

Responses to Reviewers for Genes et al., A cultural atlas of vocal variation: yellow-naped amazons exhibit contact call dialects throughout their Mesoamerican range

Note: Reviewers' comments are in regular font and our *responses are in blue italics*

Reviewer 1

The article, thanks to research that lasted a long time and in several Central American countries, describes the presence of different vocal dialects in the yellow-naped amazons, using different systems of analysis and interpretation, in particular in the work, 14 different vocal dialects are described, identified in a visual way and an undefined number but less, defined with statistical methods.

The work is well structured and makes an important contribution to the study of the dialects for the species of birds specifically of the Yellow-Naped amazons, but it should be improved in some aspects both of method and representation of the results; the numbers of individuals that have been analyzed for different areas and dialects should be reported (the total numbers in paragraph 3.1), also describing some simple statistical values such as media of songs studied for each individual by area and/or dialect. This information could be introduced not only in the text but also in Figure 4.

R: Thank you for the positive comments about the value of our study. We have attempted to address your concerns as we understood them. Specifically, we added a stand-alone supplemental data table (Supplementary Table 1) to address the concern regarding reporting the numbers of individuals analyzed. The table includes data on numbers of calls, numbers of birds, site, and year of recording.

With regard to the main results, the number of dialects, and their distribution, it would be important to describe strictly the number of dialects identified by statistical methods and to compare and comment on this result by comparing it with the classification obtained by the visual method

R: We believe this is a good suggestion, however it is important to note that the statistical methods do not provide an unequivocal answer on how many dialects are present, but rather provide visualization of call distributions that require interpretation by the observers. To clarify general patterns in these distributions, we added these sentences to section 4.1, "Overall, the SPCC and PCA plots showed more separation in the northern half of the range, like with the Apex and Island call types. The southern portion of the range had more call type overlap, like in Costa Rica."

in the materials and methods is also reported an analysis to define the statistical differences for the values of pc1 pc2 of the different dialects (posthoc analysis), but the results are not reported.

R: Thank you for this suggestion. To address it, we created Supplementary Table 3 that shows the p-values generated from the Tukey post-hoc test for each call type comparison.

In Chapter 2.5, the statistical methods reported many functions for example `leveneTest()` but should be better defined as not clear in writing.

R: We acknowledge this concern, but note that in each case where we use a statistical test we provide a brief description in the methods of the purpose of that test and how it was implemented. For example, “we used the function `leveneTest()` from the car package (Fox and Weisberg, 2019) to assess for equal variances, and the function `shapiro.test()` from the package MASS (Venables and Ripley, 2002) to evaluate the normality of our dataset.” We hope that this is sufficient detail for readers to understand and replicate our analyses if necessary.

In paragraph 3.4, the results of the analysis for main components in terms of the number of components selected, of the associated variables in terms of correlation should be better described (partly described), in this paragraph, the meaningfulness values should be simplified, and divided by classes, without putting decimal numbers (an example is better described as $p \leq 0,001$).

R: We had a hard time understanding some parts of this suggestion but attempted to address it by changing the p-values in paragraph 3.4 to show 4 significant digits (i.e. < 0.001) as suggested by the reviewer. Additionally, we added Supplemental Table 2 illustrating PCA factor loadings for the first 5 components. These additional data will allow readers to understand how acoustic variables are described by each PC variable.

The analysis by main components should be better described in relation to the group studied, and it is not clear especially in relation to Figure 4, if done overall or by study area (nation), better even if trivial to specify it.

R: Thank you for pointing out this point of confusion. Our PCA was performed on all calls combined. Our ANOVA of differences among call types for PCA 1 and PCA1 was also performed across all call types. In order to clearly illustrate differences in neighboring calls, we plotted PCA1 and PCA2 values for call types separately for each country (Figure 4). We have clarified these points in the relevant text in the methods.

In Figure 3 it is not clear what is described in the central box or its meaning

R: The central image is one of the call types from our dataset that, due to its relatively simple structure, could serve as an idea of a common origin call structure. We updated the figure to include the term “proto-dialect” on the central image.

Reviewer 2

This paper reports on a comprehensive data set that includes recordings of contact calls across most of the continuous range of the yellow-naped amazon parrot in Central America (recordings from El Salvador could not be obtained, for obvious reasons). The authors ask whether the dialect structure previously described in Costa Rica is characteristic of these birds across their

entire range. They attempt to validate visual examination of sound spectrograms as a way of defining call dialects with acoustical methods. Although the validation was not entirely successful, the scope of the study will make it interesting for researchers studying vocal learning and dialect structure. The conclusion (which the authors do not state directly), that dialects are not epiphenomena that arise because of regional demographics and habitat but are tied to species-wide characteristics such as vocal learning and social structure, could be strengthened by making some changes in the data presentation (and perhaps the analysis).

R: We thank you for your thorough review. We attempted to address your concerns regarding the direct statement of our conclusion and the data presentation throughout the paper by making various changes which we detail below. Furthermore, we think that the reviewer has nicely stated an aspect of our conclusions which we are now happy to include.

In the analysis, the contact calls were first sorted into types via visual analysis of sound spectrograms. Subsequently, the acoustic features of each call were extracted and used to check to see whether the call types defined by the human visual system were valid in terms similarity/dissimilarity measures and/or categorization by means of PCA analysis. These validating analyses were performed on subsets of the recording sites, grouped by country. This approach raises three questions:

1. The grouping by country (political boundaries defined by humans) seems questionable. Looking at the map in Figure 1, there are three sets of recording sites where none of those sites are separated by more than approximately 50 km (Mexico+Guatemala, Nicaragua+Costa Rica, and the Honduran Bay Islands). It would make much more sense to use these geographically contiguous areas as subsets for the acoustic analysis, rather than grouping sites by country - especially given the proximity of sites in southern Mexico and western Guatemala and those near the Nicaraguan/Costa Rican border (Figure 1).

R: We agree that political borders are not an ideal method of separation for interpreting range-wide vocal dialect patterns in this species, given that they are human-recognized boundaries of limited utility to the parrots. However, we note our initial plots showing calls from across the range on a single plot were indecipherable, and even when we separated calls out by region (North: Mexico/Guatemala, East: Honduras, South: Nicaragua/Costa Rica) we found that there were still too many call types in each plot to be well interpreted. Thus, we elected to break our data down by country to illustrate dialect patterns more effectively on a site-by-site level in our initial submission, and feel it is best to retain this division in the current submission. We did attempt to address the reviewer's concern to some extent, by creating a new Supplemental Figure that shows PCA analysis at the regional level as suggested by Reviewer 2.

2. The analyses of the acoustic characteristics of call dialects often do not validate the human visual sorting of sound spectrograms. The authors acknowledge this point in the discussion and provide a reasonable explanation (that cross-correlation and the acoustic features extracted may not be the best methods for acoustic comparisons of these calls, which have multiple harmonics that cover the entire bandwidth of the calls). However, they may wish to look more closely at measurements that are likely to better discriminate the call types illustrated in Figure 2: perhaps frequency modulation, entropy, spectral continuity, and amplitude modulation?

R: We have included a table of the first 5 PC factor loadings that will allow readers to determine which variables are the strongest contributors to call variation.

3. Given that the acoustic analyses did not validate many of the dialect distinctions that emerged from human visual examination of sound spectrograms, it would be good to include a figure with several examples from two similar call types to demonstrate that they do appear to be consistently distinguishable (Island and Ometepe, for example, and/or Bridge and Jump).

R: This is an excellent suggestion. We have added a Supplementary Figure 2 that shows a comparison of 2 call types which were statistically distinguishable and 2 call types that were not.

The country by country map panels in Figure 4 do not correspond to the recording sites shown in Figure 1. For example, two circles representing call types are shown in Guatemala, some distance from the border with Mexico, but there are recording sites at the border between the countries as well that do not seem to be represented in any of the maps in Figure 4. It would be good to have a comprehensive map similar to that in Figure 1 showing the distribution of call types across the entire range of recording sites. Given that the definition of vocal dialects usually includes the criterion that boundaries between dialects are relatively sharp, it is important to provide a comprehensive map so the reader can see the geographic transitions between call types clearly.

R: We thank the reviewer for bringing this mismatch to our attention. We have updated figure 1 to show only those 47 sites at which calls were recorded (rather than all sites sampled), with circles around the 3 sites in Costa Rica that were dropped from further analysis due to low numbers of birds recorded. Sites in Figure 1 now correspond directly to those in Figure 4

In the Abstract and in the Discussion, the authors propose that the similarity of acoustic characteristics in all of the call types / dialects arise because the call types stem from a common ancestral form. An alternative hypothesis is that the acoustic features of call types are similar because of physiological and anatomical features of the vocal control system that are common across all members of the species and constrain the modulations of the vocal system that produce the sounds. This alternative should be acknowledged.

R: We thank the reviewer for noting this important alternative. We have added a sentence in the abstract acknowledging it that states “Alternatively, similarity in the acoustic features of call types may also be a result of physiological and anatomical features that are common to all members of the species..”

In the Discussion in the third paragraph of section 4.1 we have added “It is also possible that call types share similar acoustic features due to anatomical and physiological characteristics of the vocal control system that are common across all members of the species, and constrain vocal system modulations that produce sounds.”

In the fifth line from the bottom of the Abstract, the authors end a sentence with “supporting our

hypothesis”. However, they have not yet presented a hypothesis. Perhaps add “, that this species has vocal dialects throughout its range”.

R: We implemented this suggestion.

In the discussion (page 10, second paragraph) the authors suggest that bilingual birds have two call types so that they can forage across dialect boundaries or move between groups with different dialects. It is also possible that they learned one dialect in their natal group, dispersed to a group with a different dialect, learned the new dialect but still retained their original dialect – in effect, an epiphenomenon of dispersal, without a functional advantage.

R: We thank the reviewer for noting this alternative explanation. In the last paragraph of section 4.1 of the Discussion we have added the following sentences and citation “It is also possible that bilingualism occurs primarily in dispersing individuals that have learned one dialect in their natal roost and then dispersed to a roost with a different dialect, and learned the new dialect while retaining the original one (Salinas-Melgoza and Wright, 2012). In this case, bilingualism would have no functional advantage.”

Salinas-Melgoza, A., & Wright, T. F. (2012). Evidence for Vocal Learning and Limited Dispersal as Dual Mechanisms for Dialect Maintenance in a Parrot. PLoS ONE, 7(11), e48667.

It would best to include a table (as supplemental data or in a data repository) showing the number of recordings / birds (and the number of calls analyzed for each recording / bird) with dates for all recording sites.

R: We have included this data as Supplemental Table 1.

Another supplemental table should report the full information on the PCA: percent variation for each component and loadings for each variable for each component, etc.

R: We created a Supplementary Table 2 showing all 26 variables measured in the PCA for the first 5 PC variables, which together explain 77% of the variation in call acoustic measurements

Other comments

Throughout:

- It is much easier to describe suggested changes/edits when papers submitted for review have line numbers.

R: We apologize for this oversight. Indeed, it would have been much easier to make suggestions using line numbers, but they were not required in the Frontiers instructions to authors so we did not include them. We include them in this revised version.

- Avoid using the words “exhibit” and “exhibited”. Instead use “have”, “had”, “use”, or “used”.
- Avoid “utilize”; replace with “use”.
- “Discreet” should be changed to “discrete”

- Avoid the passive voice whenever possible.

R: The changes described by the reviewer in the above bullet points have been made.

Abstract, penultimate line: change to “Vocal dialects in this species are ...”

page 2, para 3, line 8: change to “1996 was temporally and geographically stable.”

Page 2, para 3, line 14: change to “... showed that member of this species are less ...”

Page 2, para 3, line 15: change to “... because using the proper ...”

Page 3, para 1, line 8: change to “... we would expect to see ...”

Page 3, last para, line 2: delete “that”

Page 3, last para, line 3: change “determined” to “chosen”

Page 4, last para, line 2: change “less” to “fewer” (both cases)

Page 5, first para, penultimate line: change “modelled” to “modeled”

Page 5, para 2, line 11: delete “for”

Page 7, para 3, line 3: change “recorded” to “found” or “distinguished”

Page 7, para 3, line 6: change “give” to “gave”

Page 7, para 3, line 8: “norther” should be “northern”

Page 7, para 3, line 10: change “exhibited bilingualism” to “were bilingual”

Page 8, para 3, line 1: “2” should be “two”

Page 8, para 3: The first sentence states that there were two distinct patterns in the data. The second sentence describes the first pattern. Is the third sentence intended to describe the second pattern? If so, say so.

Page 8, para 5, line 6: “dialectal” should be “dialect”

Page 8, para 5, line 7: delete “that”

Page 8, para 5, final line: “less” should be “fewer”

Page 8, final para, line 4: change to “Based on visual categorization of sound spectrograms, we defined of 14 different call types, confirming that a pattern of dialects is present ...”

Page 9, para 2, line 14: insert “for each frequency and time point” after “amplitudes”.

Page 9, para 2, line 15: change to “Thus, when call types use the same bandwidths and very similar fundamental frequencies, the cross ...”

Page 9, para 2, line 17: perhaps insert “temporal and frequency modulation” before “differences”?

Page 9, final para, line 7: change “as to” to “of”

Page 9, final para, line 12: insert “dialects in” before ‘yellow’

Page 9, final para, penultimate line: delete “as seen”

Page 9, final para, last line: insert “and then are themselves copied accurately” just before the period. (Without that step, there can be no dialect formation).

Page 10, first para, line 1: change to “vocal dialects show”

Page 10, first para, line 3: change “like” to “such as”

Page 10, first para, line 6: delete “imperfect learning and” (see previous comment about the point that imperfect learning alone is insufficient; the term “cultural drift” includes the entire process)

Page 10, first para, line 7: change to “Long term studies by Wright and colleagues on yellow...” and delete “also”. (Since the Dahlin paper has not yet been published, it cannot be described as “previous”).

Page 10, first para, line 10: change the sentence starting on this line to “Wright (2000) found marked differences in neighboring dialects.”

Page 10, first para, line 12: change to “the species’ range”

Page 10, para 2, line 2: insert “to be” after “appeared”

Page 10, para 3, first two sentences: are these rapid population declines associated with habitat loss and fragmentation? If so, say so, in a new second sentence or a new clause at the end of the first sentence, and change the beginning of the second sentence to: This habitat fragmentation in the birds’ geographic range may serve... “

Page 10, para 3, line 3: change to “... distribution. However, there is no evidence that the current ...”

Page 10, para 3, line 4: delete “Rather, “

Page 10, last para, last line: change to “these areas, our data are likely to be an accurate ...”
(avoid basing scientific arguments on beliefs)

R: All of the above suggestions have been implemented in the manuscript.

Figure 1. The text on the figure reading “Year” is too small.

R: We made the text larger on the figure.

Figures 2 and 3 could be combined into a single figure (two panels).

R: We appreciate this suggestion. While we agree a panel figure would look nice, we are concerned that it would diminish readability and so have elected to keep them as separate figures.