Supplementary Materials

Authors: Valentin Thouzeau[1,2,*], Guillaume Dezecache[3], Philippe Schlenker[3,4], Ewan Dunbar[5,6], Emmanuel Chemla[2,†], Robin Ryder[1,†].

Affiliations:

1 Centre de Recherche en Mathématiques de la Décision, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, Paris, France.

2 Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS.

3 Institut Jean-Nicod, Département d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS.

4 Department of Linguistics, New York University, New York, New York.

5 Université de Paris, Laboratoire de Linguistique Formelle, CNRS.

6 Cognitive Machine Learning, ENS, PSL University, EHESS, CNRS, INRIA.

[*] Corresponding author, thouzeau@ceremade.dauphine.fr

[†] Joint senior authors

This file includes:

Details of the methods
Reference of the database
Table S2
Table S3
References (14 – 17)

Other Supplementary Materials for this manuscript include:

Table S1

DETAILS OF THE METHODS

Bayesian inferences

We performed Bayesian inferences using the *RevBayes* software (14). This software provides a large set of tools for phylogenetic analyses, including models for discrete character data evolutions.
We constructed each Q matrix according to each tested hypothesis, using the *R* software. We then specified the distribution of the parameters, as detailed below :

- The birth rate r_birth was drawn in an exponential distribution of parameter 0.1.
- The death rate r_death was drawn in an exponential distribution of parameter 0.1.
- The maximum of the sigmoid function R_birth was drawn in an exponential distribution of parameter 0.1.
- L the shape of the function was drawn in a uniform distribution of parameter [0, 10].
- The inflexion point of the function was drawn in a uniform distribution of parameter $[0, n_{max} + 1]$, with n_{max} the maximum number of calls.
- The rate of shift r_{shift} in an exponential distribution of parameter 0.1.
- Each rate of birth depending on the hierarchy of the calls were drawn in an exponential distribution of parameter 0.1.

We performed an estimation of those parameters for each experiment, running 10,000 generations with a burning phase of 10%. We verified that the effective sample sizes of all parameters exceeded 200, and we visually checked that the traces were standard using the Tracer software (16). We then performed a marginal likelihood estimation, using the Stepping Stone sampler method (17) with 50 stones, and running 10,000 generations with a burning phase of 10%. We then computed the log Bayes Factors using the estimated marginal likelihood of each model. We repeated with a different prior (exponential with mean 0.5); the results are in agreement with those presented here.

Model specification

In this subsection, we specify in detail the models for the Alarm category; other categories are specified similarly.
There are 5 possible meanings in the Alarm category. Each of these may be present or absent in a species' communication system, giving $2^5=32$ possible lexicons. The Q matrix is thus of size 32×32. Each lexicon can be represented by a binary vector of length 5: a 1 in position k indicates that meaning k is present in this species' system; a 0 indicates that it is absent.
A model is defined by the specification of the shape of the Q matrix.
In the following, let $i=(i_1, \ldots, i_5)$ and $j=(j_1, \ldots, j_5)$ be two lexicons, i.e. two elements of $\{0,1\}^5$. Let $A_{ij}=\sum(i_k-j_k)$ be the number of meanings present in i and not j, and $B_{ij}=\sum(j_k-i_k)$ the number of meanings present in j and not i.

<u>Model 1</u>

The value of Q_{ij} is:

- r_{birth} if $A_{ij}=0$ and $B_{ij}=1$: to go from lexicon i to lexicon j, 1 call needs to appear;
- r_{death} if $A_{ij}=1$ and $B_{ij}=0$: to go from lexicon i to lexicon j, 1 call needs to die;
- 0 otherwise

<u>Model 2</u>

The value of Q_{ij} is:

- r_{birth}^n = R_birth * (1 / (1 + exp(L * (n - D)) if $A_{ij}=0$, $B_{ij}=1$, with n= Σ_{jk}: to go from lexicon i to lexicon j, 1 call needs to appear;
- r_{death} if $A_{ij}=1$ and $B_{ij}=0$: to go from lexicon i to lexicon j, 1 call needs to die;
- 0 otherwise

(The distinction with Model 1 is that this one uses r_{birth}^{n} instead of r_{birth}.)

Model 3

The value of Qij is:

- r_{birth} if A_{ij}=0 and B_{ij}=1: to go from lexicon i to lexicon j, 1 call needs to appear;
- r_{death} if A_{ij}=1 and B_{ij}=0: to go from lexicon i to lexicon j, 1 call needs to die;
- r_{shift} if A_{ij}=1 and B_{ij}=1: to go from lexicon i to lexicon j, 1 call needs to shift meaning;
- 0 otherwise

(The distinction with Model 1 is the introduction of r_shift.)

Model 4

The value of Q_{ij} is:

- $r_{birth}^{n} = R_{birth} * (1 / (1 + \exp(L * (n - D))$ if A_{ij}=0, B_{ij}=1, with n= Σ_{jk}: to go from lexicon i to lexicon j, 1 call needs to appear;
- r_{death} if A_{ij}=1 and B_{ij}=0: to go from lexicon i to lexicon j, 1 call needs to die;
- r_{shift} if A_{ij}=1 and B_{ij}=1: to go from lexicon i to lexicon j, 1 call needs to shift meaning;
- 0 otherwise

(r_{birth}^{n} is as in Model 2; r_{shift} is as in Model 3.)

Model 5

Model 5 is as Model 1, except that the common rates r_{birth} and r_{death} are replaced with $r_{birth}^{General}$, $r_{birth}^{Intermediate}$, $r_{birth}^{Specific}$, $r_{death}^{General}$, $r_{death}^{Intermediate}$, $r_{death}^{Specific}$ depending on whether the meaning to create or remove is general, intermediate, or specific.

Model 6

In order to specified a more complex model of evolution of primate calls, we specified a collection of processes of evolution, taking into account in a more complete way the proposed hierarchy of the calls than the experiment 5, as following :

- Calls may appear at rate r_{birth}.
- Calls may disappear at rate r_{death}.
- The most general call of each category may appear at rate r_{birth}^{0}, different from r_{birth}.
- The most general call of each category may disappear at rate r_{death}^{0}, different from r_{death}.
- Calls with an immediately more general call existing may appear with a rate r_{birth}^{1}, different from r_{birth} and r_{birth}^{0}.
- Calls may split into two immediately more specialized calls at rate r_{split}.
- Calls may specialize into one immediately more specialized call at rate r_{spec}.

We computed the marginal likelihoods of this model for the three types of calls. It is not supported by the data: there is decisive evidence against it in the Alarm category (Bayes Factor = 2.32), and no evidence against it in the Food category (Bayes Factor = 0.02).

Non-Disturbance Calls

In addition to the analyses presented in the article, we also performed analyses for non-disturbance calls. The data were insufficient, leading to log Bayes factors between -0.5 and 0.5 in all cases (see Table S2), which are thus considered “barely worth mentioning” (15). In addition, in the case of experiment 2, the algorithm could not converge.

Table 1 – Log Bayes Factor comparing various models across the two categories: Alarm calls, Food calls and Non-Disturbance calls. A positive log Bayes factor (in shades of red) favours the first model (in line). Following Jeffreys (15), the evidence is considered as "substantial" if the absolute value of the log Bayes factor is above 0.5, as "strong" if it is above 1, and as "decisive" if it is above 2. (The converse applies to negative log Bayes factor, since logBF(M1, M2) = - logBF(M2, M1)).

Alarm	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	0	0.48	-2.65	-2.40	2.06
Model 2	-0.48	0	-3.13	-2.98	1.58
Model 3	2.65	3.13	0	0.25	4.46
Model 4	2.40	2.98	-0.25	0	4.71
Model 5	-2.06	-1.58	-4.46	-4.71	0

Food	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	0	0.07	-0.78	-0.56	0.80
Model 2	-0.07	0	-0.85	-0.63	0.73
Model 3	0.78	0.85	0	0.22	1.59
Model 4	0.56	0.63	-0.22	0	1.36
Model 5	-0.80	-0.73	-1.59	-1.36	0

Non-Disturb.	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	0	-	0.39	0.35	-0.10
Model 2	-	0	-	-	-
Model 3	-0.39	-	0	-0.03	-0.49
Model 4	-0.35	-	0.03	0	-0.45
Model 5	0.10	-	0.49	0.45	0

REFERENCES

14. S. Höhna, *et al.*, RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Syst. Biol.* **65**, 726–736 (2016).

15. H. Jeffreys, *The Theory of Probability* (OUP Oxford, 1998).

16. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).

17. W. Xie, P. O. Lewis, Y. Fan, L. Kuo, M.-H. Chen, Improving Marginal Likelihood Estimation for Bayesian Phylogenetic Model Selection. *Syst. Biol.* **60**, 150–160 (2011).

REFERENCES OF THE DATABASE

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (Theropithecus gelada): acoustic structure and behavioural context. Folia primatologica, 55(3-4), 109-132.

Bezerra, B. M., & Souto, A. (2008). Structure and usage of the vocal repertoire of Callithrix jacchus. International Journal of Primatology, 29(3), 671.

Bouchet, H., Blois−Heulin, C., & Lemasson, A. (2012). Age−and sex−specific patterns of vocal behavior in De Brazza's monkeys (Cercopithecus neglectus). American Journal of Primatology, 74(1), 12-28.

Casamitjana, J. (2002). The vocal repertoire of the woolly monkey Lagothrix lagothricha. Bioacoustics, 13(1), 1-19.

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. Behaviour, 142(4), 397–421.

Fuller, J. L. (2014). The vocal repertoire of adult male blue monkeys (Cercopithecus mitis stulmanni): a quantitative analysis of acoustic structure. American journal of primatology, 76(3), 203-216.

Gros-Louis, J. J., Perry, S. E., Fichtel, C., Wikberg, E., Gilkenson, H., Wofsy, S., & Fuentes, A. (2008). Vocal repertoire of Cebus capucinus: acoustic structure, context, and usage. International Journal of Primatology, 29(3), 641-670.

Hammerschmidt, K., & Fischer, J. (1998). The vocal repertoire of Barbary macaques: a quantitative analysis of a graded signal system. Ethology, 104(3), 203-216.

Jeffreys, H. (1998). *The Theory of Probability*. OUP Oxford.

Kudo, H. (1987). The study of vocal communication of wild mandrills in Cameroon in relation to their social structure. Primates, 28(3), 289-308.

Macedonia, J. M. (1993). The vocal repertoire of the ringtailed lemur (Lemur catta). Folia primatologica, 61(4), 186-217.

Range, F., & Fischer, J. (2004). Vocal repertoire of sooty mangabeys (Cercocebus torquatus atys) in the Taï National Park. Ethology, 110(4), 301-321.

Rowell, T. E., & Hinde, R. A. (1962, March). Vocal communication by the Rhesus Monkey (Macaca Mulatta). In Proceedings of the Zoological Society of London (Vol. 138, No. 2, pp. 279-294). Oxford, UK: Blackwell Publishing Ltd.

Salmi, R., Hammerschmidt, K., & Doran−Sheehy, D. M. (2013). Western gorilla vocal repertoire and contextual use of vocalizations. Ethology, 119(10), 831-847.

de Waal, F. B. (1988). The communicative repertoire of captive bonobos (Pan paniscus), compared to that of chimpanzees. Behaviour, 106(3), 183-251.