



## LJMU Research Online

Lameira, AR, Alexandre, A, Gamba, M, Nowak, MG, Vicente, R and Wich, SA

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

<http://researchonline.ljmu.ac.uk/id/eprint/15527/>

### Article

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Lameira, AR, Alexandre, A, Gamba, M, Nowak, MG, Vicente, R and Wich, SA (2021) Orangutan information broadcast via consonant-like and vowel-like calls breaches mathematical models of linguistic evolution. *Biology Letters*, 17 (9). ISSN 1744-9561**

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact [researchonline@ljmu.ac.uk](mailto:researchonline@ljmu.ac.uk)

<http://researchonline.ljmu.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<sup>1,2\*</sup>, António Alexandre<sup>3</sup>, Marco Gamba<sup>4</sup>, Matthew G. Nowak<sup>5,6</sup>, Raquel**  
5 **Vicente<sup>3</sup>, Serge Wich<sup>7,8</sup>**

6  
7 <sup>1</sup>Department of Psychology, University of Warwick, Coventry, UK

8 <sup>2</sup>School of Psychology and Neuroscience, University of St Andrews, Scotland, UK

9 <sup>3</sup>Independent researcher

10 <sup>4</sup>University of Turin, Department of Life Sciences and Systems Biology, Turin, Italy

11 <sup>5</sup>Sumatran Orangutan Research Program (PanEco-YEL), North Sumatra, Indonesia

12 <sup>6</sup>Department of Anthropology, Southern Illinois University, Illinois, USA

13 <sup>7</sup>School of Natural Sciences and Psychology, Liverpool John Moores University, UK

14 <sup>8</sup>Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, the Netherlands

15 \*Corresponding author: [adriano.lameira@warwick.ac.uk](mailto: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 two<sup>1</sup>.  
42 With regards to language origin – the last major evolutionary transition of life on Earth<sup>2</sup> – much  
43 attention has been dedicated to the role of the encoder (its anatomical<sup>3–6</sup> and motoric attributes<sup>7–</sup>  
44 <sup>11</sup>), the receiver (its anatomical<sup>12,13</sup> and perceptual attributes<sup>14–18</sup>) and the interactions between the  
45 two<sup>19</sup>. Surprisingly, however, the role of the channel<sup>1</sup> – the interval between encoder and decoder  
46 that a signal must traverse – in the emergence of language has remained virtually ignored<sup>20</sup>.

47 This knowledge gap is particularly problematical in light of game theory and mathematical  
48 models of language evolution<sup>21–23</sup>. 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”<sup>21–23</sup>, 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 continent<sup>24</sup>, 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 calls<sup>7</sup>. These calls exhibit articulatorily homology with their human counterparts,  
66 and therefore, represent living proxies of spoken language’s putative pre-linguistic units<sup>25–27</sup>.  
67 Namely, we played-back consonant-like kiss-squeaks and vowel-like grumphs<sup>28</sup> 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 combinations<sup>29</sup>,  
70 therefore, presenting a privileged hominid model for this study<sup>30</sup>.

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 prevenient from the same syllable-like combination were used for playback. We  
78 extracted four acoustic paraments over distance. We used individual, contextual and geographical  
79 acoustic signatures<sup>25</sup> 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 barrier dispersal barrier for orangutans at this altitude<sup>31</sup>. 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 1990s<sup>32</sup>).

92

### 93 *Data collection*

94 Recordings for the playback playlist were previously collected at three research stations: Tuanan  
95 and Gunung Palung (Central and West Kalimantan, respectively, Indonesian Borneo) and Sampan  
96 Getek (North Sumatra, Indonesia). The playback playlist included 120, 118 and 249 calls to assess  
97 individual ID, context and population ID information, respectively (see more in ESM). Orangutan  
98 kiss squeaks<sup>28</sup> were used as living proxies of voiceless proto-consonants, orangutan grumphs<sup>28</sup> as  
99 living proxies of voiced proto-vowels.

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

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

116 Playbacks were conducted twice at two locations (i.e. along two transects), once at each  
117 location. Re-recordings were conducted every 25m along the two transects across the forest up  
118 until 100m away, at which point playbacks became too faint to be analysed. Transects started  
119 within 10m from each other and advanced forward in oblique direction one from other. Using  
120 different transects allowed to assess the impact of particular phonological features (e.g. larger tree  
121 trunks, leaf density) on broadcast performance. Transects were straight, flat and included no  
122 obvious canopy openings or clearings. Playbacks were re-recorded using ZOOM H4next Handy  
123 Recorder (ZOOM Corporation, Tokyo, Japan) connected to a RØDE NTG-2 directional microphone  
124 (RØDE LLC, Sydney, Australia). Audio data were recorded using the WAVE PCM format at 16-bit.  
125 The microphone was set at 1 - 1,5m from the ground. Data for distance zero were extracted from

126 the original playback recordings. In total, 7826 calls (incl. original at 0m and re-recordings up to  
127 100m) were collected (see ESM for sample breakdown). For each transect, three playbacks  
128 sessions were conducted, one for each information type: one playlist comprised recordings varying  
129 in individual subjects, the other in context and the other in population.

130

#### 131 *Data measurements*

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

144

#### 145 *Data analyses – Acoustic performance*

146 To assess acoustic broadcast performance during transmission, Linear Mixed Models (model type:  
147 III sum of squares; test model terms: Satterthwaite, using restricted maximum likelihood) were  
148 conducted using JASP<sup>34</sup> (version 0.14.1). One model was generated per acoustic parameter (x4)  
149 per call type (x2), with a total of 8 models. Per model, the acoustic parameter was inserted as  
150 dependent variable (N=3560 per call type). Distance (treated as ordinal: 0, 25, 50, 75, 100m),  
151 transect (2 levels), context (3 levels: towards human observers, tiger-patterned predator-model,  
152 plain-white predator-model)<sup>29</sup> and population (3 levels: Tuanan, Gunung Palung, Sampan Getek)  
153 were inserted as fixed effect variables. Individual (20 levels) and call number (N=249 per call type)  
154 were inserted as random effect, since some calls were re-used for different playbacks and from the  
155 same individual. Random slopes for distance and transect were allowed to vary per individual. No  
156 explicit indication of nested variables (e.g. individual within population) was provided since this is  
157 automatically identified by the model (see<sup>25</sup> and ESM).

158

#### 159 *Data analyses – Information performance*

160 To assess information broadcast performance, we conducted discriminant function analyses (DFA)  
161 per distance<sup>33</sup>. All analyses were based on the four measured acoustic parameters simultaneously.

162 Six analyses were conducted to test information content (x3; individual ID, context, population ID)  
 163 for each call type (x2). LMM results indicated that “transect” had a significant effect acoustic  
 164 performance over distance, hence, all (p)DFA analyses were conducted using one transect only.  
 165 We conducted DFA with leave-one-out procedure using SPSS (IBM SPSS Statistics, version 27;  
 166 ESM) to assess information content about individual identify (same context used across  
 167 individuals). To assess information content about context and population, we performed permuted  
 168 DFAs (pDFA) with cross-classification<sup>35</sup>: crossed pDFA for context (to control for individual  
 169 variation) and nested pDFA for population (individual variation nested within population; ESM).  
 170 pDFA was conducted in R<sup>36</sup> with MASS<sup>37</sup> and using a function provided by R. Mundry<sup>35</sup>. Because  
 171 crossed pDFAs do not tolerate null data, only three individuals with calls in all contexts were  
 172 included. Figures were prepared using ggplot2<sup>38</sup> and gridExtra<sup>39</sup>. A script example was:  
 173 pdfa.res=pDFA.crossed(test.fac="Context", contr.fac="Individual", variables=c("Duration", "Max  
 174 frequency", "Max time", "Max power"), n.to.sel=NULL, n.sel=100, n.perm=1000,  
 175 pdfa.data=test.data).

176

## 177 **Results**

### 178 *Acoustic performance over distance*

179

180 Consonant-like and vowel-like call acoustic parameters changed significantly during transmission  
 181 (Table 1, Fig. 1, ESM). This was expected since different parameters interact differentially with the  
 182 environment (e.g. max power declines over distance following the general inverse square law of  
 183 sound attenuation). Several significant differences were found between transects (ESM),  
 184 confirming that acoustic performance was (partly) dictated by the physical structure of the  
 185 transmission channel. Context had a significant effect on the acoustic performance of some  
 186 parameters (ESM). Given that both call types are known to exhibit marked contextual variation<sup>25</sup>,  
 187 this shows that the acoustic features of different contextual sub-types affect how their transmission  
 188 plays out. For both consonant-like and vowel-like calls, population had a significant effect on some  
 189 acoustic parameters (ESM), suggesting that geographic accents<sup>25</sup> may endow calls with better  
 190 transmission properties. Given that forest structure is no longer pristine across virtually all  
 191 orangutan sites, it is unclear whether these gains can be attributed to adaptive selection in some  
 192 populations.

193

	Consonant-like calls (Kiss-squeaks)			Vowel-like calls (Grumps)		
	df	F	p	df	F	p
Duration (s)	<b>4, 16.81</b>	<b>14.492</b>	<b>&lt;0.001</b>	<b>4, 20.35</b>	<b>51.298</b>	<b>&lt;0.001</b>
Max frequency (Hz)	<b>4, 19.22</b>	<b>8.453</b>	<b>&lt;0.001</b>	<b>4, 14.11</b>	<b>17.600</b>	<b>&lt;0.001</b>

Max power (dB <sup>uncalibrated</sup> )	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

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

195 [approximate position of Fig. 1]

196

197 *Information performance over distance*

198

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

208

209

	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	<b>0.017</b>	0.683	0.133	0.133	0.45

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

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

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

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

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

215

216 [approximate position of Fig. 2]

217

218

219 **Discussion**

220



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

226 Orangutan consonant-like calls exhibited extreme spectral differences compared with their  
227 vowel-like counterparts (i.e. frequency centered at ~4000 vs. 250Hz, respectively, Fig. 1A, 1D).  
228 However, both can be information-dense<sup>25</sup> and their information performance was equivalent. This  
229 suggests that similar results would have been likely when other nonhuman hominid consonant- and  
230 vowel-like calls had been selected. Our analyses covered a wide frequency band wherein the actual  
231 (but now extinct) proto-linguistic units of language have probably laid.

232 Information loss was assessed by measuring calls' biometric information content (i.e. about  
233 individual ID, context and population ID). There is no evidence that other types of informational  
234 content (e.g. culturally conventionalized arbitrary information, such as, a word's meaning) transmit  
235 differently via the same acoustic signals. Some orangutan consonant-like calls exhibit arbitrary  
236 function<sup>40</sup> and other great ape consonant-like and vowel-like calls are transmitted culturally<sup>7,10,11,41-  
237 46</sup>. Thus, these calls are not unescapably limited to the transmission of biometric information, even  
238 though this was the information used for our empirical validation.

239 Findings offer three insights into language origin and linguistic evolution. First, proto-  
240 consonants and -vowels encoded ample information<sup>25</sup> and were resilient against information loss  
241 up to 100m distance across channels adverse to signal transmission.

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

248 Third, mathematical and computational approaches to language evolution have not, thus  
249 far, explicitly or implicitly modeled hominid behavior. Theoretically, current models could apply to  
250 any communication system transitioning to a combinatorial state, not necessarily within the hominid  
251 family. The fact that language transpired in the human clade, but none other, implies, thus, that  
252 "being a hominid" cannot be discounted from theoretical incursions that might stand a chance to  
253 enlighten how linguistic evolution ensued from the repertoire of an ape-like ancestor<sup>47</sup>. While  
254 current models assuredly encapsulate a possible evolutionary scenario, this was not the one to  
255 have likely catalyzed language. The most beneficial future theoretical models will be those that

256 conform with, and factor in, the (consonant-vowel-based) combinatorics shared between great  
257 apes and humans.  
258

259 **Acknowledgments**

260

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

266

267 **References**

268

- 269 1. Shannon, C. A mathematical theory of communication, Part I, Part II. *Bell Syst. Tech. J.* **27**, 623–656  
270 (1948).
- 271 2. Szathmáry, E. & Smith, J. M. The major evolutionary transitions. *Nature* **374**, 227–232 (1995).
- 272 3. Lieberman, P., Klatt, D. & Wilson, W. Vocal tract limitations on the vowel repertoires of rhesus  
273 monkey and other nonhuman primates. *Science* (1969).
- 274 4. Boë. Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech  
275 Precursors. *PLOS ONE* **12**, e0169321 (2017).
- 276 5. Boë, L.-J. *et al.* Which way to the dawn of speech?: Reanalyzing half a century of debates and data in  
277 light of speech science. *Sci. Adv.* **5**, eaaw3916 (2019).
- 278 6. Fitch, T. W., Boer, B., Mathur, N. & Ghazanfar, A. A. Monkey vocal tracts are speech-ready. *Science*  
279 *Advances* **2**, e1600723–e1600723 (2016).
- 280 7. Lameira, A. R. Bidding evidence for primate vocal learning and the cultural substrates for speech  
281 evolution. *Neuroscience & Biobehavioral Reviews* **83**, 429–439 (2017).
- 282 8. Lameira, A. R. & Shumaker, R. W. Orangutans show active voicing through a membranophone. *Sci*  
283 *Rep* **9**, 12289 (2019).
- 284 9. Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A. & Shumaker, R. W. Vocal fold control beyond the  
285 species-specific repertoire in an orang-utan. *Scientific reports* **6**, 30315 (2016).

- 286 10. Lameira, A. R. *et al.* Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-  
287 ended call repertoire: A replication and extension. *Journal of the Acoustical Society of America* **134**,  
288 1–11 (2013).
- 289 11. Lameira, A. R. *et al.* Speech-like rhythm in a voiced and voiceless orangutan call. *PloS one* **10**,  
290 e116136 (2015).
- 291 12. Ramsier, M. A., Cunningham, A. J., Finneran, J. J. & Dominy, N. J. Social drive and the evolution of  
292 primate hearing. *Philosophical transactions of the Royal Society of London. Series B, Biological*  
293 *sciences* **367**, 1860–1868 (2012).
- 294 13. Quam, R. *et al.* Early hominin auditory capacities `{\textbar}` Science Advances. *Sci Adv* **1**, e1500355  
295 (2015).
- 296 14. Ghazanfar, A. A. Language evolution: neural differences that make a difference. *Nature Neuroscience*  
297 **11**, 382–384 (2008).
- 298 15. Schlenker, P., Chemla, E. & Zuberbuhler, K. What Do Monkey Calls Mean? *Trends in Cognitive*  
299 *Sciences* **0**, (2016).
- 300 16. Hopkins, W. D. *et al.* Genetic Factors and Orofacial Motor Learning Selectively Influence Variability in  
301 Central Sulcus Morphology in Chimpanzees (*Pan troglodytes*). *Journal of Neuroscience* **37**, 5475–  
302 5483 (2017).
- 303 17. Lameira, A. R. & Call, J. Time-space–displaced responses in the orangutan vocal system. *Sci Adv* **4**,  
304 eaau3401 (2018).
- 305 18. Watson, S. K. *et al.* Nonadjacent dependency processing in monkeys, apes, and humans. *Sci. Adv.* **6**,  
306 eabb0725 (2020).
- 307 19. Townsend, S. W. *et al.* Exorcising Grice’s ghost: an empirical approach to studying intentional  
308 communication in animals: Intentional communication in animals. *Biological reviews of the*  
309 *Cambridge Philosophical Society* **92**, 1427–1433 (2016).

- 310 20. Snowdon, C. T. Chapter 7 Plasticity of Communication in Nonhuman Primates. in *Advances in the*  
311 *Study of Behavior* (eds. Marc, N., Klaus, Z., Nicola, C. S. & Vincent, J. M.) 239–276 (Academic Press,  
312 2009).
- 313 21. Nowak, M., Krakauer, D. & Dress, A. An error limit for the evolution of language. *Proceedings.*  
314 *Biological sciences / The Royal Society* **266**, 2131–2136 (1999).
- 315 22. Plotkin, J. B. & Nowak, M. A. Language Evolution and Information Theory. *Journal of Theoretical*  
316 *Biology* **205**, 147–59 (2000).
- 317 23. Nowak, M. & Krakauer, D. The evolution of language. *Proceedings of the National Academy of*  
318 *Sciences* **96**, 8028–8033 (1999).
- 319 24. Blumenthal, S. A. *et al.* Aridity and hominin environments. *Proceedings of the National Academy of*  
320 *Sciences* **220**, 201700597 (2017).
- 321 25. Lameira, A. R. *et al.* Proto-consonants were information-dense via identical bioacoustic tags to proto-  
322 vowels. *Nature Human Behaviour* **1**, 0044 (2017).
- 323 26. Lameira, A. R., Maddieson, I. & Zuberbuhler, K. Primate feedstock for the evolution of consonants.  
324 *Trends in cognitive sciences* **18**, 60–62 (2014).
- 325 27. Lameira, A. R. The forgotten role of consonant-like calls in theories of speech evolution. *Behavioral*  
326 *and Brain Sciences* **37**, 559–560 (2014).
- 327 28. Hardus, M. E. *et al.* A description of the orangutan’s vocal and sound repertoire, with a focus on  
328 geographic variation. in *Orangutans* (eds. Wich, S., Setia, M. T., Utami, S. S. & Schaik, C.) 49–60  
329 (Oxford University Press, 2009).
- 330 29. Lameira, A. R. *et al.* Predator guild does not influence orangutan alarm call rates and combinations.  
331 *Behavioral Ecology and Sociobiology* **67**, 519–528 (2013).
- 332 30. Lameira, A. R. & Call, J. Understanding Language Evolution: Beyond *Pan* -Centrism. *BioEssays* **42**,  
333 1900102 (2020).

- 334 31. Arora, N. *et al.* Effects of Pleistocene glaciations and rivers on the population structure of Bornean  
335 orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences* **107**, 21376–21381  
336 (2010).
- 337 32. Knop, E., Ward, P. I. & Wich, S. A. A comparison of orang-utan density in a logged and unlogged forest  
338 on Sumatra. *Biological Conservation* **120**, 183–188 (2004).
- 339 33. Lameira, A. R. & Wich, S. Orangutan Long Call Degradation and Individuality Over Distance: A  
340 Playback Approach. *International Journal of Primatology* **29**, 615–625 (2008).
- 341 34. JASP Team (2020). JASP (v. 0.14.1) [Computer software].
- 342 35. Mundry, R. & Sommer, C. Discriminant function analysis with nonindependent data: consequences  
343 and an alternative. *Animal Behaviour* **74**, 965–976 (2007).
- 344 36. Team, R. R: A language and environment for statistical computing. (2013).
- 345 37. Venables, W. N. & Ripley, B. D. Survival Analysis. in *Modern Applied Statistics with S* 353–385  
346 (Springer New York, 2002). doi:10.1007/978-0-387-21706-2\_13.
- 347 38. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, 2009).
- 348 39. Auguie, B. gridExtra: functions in Grid graphics. *R package version 0.9 1*, (2012).
- 349 40. Lameira, A. R. *et al.* Population-specific use of the same tool-assisted alarm call between two wild  
350 orangutan populations (*Pongo pygmaeus wurmbii*) indicates functional arbitrariness. *Plos One* **8**,  
351 e69749 (2013).
- 352 41. Wich, S. A. *et al.* Call cultures in orang-utans? *PloS one* **7**, e36180 (2012).
- 353 42. Taglialatela, J. P., Reamer, L., Schapiro, S. J. & Hopkins, W. D. Social learning of a communicative  
354 signal in captive chimpanzees. *Biology letters* **8**, 498–501 (2012).
- 355 43. Russell, J. L., Joseph, M., Hopkins, W. D. & Taglialatela, J. P. Vocal learning of a communicative signal  
356 in captive chimpanzees, *Pan troglodytes*. *Brain and Language* **127**, 520–525 (2013).
- 357 44. Wich, S. *et al.* A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–  
358 64 (2009).

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

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

371

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

379