



**HAL**  
open science

# Chimpanzees Have no Expanded Formant Space and No Correlated F1 and F2: A Comment on Grawunder et al. (2021)

Louis-Jean Boë, Thomas R Sawallis, Pierre Badin, Jean-Luc Schwartz

## ► To cite this version:

Louis-Jean Boë, Thomas R Sawallis, Pierre Badin, Jean-Luc Schwartz. Chimpanzees Have no Expanded Formant Space and No Correlated F1 and F2: A Comment on Grawunder et al. (2021). *International Journal of Primatology*, 2023, 44 (December), pp.1046-1049. 10.1007/s10764-023-00396-1 . hal-04239700

**HAL Id: hal-04239700**

**<https://hal.science/hal-04239700>**

Submitted on 14 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Chimpanzees have no expanded formant space and no correlated F1 and F2: A comment on Grawunder *et al.* (2021)**

International Journal of Primatology – <https://doi.org/10.1007/s10764-023-00396-1>

**Louis-Jean Boë<sup>1</sup>, Thomas R. Sawallis<sup>2,\*</sup>, Pierre Badin,<sup>1</sup> Jean-Luc Schwartz<sup>1,\*</sup>**

<sup>1</sup>Université Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France.

<sup>2</sup>New College, The University of Alabama, Tuscaloosa, AL, USA.

\* Jean-Luc Schwartz - jean-luc.schwartz@gipsa-lab.grenoble-inp.fr

\* Thomas R. Sawallis - tom.sawallis@gmail.com

We enthusiastically support investigations of human speech emergence through vocalizations of living primates, as in Grawunder *et al.* (2021), but we seriously doubt their supplement's formant data, which we believe reveals substantial errors. These errors invalidate the surprising main claims in the paper: F1-F2 correlation and formant space expansion through extreme F1s and F2s.

The latter claim drew our attention because previous research shows (e.g., Boë *et al.* 2017, 2019) that such expansion is impossible. Basically, for any variable tube of a fixed length closed at one end (like a vocal tract (VT) closed at the larynx), there is an F1-F2 “Maximal Acoustic Space” (MAS), of the grossly triangular shape familiar from over a century of research on human speech and phonetics. For any configuration of such a tube, the F1-F2 pairs necessarily lie within the MAS, regardless of number, length, or area of its component tubes and of the resulting form of the overall tube. Long-tailed macaques (*Macaca fascicularis*) (Fitch *et al.*, 2016), and Guinea baboons (*Papio papio*) (Boë *et al.*, 2017) conform to that restriction, while the position and size of their MAS within the F1-F2 space depends naturally on each species' VT length (VTL). The areas covered within the MAS depend on the species' control of VT shape during vocalization. Several other primate species have been shown to conform as well, by using a species' typical VTL to normalize published data to a common MAS (Boë *et al.*, 2019).

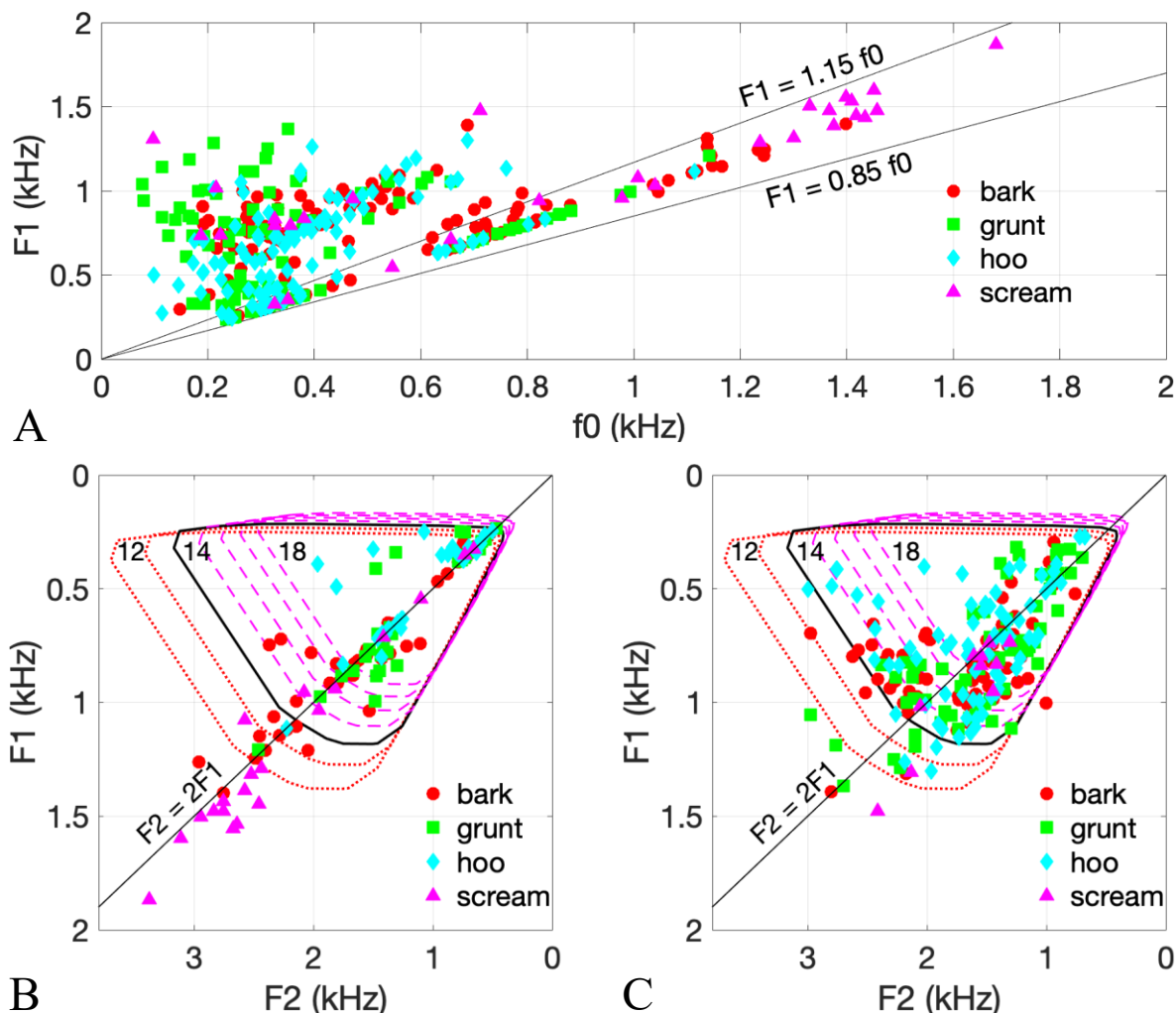
We know of no published measurements of chimpanzee VTL. Nishimura (2005) published MRIs of chimpanzee (*Pan troglodytes*) subjects 8 and 12 with a scale, and we measured their VTLs as 12.3 and 14.5 cm respectively, somewhat less than the sum of the measured horizontal and vertical portions of their supraglottal VTs in Nishimura's Table 3. From Table 3 data for subjects 11-15, comparable to subjects in Grawunder *et al.*, we estimate sub-adult and adult chimpanzee VTLs range approximately from 12 to 18 cm. Moreover, during vocalization, VTL can vary due to either larynx or lip movements, so Fig. 1 shows MASs from 12 to 18 cm in 1 cm steps.

Linear predictive coding (LPC), when used for formant detection, is prone to various biases and errors which worsen as fundamental frequency ( $f_0$ ) rises and the harmonics (multiples of  $f_0$ ) spread wider apart. Formant values delivered by LPC therefore require expert critical examination to avoid confusing  $f_0$  and its harmonics with formants.

Fig. 1 shows exactly that confusion. Panel A shows the  $f_0$  - F1 relationship for the 361 utterances in the Grawunder *et al.* supplement for the four main call types with all values of  $f_0$ , F1 and F2 available (corresponding to Fig. 3a in Grawunder *et al.*). Using a closeness criterion of  $0.85 \leq F1/f_0 \leq 1.15$  (i.e.,  $\pm 15\%$ ), 136 utterances (38%) have  $F1 \approx f_0$ , including, strikingly, all utterances where  $f_0 > 800$  Hz.

**ACCEPTED MANUSCRIPT / CLEAN COPY**

(Implausibly, 93 of the 136, 26% of the total, have *identical* values for  $f_0$  and  $F_1$ , twice down to the 11<sup>th</sup> decimal!) Further analysis shows that of those 136 vocalizations, 110 (81%, or 30% of the total) also have  $F_2 \approx 2f_0$ . Since laryngeal frequency and VT resonances are anatomically and acoustically independent, having  $f_0 \approx F_1$  is technically possible, but the high proportion points instead to measurement confusion between harmonics and formants. Further confusions are possible, particularly for  $f_0 > 400$  Hz where, beyond  $F_1 \approx f_0$ , most other  $F_1$  values are near the second harmonic,  $2f_0$ .



**Fig. 1.** Documenting the formant/harmonics confusion. A shows the  $F_1$ - $f_0$  relationship for the Grawunder et al. chimpanzee utterances and our  $\pm 15\%$  operational boundaries for  $F_1 \approx f_0$ . B and C show the same utterances against Maximal Acoustic Spaces (MASs) for vocal tract lengths (VTLs) from 12 to 18 cm in  $F_1$ - $F_2$  space with the axes' origins to the upper right, following convention in speech and phonetics. B and C also show the  $F_2 = 2F_1$  line to help judge possible formant correlation. B shows the  $F_1$ - $F_2$  data for utterances where  $F_1 \approx f_0$  while C shows the remaining  $F_1$ - $F_2$  data where  $F_1 \neq f_0$ .

The utterances in A are then sorted into one of the two lower panels. B shows the data where  $F_1 \approx f_0$ , while C shows the remaining data where  $F_1 \neq f_0$ . B includes most of the points that are beyond any plausible MAS. A MAS encompassing those points would require a VTL of 8 cm or less, which is unrealistic for adult or sub-adult chimpanzees. Removing these points keeps proportionally more data inside the MASs and reduces any  $F_1 - F_2$  correlation, as shown in C. Most points there fit into the MAS

with a 14 cm VTL, in accordance with our measurement of Nishimura's subject 12. MAS size is inversely proportional to VTL, so utterances with protruded lips would lower formants and shift data points up and right into a smaller MAS. Some of the remaining points outside the 14 cm MAS (e.g., 4 barks and 2 screams) may come from small subjects or shortened VTLs, others may stem from different kinds of errors, but not from the F1-F2 correlation.

The two major claims in Grawunder et al. (2021) are: (1) expansion of F1-F2 data outside the vowel triangle of human speech – i.e. outside the MAS (“There is also indication that non-human primates may use vowel space *not* used by humans, particularly when both F1 and F2 are high.”, p. 9) and (2) F1-F2 correlation (“Comparing our results with published primate data, humans show less F1-F2 correlation”, Abstract). We believe both conclusions have the same cause, a misidentification of the fundamental and its harmonics as formants. This results in substantial errors in the published formant measurements (probably about 38%), leading in turn to mistaken findings.

These erroneous measurements contradict long established understandings about formants. They probably also contribute to a further problem here. The authors ascribe formant values to call categories which are acoustically incompatible with the facial configurations their Fig. 1 claims as characteristic of those categories: open-jawed barks cannot make the claimed low formant values, nor can their respective close jaws and protruded lips produce the high formant values that keep 7 grunts and 5 hoos outside the 14 cm MAS.

Grawunder and colleagues are stewards of a large, valuable multimodal corpus that is interesting and important for science. Unfortunately, erroneous formant data due to confusions between formants and f0 harmonics lead to unwarranted conclusions about vocal expansion and inter-formant correlation. Pointing out the erroneous data and misguided ideas about chimpanzee vocalizations should help avoid misdirected research in primate communication.

### Acknowledgements

The authors are grateful to the editor and also for very helpful comments by two anonymous reviewers. Louis-Jean Boë thanks Marion Laporte and Guillaume Barbier for fruitful discussion on this

### References

- Boë L-J, Berthommier F, Legou T, Captier G, Kemp C, Sawallis TR, Becker Y, Rey A, Fagot J. 2017 Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS ONE* 12, e0169321. (doi:10.1371/journal.pone.0169321) (Ref. 7 in Grawunder et al., 2021).
- Boë L-J, Sawallis TR, Fagot J, Badin P, Barbier G, Captier G, Ménard L, Heim J-L, Schwartz J-L. 2019 Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Sci. Adv.* 5, eaaw3916. (doi:10.1126/sciadv.aaw3916) (Ref. 4 in Grawunder et al., 2021).
- Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016 Monkey vocal tracts are speech-ready. *Sci. Adv.* 2, e1600723. (doi:10.1126/sciadv.1600723) (Ref. 46 in Grawunder et al., 2021).
- Grawunder, S., Uomini, N., Samuni, L., Bortolato, T., Girard-Buttoz, C., Wittig, R. M., & Crockford, C. (2021). Chimpanzee vowel-like sounds and voice quality suggest formant space expansion through the hominoid lineage. *Philosophical Transactions of the Royal Society B*, 377(1841), 20200455
- Nishimura, T. (2005). Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *American Journal of Physical Anthropology*, 126(2), 193–204. <https://doi.org/10.1002/ajpa.20112>