David Jansen <david.awam.jansen@gmail.com>

MS: 3399353257711615 - Temporal segregation of individual and behavioral signaturesin banded mongoose close calls

Kester Jarvis <bmcbiologyeditorial@biomedcentral.com>   Wed, Sep 5, 2012 at 4:45 AM
Reply-To: Kester Jarvis <bmcbiologyeditorial@biomedcentral.com>
To: Mr David Jansen <david.awam.jansen@gmail.com>

MS: 3399353257711615
Temporal segregation of individual and behavioral signaturesin banded mongoose close calls
David A.W.A.M. Jansen, Michael A. Cant and Marta B. Manser

Dear Mr. Jansen,

Thank you for your submission to BMC Biology, which has now been seen by three referees, plus two more whom we asked to adjudicate on the apparent novelty of your work after seeing the content of the first three reports. You will find all of these reports at the foot of this email, and will see that the referees, including those we asked to adjudicate, are generally supportive of publication. However, you will also see that all the referees feel that the setting of your results into the context of other research in the field could be improved, and that in particular it is important to make clear exactly what is new and interesting about the calls you report.

We would therefore be willing to publish your paper in BMC Biology, provided that you are able to fully address the comments of the referees. In particular, it will be important to ensure that all of the relevant literature is cited and that the novelty of the results reported in your paper is clearly explained relative to those which precede it: do remember that BMC Biology is aimed at a general biological audience and that it is important to set your article into the relevant context, which may not be immediately familiar to readers as it would in a more specialist journal. In this spirit, we would also ask you to present your data in the manuscript more explicitly. We think referee 5?s suggestion to include more spectrograms in figure 1 is a good one, and it would be useful to include an explicit illustration or description of the differences between the different call types, if this is possible. You should also pay attention to our data deposition policy (http://www.biomedcentral.com/about/supportingdata) and make all raw data publicly available as far as possible.

We will look forward to receiving your revised manuscript by 26th September, and you should include a cover letter detailing your point-by-point responses to the referees? comments. It is also useful if you can mark any changes on the manuscript itself, which allows a more rapid editorial assessment. Do let us know if you will need any more time, and we can adjust our expectations for the delivery date accordingly.

Please get in touch if you?ve any questions or concerns.

Kind regards,
Kester

Kester Jarvis, PhD
Senior Assistant Editor
BMC Biology
BioMed Central
Floor 6, 236 Gray's Inn Road
London, WC1X 8HB

Tel: +44(0)20 3192 2307
Email: kester.jarvis@biomedcentral.com, BMCBiologyEditorial@biomedcentral.com
Website: http://www.biomedcentral.com

REFEREES? REPORTS

REFEREE 1
Reviewer's report
Background
?Vocal signatures? needs to be defined. Do you mean ?calls with individually distinctive acoustic properties?? If so, why

not say this and then say (for simplicity, ?vocal signatures?).

Methodology/Principle findings
?The first evidence?? This is an exaggeration. What about the grunts given by vervet monkeys that differ depending on whether the animal is approaching a dominant, a subordinate, moving into an open area, or seeing another group? The issue of variation along two dimensions addressed in this study by first showing that one individual varied specific acoustic features from one context to the next and then showing that, despite inter-individual differences, other individuals varied the same acoustic measures.
And what about the grunts of baboons (Owren?s papers) that are both individually distinctive and acoustically different depending upon whether the vocalizer is on a move or approaching a female with infant? There are many such examples of acoustic subtypes. This section needs to be rewritten.

Conclusions/Significance
What do you mean by ?temporal segregation?? Calls given at different times? In different contexts? Once again, this is by no means ?the first evidence? of its kind. This is extremely over-stated! ?We argue that?? but this idea has been in the literature for decades, ever since the call subtypes in vervet monkey grunts were first shown to exist through playback experiments. Of course you can argue that call subtypes allow a signaler with a limited vocal repertoire to overcome such limits and send many messages ? this is obvious ? but the important question is: why so few subtypes? Why not more?

Text
?Animals can encode?? It?s not at all clear what you are claiming here, in large part because acoustic signatures is not defined. Is your point that acoustic signatures are rare? If so, this is simply wrong: they?ve been known for a long time, particularly in primates. Is it that acoustic signatures based on temporal features are rare? If so, this is also incorrect because variation in temporal features help define the subtypes of vervet grunts, and probably the context-specific subtypes of many other species.
?However because acoustic space is limited and many acoustic parameters are correlated with one another, the amount of variation that can be used by signalers to encode different signature types is ultimately constrained.? Why? It?s not in human speech, where speakers use a very narrow bandwidth and acoustic parameters are highly correlated.
??there currently exist no studies that refer to it within a single call type.?
This, again, is simply wrong, both in the empirical details and in the claim that the ?trade-off? argument has never been proposed before. In fact, Rendall & Owren have worked for years on the problem of how individual signalers simultaneously encode caller identity and context-specific variation in their calls. This is what has led them to argue that, in primates at least, individual identity is encoded primarily in formant structure whereas context-specific information is encoded in other acoustic parameters. They support this argument with extensive data on baboon grunts.
Further, when Green presented his original data on Japanese macaque coos, he was criticized in a subsequent paper by Lillehei & Snowdon, who emphasized the importance of controlling for individual variation and looking simultaneously at the encoding of caller ID and context.
Finally, it?s not at all clear why you are making this claim without mentioning Manser?s work on meerkats alarm calls, where two types of information are simultaneously encoded in a single vocalization.
The idea that the data presented here are the ?first of their kind? is simply wrong. It greatly distracts from the paper.

REFEREE 2
Reviewer's report
Review of ms 3399353257711615 ?Temporal segregration of individual and behavioural signatures in banded mongoose close calls?

The authors recorded the close calls of banded mongooses and then used acoustic and statistical analyses to show that they contain temporally separated signatures for individual identity and behavioural context. This is quite remarkable, given that these calls are very short (less than 150 ms), quiet, and structurally simple. The study provides one of the few examples of a simple acoustic signal communicating mulitple signatures and will thus appeal to a broad audience, including anyone interested in complex communication and social behaviour. The data and analyses are solid and fully support the paper?s conclusions. The manuscript is very well written and I really enjoyed reading it. I have a few comments to help improve the paper.

Important revisions that are not essential to the validity of the paper

1. The authors emphasize the paper?s novelty by repeatedly stating that it is the first example of an animal encoding multiple signatures into different elements of a single type of signal. Yet, they provide examples of two other well-studied systems that also encode multiple signatures into a single signal. White-crowned sparrows encode individual identity and group membership into temporally separated elements of a single song (this paper specifically discusses Marler?s segregation of information hypothesis), whereas túngara frogs encode species identity and individual quality into contiguous, and often partially overlapping, elements of a single advertisement call. Another example is provided by meadow pipits, which encode individual identity into the beginning of their song and species identity into the end of their song (Elfström 1990 Anim Behav 40:786-788). The authors should state clearly how their findings are different from those

of these other studies. Is it because this is the first example in a mammal, because the elements containing the signatures are contiguous (i.e., not temporally separated, as they are in bird song), or because the other examples aren?t considered a ?single call type?? The whine and chuck series of the túngara frog?s advertisement call would certainly qualify as a single signal, and most ornithologists would also consider the song of the white-crowned sparrow and the meadow pipit to be a single signal, despite having multiple temporally separated elements, such as trills and note complexes. This does not diminish the value of the current study, but the differences between it and these other studies should be stated more clearly.

2. I applaud the authors for using a permuted discriminant function analysis (pDFA) that properly accounts for non-independence among multiple calls from the same individuals. This is very rare in animal behaviour research, and is one of the greatest strengths of the current paper. Although I have no concerns about the general statistical approach, I do have some suggestions to help clarify the description of the statistical methods. First, the purpose of the linear mixed model is a bit unclear. Was it performed solely for the purpose of calculating variance inflation factors and obtaining a subset of acoustic parameters that was free from multicollinearity for use in the pDFA, or was it used in addition to the pDFA to compare calls among behavioural contexts? Were the acoustic variables used as covariates in a single linear mixed model (if so, what was the dependent variable), or were they used as dependent variables in separate models? Second, it is unclear exactly how many pDFAs were conducted. I understand that the whole call, the noisy base of the call, and the harmonic part of the call were analyzed separately. But, for each of these call components, did the authors conduct a single pDFA that included multiple factors for individual, behavioural context, group, and sex (i.e. a total of 3 pDFAs), or were separate pDFAs conducted for each group and sex? The fact that Table 1 presents separate classifications for each group suggests that separate pDFAs were conducted for each group and for each sex (i.e. 3 call components x 2 sexes x 4 groups = 24 pDFAs). If this is the case, do the results presented in the results section correspond to males or to females. One or two sentences near the end of the statistical analysis section would really help to clarify how many analyses were used, and which factors were used in each analysis.

3. At present, there is no way to assess which acoustic variables were included in the pDFAs following the test for multicollinearity and the stepwise selection procedure, or which of the variables in the final models had the greatest effect on the disciminibility of individuals and behavioural contexts. If only 3 pDFAs were conducted (see comment 2 above), then the authors could include a small table (perhaps as an ?Additional Material Files?) that shows the final discriminant functions and their corresponding loadings. If more pDFAs were conducted, then perhaps the authors could include a sentence or two that states, in general terms, which variables were used in the final analyses and which were the most important for discrimination. This is also important because an assumption of the pDFA is that the number of variables included must be smaller than the number of individuals included in the smallest class. Groups 11 and 15 had only 7 individuals, so the corresponding analyses should have included no more than 6 predictor variables.

Suggested revisions that are not essential to the validity of the paper

4. The authors should mention that future research should use playback experiments to test if the behavioural context signature is important to receivers. This could go in the first paragraph of the discussion following the sentence that describes receiver responses to individual signatures.

5. The acoustic analysis methods suggest that there are 19 acoustic variables (16 from the automatic measurement procedure, 2 derived variables, and the number of pulses), yet Table 3 only shows 14 variables. Also, some of the acoustic variables are not described in sufficient detail to be replicated. ?Bandwidth?, ?Maximum Frequency?, and ?Minimum Frequency? need to be defined relative to the call?s ?peak amplitude? (this value is available directly from Avisoft?s automatic measurements dialogue box). For example, maximum frequency could be defined as the highest frequency that was within 10 dB of the peak amplitude. Otherwise, minimum frequency would always be 0 Hz and maximum frequency and bandwidth would always be the highest frequency of the recording (probably 22050 Hz). Also, the frequency variables listed in Table 3 should have ?Hz? listed in parentheses as their unit of measurement. Finally, the authors should explain where within the call the frequency measurements were obtained. Table 3 states ?mean? or ?max? in parentheses, but these terms are unclear and will not be meaningful to anyone who is unfamiliar with Avisoft software. Does ?mean? indicate that measures were obtained from an averaged spectrum, or that they were obtained from every spectrum within the call and then averaged. These two methods produce very different results, so the method should be defined clearly in the text. For example, the authors could state ?all frequency measures were obtained from the mean spectrum of each call or call component, and the 3 quartiles were also measured from the point within the call or call component that had the maximum amplitude.?

6. Table 4 should be re-labelled as ?Table 1? since it is mentioned first in the text.

7. The footnotes in Tables 1 and 2 state that results are included for ?the whole call and the noisy part? of the call, but should also state that results are included for the ?harmonic part? of the call.

8. In the recording methods, please state the file format, number of bits, and sampling rate used on the two recorders.

9. I have a few suggested wording changes and typo corrections. In the last line of the Background in the Abstract, ?

emitted close graded close calls? should be ?emitted graded close calls.? In the last sentence of the Introduction, delete ? in their close calls? and change ?aspects of a graded call? to ?aspects of this graded call.? In the results, change ? considerable? to ?considerably? and change ?neither? and ?nor? to ?either? and ?or? (or change ?No evidence? to ? Evidence?). In the Discussion, ?tungara? should have an accent above the ?ú?. In the fourth paragraph of the Discussion, change ?Similar? to ?Similarly? and change ?unambiguity? to ?clarity. In the last sentence of the Discussion, change ?avoid the lack of ambiguity? to ?avoid the ambiguity.? In the Conclusion, change ?provide equally? to ?provide an equally? and change ?in animals, and in humans? to ?in human and non-human animals.? In the first line of the Recording Methods, should it be ?more than? instead of ?less then?? In reference 48, delete the second ?d? from ? handbookd?. In Table 1 footnote, change ?change? to ?chance.? Throughout the Methods, change the 3 instances of ? analyzes? to ?analyses? since this in the noun form of the word.

REFEREE 3
Reviewer's report
This is a well written and interesting paper showing how individuality and behavioural context are encoded in banded mongoose calls. The authors found that individuality and context are encoded in different parts of the calls, therefore providing a nice example of temporal segregation of information. I have some suggestions to improve the clarity of the manuscript.

General comment. The term ?signature? is typically used to describe information encoded in vocalisations about the identity of an individual (e.g. species, group, individual), similarly as in humans (?A distinctive mark, characteristic, or sound indicating identity?). It does not sound right to me to use it in order to describe behavioural context, particularly because the context in which an individual calls is not fixed over time. Therefore, I suggest changing ?signature? to ? information? or ?cues?, or else, when referring to behaviour/context throughout the ms. Note that the literature cited in the background P3, L5-7 did not use the term ?signature? to describe vocal indicators of male quality or reproductive state.

Specific comments.

P3-L21. I would emphasize here that for individual recognition, we need variation between individuals (and stereotypy within individual), not just ?variation?.

P3-L24. I would mention here that segregation of information can be done by using either different features (temporal segregation) or different ranges of the same feature (e.g. frequencies) to convey different information (Marler 1960).

P3-L26. Another good bird example of temporal segregation is the Ortolan bunting, in which the initial song part codes for the individual identity and varies between males and the final part codes for the population identity (e.g. Osiejuk et al. 2005, Behavioural Processes 68:69?83).

P5. I suppose that all the ncce provided in the results are percentages, not just numbers? If yes, ?%? could be indicated after each of them.

P5-L8. This ncce must be wrong, it is exactly the same as the one mention above and the p value is significant (contrary as what is written in the text).

P5, L11-19. It would be useful if contexts could be described before these results, in the introduction, or at least, mention Table 4 here.

P5, L16-18. Table 2 only shows general differences between contexts. If such claims are being made here (e.g. duration increases, etc.), then statistical results comparing each acoustic parameter value between contexts should be provided. Such tests are mentioned in the methods (linear mixed effect models), but do not seem to appear anywhere in the results.

P6, L1-3. This is the case if individuals from a given group are more genetically related than individuals from different groups. The other mechanism that can arise when individuals change groups and in species capable of vocal flexibility is group convergence (e.g. Knörnschild et al. In Press. Animal Behaviour; Briefer and McElligott 2012, Animal Behaviour 83, 991-1000).

P6, L5-13. This kind of segregation of information is not rare in songs of birds or even mammals. The difference between ?songs? and ?calls? is quite subjective (at least ?human subjective?). The term ?song? is usually used for vocalisations composed of sequences of elements. This presence of multiple elements facilitates temporal segregation of information. ? Calls? is used for simpler vocalisation composed generally of one single element. If only one simple element is present, temporal segregation of information is not possible. These definitions of ?songs? and ?calls? established by researchers are the only reason why this paper is the first to find segregation of information in a call. See for example hyrax, which produce a complex call termed ?song? by researchers. This song encodes information regarding individual identity, body weight, size and condition, social status and hormonal state in different components (Koren and Greffen 2009, Behav Ecol Sociobiol 63:581?590). Therefore, I suggest revising this claim on the novelty of the paper. This would not make the

paper less interesting than it is, because it still shows a very nice case of segregation of information.

P8-L14. What do you mean by ?less than 1 year??

P9, L2. Maybe mention what is the lowest fundamental frequency measured in banded mongoose here, to justify the filter.

P9-Acoustic analysis. Some parameters like Max. Freq. and Min. Freq., listed in Table 3 are not described in the methods.

P9. Statistical analyses. I did not find where the results of the linear mixed effect model were. It would be useful to present them in a table (e.g. in Table 3).
The part on parameters with VIF ?2.5 included in the analyses or not is not clear. In which analyses where these parameters included? ?the remaining parameters were entered into a DFA?: does that mean that the parameters entered in the DFA were mutlicollinear? Can these parameters included in the DFA be listed somewhere? One of the assumptions of DFA is an absence of perfect multicollinearity, which means that parameters entered in the DFA should not be too redundant. If this is the case, a PCA can be carried out before the DFA. Is that the case here?
How were correct classification probabilities determined for behavioural context while controlling for individual and vice-versa? Was a permuted DFA used (pDFA is mentioned in the methods, but later on)? More generally, it is not very clear how sex, group, behavioural context and individuality were all controlled for in the pDFAs, given that a pDFA can control for only one factor at a time (one test and one control factor). From the results, it seem that groups were analysed separately, which seems ok, but then this should be explained in details in the methods.
P10,L3-4. Which criteria on the basis of results from the linear effect models were used to select parameters?

Table 2. Maybe include here also the random %, similarly as in Table 1. Also, is the ncce for digging-searching on the whole call really only 3.34%? This ncce is listed as p<0.1.

REFEREE 4
Reviewer's report
I agree the novelty of the findings seems to rest on the definition of a call vs a song and the importance of the temporal segregation of the info (as reviewer 1 points out many primate calls are context specific and individually distinctive, but the identity ?signature? in these calls is not a separate temporal element, so personally I don?t think the vervet findings are directly comparable). I think the results need more firmly embedded in the existing literature (where it is more explicit as to why these findings are different or similar to other findings), however I do think this will still be a paper of wide interest to biological researchers. Complexity of communication systems seems to be a topic attracting wide interest across taxa currently, so I think these findings remain interesting and suitable for BMC Biology.

REFEREE 5
Reviewer's report
This interesting paper demonstrates that different types of information are encoded within different temporally-segregated portions of the same short contact call, in banded mongooses. Specifically, the initial noisy part of the call encodes individual and group identity, while the (optional) more harmonic second part encodes the behavioral context, or current activity of the animal.

I think this paper is interesting and worth publishing, but I agree with the first reviewer that its overarching presentation and contextualization need some work. Specifically, the non-specialist reader will be left wondering what, exactly, is new about this finding and why it is important.

There are two problems: the first is the use of the term "signature" which is used in a non-standard and, particularly in this system, non-helpful way. "Signature" is typically used to discuss acoustic cues to individual and/or group identity, and NOT sex, size, quality etc. For the latter "vocal cues to?" or "acoustic indicators of?" would be more standard or more appropriate in this case.

The second is that "temporal segregation" is too vague: this would include sequential combination of separate acoustic entities of any size or continuity. Thus, many complex birdsongs encode individuality in the song, specifically in the sequence of individual song syllables (e.g. nightingales), or the co- and qui- notes of the coqui frog encode male- versus female-directed information. Such encoding has also been shown in cetaceans and primates and is nothing new, even in mammals.

What IS new is that a single short syllable contains two different types of information, but in temporally-segmented fashion. This is the way human consonant-vowel syllables ("ba" vs "pa" /"ba" vs. "bo") work, and this seems a far more appropriate analogy than the syntactic one ("I John eat" versus "I John move") used here. So the key novelty is in the within-syllable encoding aspect of the findings presented here, which might be termed _segmental concatenation_, or "phonological". It is the within-syllable aspect of these findings that is novel and important.

So my suggestion is to reframe the result in these terms:
"Segmental concatenation of individual signatures and context cues within single syllables of banded mongoose contact calls"

"While many animal signaling systems use concatenation of acoustically-separate syllables to enrich and extend the signaling space (e.g. birdsong), human speech also uses a different type of encoding into individual syllables, at a phonological level. Thus a stop consonant like /b/ versus /p/ can be combined with a vowel like /a/ or /o/ to create a richer signalling unit than either class alone could provide. Such combinations (versus "syntactic" concatenation of syllables and words) are a core component of the phonological component of human spoken language"

The authors might also cite that specific segment types commonly signal group (dialect) differences in human speech (e.g. trilled versus uvular /r/ in German dialects, vowel identity in English dialects) which is close to their "group signature" idea.

Of course, the analogy is imperfect since humans don't (typically) have individual specific segments (though an individual lisp or similar abnormalities might reliably signal identity in some cases). But nor do humans walk around saying "I John eat". No analogy is perfect.

While this way of framing the results doesn't do full justice to either human syntax or to phonology, it at least will make clear to most readers what this study is all about.

I think the authors should present spectrograms as Figure 1, with some rather extreme examples to illustrate the phenomenon.

To submit your revised manuscript

When you have revised your manuscript in light of the reviewer's comments and made any required changes to the format of your paper, please upload the revised version by following these instructions:

1. Go to http://www.biomedcentral.com/manuscript/login/man.asp?txt_nav=man&txt_man_id=3399353257711615 and log on with your email address and password.
2. With the 'Manuscript details' tab, please update the title, abstract and author details if they have changed since the previous version. It is very important that all changes are updated on this page, as well as in the manuscript file as the information on this page will be used in PubMed and on BioMed Central if your manuscript is accepted for publication.
3. With the 'Cover letter' tab, please provide a covering letter with a point-by-point description of the changes made.
4. With the 'Upload files' tab, please upload the revised version of the manuscript and press 'Submit new version'. Please wait for the confirmation page to appear - this may take a few moments.