W. Vocal fold control beyond the
288 species-specific repertoire in an orang-utan. *Scientific reports* **6**, 30315 (2016).

- 289 10. Lameira, A. R. *et al.* Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-
290 ended call repertoire: A replication and extension. *Journal of the Acoustical Society of America* **134**,
291 1–11 (2013).
- 292 11. Lameira, A. R. *et al.* Speech-like rhythm in a voiced and voiceless orangutan call. *PloS one* **10**,
293 e116136 (2015).
- 294 12. Ramsier, M. A., Cunningham, A. J., Finneran, J. J. & Dominy, N. J. Social drive and the evolution of
295 primate hearing. *Philosophical transactions of the Royal Society of London. Series B, Biological*
296 *sciences* **367**, 1860–1868 (2012).
- 297 13. Quam, R. *et al.* Early hominin auditory capacities `{\textbar}` Science Advances. *Sci Adv* **1**, e1500355
298 (2015).
- 299 14. Ghazanfar, A. A. Language evolution: neural differences that make a difference. *Nature Neuroscience*
300 **11**, 382–384 (2008).
- 301 15. Schlenker, P., Chemla, E. & Zuberbuhler, K. What Do Monkey Calls Mean? *Trends in Cognitive*
302 *Sciences* **0**, (2016).
- 303 16. Hopkins, W. D. *et al.* Genetic Factors and Orofacial Motor Learning Selectively Influence Variability in
304 Central Sulcus Morphology in Chimpanzees (*Pan troglodytes*). *Journal of Neuroscience* **37**, 5475–
305 5483 (2017).
- 306 17. Lameira, A. R. & Call, J. Time-space–displaced responses in the orangutan vocal system. *Sci Adv* **4**,
307 eaau3401 (2018).
- 308 18. Watson, S. K. *et al.* Nonadjacent dependency processing in monkeys, apes, and humans. *Sci. Adv.* **6**,
309 eabb0725 (2020).
- 310 19. Townsend, S. W. *et al.* Exorcising Grice’s ghost: an empirical approach to studying intentional
311 communication in animals: Intentional communication in animals. *Biological reviews of the*
312 *Cambridge Philosophical Society* **92**, 1427–1433 (2016).

- 313 20. Snowdon, C. T. Chapter 7 Plasticity of Communication in Nonhuman Primates. in *Advances in the*
314 *Study of Behavior* (eds. Marc, N., Klaus, Z., Nicola, C. S. & Vincent, J. M.) 239–276 (Academic Press,
315 2009).
- 316 21. Nowak, M., Krakauer, D. & Dress, A. An error limit for the evolution of language. *Proceedings.*
317 *Biological sciences / The Royal Society* **266**, 2131–2136 (1999).
- 318 22. Plotkin, J. B. & Nowak, M. A. Language Evolution and Information Theory. *Journal of Theoretical*
319 *Biology* **205**, 147–59 (2000).
- 320 23. Nowak, M. & Krakauer, D. The evolution of language. *Proceedings of the National Academy of*
321 *Sciences* **96**, 8028–8033 (1999).
- 322 24. Blumenthal, S. A. *et al.* Aridity and hominin environments. *Proceedings of the National Academy of*
323 *Sciences* **220**, 201700597 (2017).
- 324 25. Lameira, A. R. *et al.* Proto-consonants were information-dense via identical bioacoustic tags to proto-
325 vowels. *Nature Human Behaviour* **1**, 0044 (2017).
- 326 26. Lameira, A. R., Maddieson, I. & Zuberbuhler, K. Primate feedstock for the evolution of consonants.
327 *Trends in cognitive sciences* **18**, 60–62 (2014).
- 328 27. Lameira, A. R. The forgotten role of consonant-like calls in theories of speech evolution. *Behavioral*
329 *and Brain Sciences* **37**, 559–560 (2014).
- 330 28. Hardus, M. E. *et al.* A description of the orangutan’s vocal and sound repertoire, with a focus on
331 geographic variation. in *Orangutans* (eds. Wich, S., Setia, M. T., Utami, S. S. & Schaik, C.) 49–60
332 (Oxford University Press, 2009).
- 333 29. Lameira, A. R. *et al.* Predator guild does not influence orangutan alarm call rates and combinations.
334 *Behavioral Ecology and Sociobiology* **67**, 519–528 (2013).
- 335 30. Lameira, A. R. & Call, J. Understanding Language Evolution: Beyond *Pan* -Centrism. *BioEssays* **42**,
336 1900102 (2020).

- 337 31. Arora, N. *et al.* Effects of Pleistocene glaciations and rivers on the population structure of Bornean
338 orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences* **107**, 21376–21381
339 (2010).
- 340 32. Knop, E., Ward, P. I. & Wich, S. A. A comparison of orang-utan density in a logged and unlogged forest
341 on Sumatra. *Biological Conservation* **120**, 183–188 (2004).
- 342 33. Lameira, A. R. & Wich, S. Orangutan Long Call Degradation and Individuality Over Distance: A
343 Playback Approach. *International Journal of Primatology* **29**, 615–625 (2008).
- 344 34. JASP Team (2020). JASP (v. 0.14.1) [Computer software].
- 345 35. Mundry, R. & Sommer, C. Discriminant function analysis with nonindependent data: consequences
346 and an alternative. *Animal Behaviour* **74**, 965–976 (2007).
- 347 36. Team, R. R: A language and environment for statistical computing. (2013).
- 348 37. Venables, W. N. & Ripley, B. D. Survival Analysis. in *Modern Applied Statistics with S* 353–385
349 (Springer New York, 2002). doi:10.1007/978-0-387-21706-2_13.
- 350 38. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, 2009).
- 351 39. Auguie, B. gridExtra: functions in Grid graphics. *R package version 0.9 1*, (2012).
- 352 40. Lameira, A. R. *et al.* Population-specific use of the same tool-assisted alarm call between two wild
353 orangutan populations (*Pongo pygmaeus wurmbii*) indicates functional arbitrariness. *Plos One* **8**,
354 e69749 (2013).
- 355 41. Wich, S. A. *et al.* Call cultures in orang-utans? *PloS one* **7**, e36180 (2012).
- 356 42. Tagliatela, J. P., Reamer, L., Schapiro, S. J. & Hopkins, W. D. Social learning of a communicative
357 signal in captive chimpanzees. *Biology letters* **8**, 498–501 (2012).
- 358 43. Russell, J. L., Joseph, M., Hopkins, W. D. & Tagliatela, J. P. Vocal learning of a communicative signal
359 in captive chimpanzees, *Pan troglodytes*. *Brain and Language* **127**, 520–525 (2013).
- 360 44. Wich, S. *et al.* A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–
361 64 (2009).

- 362 45. Perlman, M. & Clark, N. Learned vocal and breathing behavior in an enculturated gorilla. *Animal*
363 *cognition* **18**, 1165–1179 (2015).
- 364 46. Hayes, C. *The ape in our house*. (Harper, 1951).
- 365 47. Gomez-Marin, A. & Ghazanfar, A. A. The Life of Behavior. *Neuron* **104**, 25–36 (2019).
- 366

367 **Figure 1.** Spectrographic representation of orangutan consonant-like and vowel-like calls (A),
368 graphic representation of the experimental set up (B) and acoustic performance during
369 transmission (C-F; based on raw data). ^{uncal.}: uncalibrated. Box plots represent median and
370 25-75% interquartile range, whiskers represent lowest/highest value within 1.5 times
371 interquartile range below/above, outliers omitted for clarity. Linear trend lines represented
372 across distance are for visual aid only (based on raw data). *: $p < 0.001$ (LMM ANOVA; see
373 Table 1).

374

375 **Figure 2.** Graphic representation of information performance of orangutan consonant-like and
376 vowel-like calls during transmission, as measured by percentage of correctly assigned
377 cases over distance. Black dotted lines: chance level. A: Continuous lines: Correctly
378 classified cases (DFA); Dashed lines: Correctly classified cross-validated cases (DFA
379 Leave-one-out). B and C: Continuous lines: Correctly cross classified cases (pDFA);
380 Dashed lines: Correctly classified selected cases (pDFA). *: $p < 0.05$ (Spearman's
381 correlation; see Table 2).

382