

## Editor

Presently, we have the review results from the same reviewers.

I recommend you to revise the manuscripts following these suggestions. I am waiting for a second revision.

RE: We thank the reviewers for their helpful comments. We believe that the current changes have significantly improved the manuscript and that you will find the article suitable for publication. Detailed responses to each comment are included below, with appropriate changes made to the manuscript.

## Reviewer: Sharon Glaeser

### **Basic reporting**

The basic reporting is significantly improved and meets all of the criteria in reviewer guidelines. The expanded coverage on acoustic communication in general and specifically on African elephant rumbles is a significant improvement. The details added on methods eliminates previous concerns. (I also appreciate that my editorial suggestions were incorporated.)

### **Experimental design**

The experimental design meets all of the criteria in reviewer guidelines. I had only one previous concern on the acoustic parameters measured, which has been addressed.

### **Validity of the findings**

The validity of findings meets all of the criteria in reviewer guidelines. The detail added addressed all previous concerns.

### **Comments for the author**

Remarkable improvement. I have only a few editorial suggestions plus citations to add for hormonal correlates of musth. See attached file.

RE: Thank you for your comments, we are delighted to hear that you are satisfied with the changes we made to the manuscript. Thank you for the comments in the pdf – we have incorporated these suggestions into the manuscript.

## Reviewer 1

### **Basic reporting**

Review:

Thank you for addressing my comments: however, I have to correct you, formant frequencies do also

contribute to individual distinctiveness (they are not only coding size) in many mammals, including elephants. Why? This is easily explained: the supralaryngeal vocal tract is distinctive from individual to individual. Therefore, leaving out this parameter cannot be really explained by your argument. Formant values in Stoeger et al. had the PIC (potential for individual coding) above 1, and thus contributed to the result.

Second: of course it is difficult to collect vocalizations in the wild, but if you do not have the sample size for a certain study (or investigation), than this is to be accepted. You address this now with changing statistics, still, sample size is so low for investigating individual differences. I would request the authors to tone down their arguments /discussion / results. So for example as mentioned below- don't state that you show an individual signature, for example. This is a very strong term.

In addition I still have some concerns about the acoustic measurements, that need further explanation.

RE: The reviewer makes a good point. We have changed the terms and reworded the discussion to account for this. We also now include a limitations section in which we highlight these issues (L260-275). Please see below for an explanation about the acoustic measurements.

Abstract: vocal signature is a very strong term. We know that from signature whistles from dolphins, for example, where we can clearly see individual differences in the spectrograms. Could you stick to the term "individual differences". To me, the paper does not significantly enough show a signature. (for this, more acoustic parameter should have been measured) and a true signature should be visible. Although you argue different, you did not measure all relevant acoustic parameter. Stoeger Baotic 2016 do not state that formants are not relevant for individual differences, in fact they are as well.

RE: Agreed. 'Individual signature' was now rephrased to 'individual differences'.

And there are some referencing that I believe should be adjusted:

Line 71: references: these are not the proper references that should be used. The author of the chapter about vocalizations in the Moss et al book is Joyce Poole. So you should cite her chapter instead of simply citing the entire book.

RE: Done

Line 89: again, please cite the chapter within the book that is addressing male societies.

RE: Done

Introduction: you gave some arguments of why you believe that it is important to investigate whether individuality is persistent over time in the rebuttal letter. But not so in the paper, thus the rationale for

your study is still not really addressed yet. Can you please add that into the introduction, because you not only need to convince me, but ultimately also the reader of your paper.

RE: Done

### **Experimental design**

Methodology: acoustic analysis. I still do not understand why you omit the common acoustic parameter. Your arguments do not convince me, and are partly wrong.

I repeat my comment from the first round:

“Why not analyzing source and filter parameter and the ones you think are more important, and then you can statistically analyze which parameters are more relevant and contribute more to individual discrimination. This would be the proper scientific approach. Not simply omitting parameter without any comprehensive reason. “

Formant frequencies do contribute to individual coding in many mammals, including elephants. I tend to assume that maybe it was not possible due to quality reasons to measure some of these parameter? This can happen if you record in the wild, at maybe greater distance to the elephants. If this is the case, simple mention it, and tone down conclusions.

If you decide not to add those parameter, tone down your conclusions and mention that it would be important in the future to measure the other parameter as well in order to reveal an acoustic signature. (but not state that you now revealed an acoustic signature).

RE: We have attempted to conduct formant analyses, however, due to the nature of collecting data from wild animals (distance from animals, as well as background noise), the quality of all recordings were not sufficient to extract this information. Without information on formants from all of our recordings (due to an already limited sample size), we are not able to conduct robust statistical analyses. Consequently, following the reviewer's comment, we toned down the discussion and added the suggested information. Furthermore, we added a paragraph stating the limitations of our study (L260-275), highlighting these constraints.

### **Validity of the findings**

Ok in principle, but tone down conclusions.

RE: Done

### **Comments for the author**

Please tone down conclusions; due to low sample size, missing acoustic analysis.

RE: Thank you for these insightful comments. All of the suggestions have now been incorporated into the manuscript.

## Reviewer 3

### Basic reporting

- The authors tell us that filter (formant) frequencies are associated with the length of the vocal tract and therefore correlated to the maturity and age of the bull elephants (Stoeger & Baotic, 2016). They also go on to say that McComb et al. 2003 mentions that the ability of formant frequencies to carry individual identity over long distances is likely to be severely reduced as was confirmed by rerecording measurements but the recordings taken of the elephants in this manuscript do not take into account long distance communication as some of the rumbles were produced in social situations and during feeding – we have no idea if they're being produced in the context of long distance communication. Therefore, it doesn't seem to make sense to just completely remove it from the analysis and not compare the results with it and without it. Just because it was found to mostly correlate with maturity, shouldn't it be incorporated with the wild bull population to rule it out? Because one of the big differences between this manuscript and the Stoeger paper is that it was conducted on wild elephants, shouldn't a formant analysis be included? Wouldn't this be a great way to show that formants are not that important for individual vocal identity?

- o The authors mentioned the difficulty of collecting data from dangerous animals but is it not possible to analyze formant frequencies from the data already collected? If Raven doesn't have this feature, Praat does. It seems that this would be an important component to include since the only distinction these authors make from Stoeger & Baotic is the fact that their data are collected on wild adult males. And yet they don't include one of the measures that the other authors do in order to make the study equally as robust.

RE: The reviewer makes a good point. We have attempted to conduct formant analyses, however, due to the nature of collecting data from wild animals (distance from animals, as well as background noise), the quality of all recordings was not sufficient to extract this information. Without information on formants from all of our recordings, we would not be able to conduct robust statistical analyses. We realise that these are limitations of our study (L260-275) and following the suggestions of one of the other reviewers have now discussed this in the text as well as toned down the conclusions made based on our results.

### Experimental design

- The table reporting the significance between individuals was removed. The authors say it does not provide additional information. Yes, the PERMANOVA tests for differences between all groups at once but like they said, to provide additional information about the intricacies of the differences

between dyads, it was revealed that most are significantly different but some are not. This could be interesting for future studies, especially since there could be a reason that some individuals vocalizations are more similar than others (perhaps relatedness or some other reason?). Perhaps this could be in the supplemental materials?

RE: Agreed. We included the information about pairwise comparisons into the methods section and the discussion. Because of this we also put the table back into the main manuscript. We do suspect that relatedness or rate of social associations may be a factor influencing this and it is something that we are planning on looking into in the future. We added this possibility in the discussion (L213-221). We also included a limitations section in the discussion (L260-275) where we discuss the possibility of the dissimilarity differences possibly being an artefact of limited sample sizes or acoustic parameters measured.

## Results Section

Table 2: Please confirm that these numbers are correct for the 95%, 5% and Center frequencies. The 5% frequencies are very high.

RE: Thank you for spotting this error. Indeed, the data was mislabelled. This is now corrected.

## Validity of the findings

### Discussion Section

- Lines 229-239: The authors discussed the possibilities of increased hormone levels of the bulls in “pre” or “post” musth, therefore making their results even more robust (because testosterone has been found to influence male vocalizations). I recommend removing this part of the discussion because they do not have any data on this and it seems unnecessary to mention.

RE: Agreed, this fragment was removed.

- Perhaps they should discuss the limitations of their data in the discussion section so it's clear to readers that it was challenging to collect all of the parameters used in previous studies.

RE: Thank you for this suggestion. We added a paragraph stating the limitations of our study (L260-275), highlighting the constraints of the study identified by reviewers in their comments.

## Comments for the author

### Grammar

- Line 106, sentence starting with “This will allow for a better...” needs to be rephrased

RE: Done

- Lines 173-174, the authors wrote “limit recordings to a specific behavioral or social contexts” so contexts should be singular.

RE: Done

- Line 186, “used” should be “found”?

RE: Done

- Line 158 – says “McComb et al., 2001” and is not cited in the reference section. Should this be 2003?

RE: Yes, done.

- Line 166 & 168 – Charif and Sharif – spelled two different ways; if it is Charif, it is not in alphabetical order in the references section

RE: We corrected both the in-text citations and the reference.