



Shared calls in repertoires of two locally distant gray parrots (*Psittacus erithacus*)

Tereza Roubalová, Nicolas Giret, Dalila Bovet, Richard Policht, Jitka Lindová

► To cite this version:

Tereza Roubalová, Nicolas Giret, Dalila Bovet, Richard Policht, Jitka Lindová. Shared calls in repertoires of two locally distant gray parrots (*Psittacus erithacus*). *Acta ethologica*, 2020, 23 (3), pp.155-171. 10.1007/s10211-020-00350-x . hal-02929991

HAL Id: hal-02929991


<https://hal.science/hal-02929991>

Submitted on 9 Nov 2020

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.

1 **Shared Calls in Repertoires of Two Locally Distant Grey Parrots (*Psittacus erithacus*)**

2 Tereza Roubalová^{1,*} , Nicolas Giret² , Dalila Bovet³,  Richard Policht^{4,5}  and Jitka
3 Lindová¹ 

4 ¹ Faculty of Humanities, Charles University, U Kříže 8, Prague 5, 15800, Czech Republic

5 ² Neuroscience Paris Saclay Institute, UMR9197 CNRS, Université Paris Sud, Université
6 Paris Saclay, 91400 Orsay, France.

7 ³ Laboratoire Ethologie Cognition Développement, UPL, Univ Paris Nanterre, F92000
8 Nanterre France

9 ⁴ Department of Ethology, Institute of Animal Science, Přátelství 815, Prague-Uhřetěves, 104
10 00, Czech Republic

11 ⁵ Department of Game Management and Wildlife Biology, Faculty of Forestry and Wood
12 Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Prague 6, 165 21, Czech
13 Republic

14
15 *Corresponding author: *E-mail address:* TerezaRoubalova@seznam.cz, (T. Roubalová)

17 **Acknowledgements**

18 We wish to thank Anna Pilátová, Ph.D. for her suggestions on improving the manuscript. We
19 also thank Denisa Kováčsová, Katarina Prikrylová and others for taking care of parrots in the
20 Czech laboratory. This study was supported by The Ministry of Education, Youth and Sports
21 – Institutional Support for Longterm Development of Research Organizations – Charles
22 University, Faculty of Humanities (2019), the Charles University Grant Agency (GAUK

822119), and Specific Academic Research Projects 2019 of the Charles University, Faculty of Humanities, No. 260 469. Funding for the French laboratory was provided by the 6th PCRD Nest Pathfinder “What it means to be human”. All experiments comply with Czech and French legislation concerning animal care.

Abstract

Parrots belong to the handful of animal species capable of vocal production learning. They are considered to be open-ended learners with complex and variable vocalizations. It is not known, however, to what extent their repertoires are shared among individuals within a group or between vocally isolated individuals or groups. In Study 1, we mapped the repertoire of four captive African grey parrots (*Psittacus erithacus*) using a combination of three acoustic analyses. In Study 2, we compared the repertoire of two female African grey parrots from two different laboratories who had never been in vocal contact with each other or any member of the other parrot's social group. Results of Study 1 showed a relatively large agreement between all three methods used to analyze the vocalizations. Almost three quarters (72.8%) of categories determined by visual-acoustic analysis were confirmed by at least one of the two computer-aided methods used, i.e., by spectrographic cross-correlation and/or a multiparametric statistical method. In Study 2, we found a relatively large proportion of calls shared between the repertoires of the two grey parrot subjects. Over half and over a fourth of calls produced by parrots with the smaller and the larger repertoire, respectively, were shared between the two. No previous study identified such a large proportion of intergroup shared calls within this parrot species. It seems that some calls tend to reappear in vocally isolated groups based on inherited predispositions, similarly as has been documented in songbirds.

Keywords: vocal learning; African grey parrot; vocalization; spectrographic cross-correlation

Introduction

Social vocal production learning, which takes place when individuals modify their vocalizations by imitating others (Janik and Slater 2000), requires complex neural adaptations (Jarvis 2004) and is considered relatively rare. So far, it has been identified only in a few mammalian species, including humans, cetaceans, pinnipeds, bats, and elephants (Jarvis 2004; Janik 2014; Reichmuth and Casey 2014; Knörnschild 2014; Stoeger and Manger 2014), and in three groups of birds, namely songbirds, hummingbirds, and parrots (Brainard and Doupe 2002; Mooney 2009; Araya-Salas and Wright 2013; Pepperberg 2010).

Unlike songbirds, practically all parrot species are considered open-ended vocal learners (Bradbury 2003; Dahlin et al. 2014), which means they have the ability to modify their repertoires even in adulthood (Todt and Geberzahn 2003). Their vocal repertoires are highly complex and variable (Bradbury 2003), which is probably why parrots so far received less attention than songbirds in terms of repertoire mapping. Nevertheless, there are several studies on the vocal repertoire of parrot species such as budgerigars (Brittan-Powell et al. 1997; Hall et al. 1997), lilac-crowned amazons (Montes-Medina et al. 2016), keas (Schwing et al. 2012), orange-fronted parakeets (Bradbury et al. 2001), and yellow-faced parrots (De Araújo et al. 2011).

Among parrots, the African grey parrot (*Psittacus erithacus*) is a species universally known for its outstanding ability to mimic both animate vocalizations and inanimate sounds (Cruickshank et al. 1993; Pepperberg, 1999). It has been demonstrated that they can learn to imitate human vocalizations and acquire some understanding of the concepts related to human words. Specifically, using the Model/Rival (M/R) technique based on imitation of speech between two humans by a parrot motivated to take part in the interaction, grey parrot Alex

learned to label numerous objects, quantities, materials, shapes, and colors, and to use the concept of the same and different (Pepperberg 1999, 2002, 2010).

Social learning of vocalizations in semi-natural conditions has been studied in budgerigars (e.g. Hile and Striedter 2000; Hile et al. 2000) and parrotlets (e.g. Berg et al. 2012; Wanker et al. 2005), whose vocal repertoire is, however, significantly narrower than that of grey parrots. One should not, therefore, generalize the findings of these studies to parrot species with rich vocalization repertoires.

The vocal repertoire of grey parrots has been studied both in captive individuals and in the wild (Cruickshank et al. 1993; Giret et al. 2012; May 2004) with variable results. Comparisons between repertoires of grey parrot groups are scarce and existing studies offer no conclusions regarding the extent to which repertoires are shared or unique to individuals or groups. Cruickshank described 88 call types in a wild grey parrot pair, including nine call types of other bird species and one bat call. May (2004) found 39 acoustic call types sorted in four classes in wild grey parrots, while Giret et al. (2012) distinguished 168 different grey parrot calls in their captive group, 30 of which were imitations of French labels. These authors also noted it was difficult to find calls shared between captive parrots and parrots in the wild and suggested that most grey parrot call types are learned. Begging calls, used shortly after hatching to indicate distress and elicit allofeeding, seem to be the only calls considered clearly innate (Hall et al. 1997).

While literature is far from unanimous with respect to the composition and size of grey parrots' vocal repertoire, our knowledge of the role of social environment in its development is even more limited. To improve our understanding of these subjects, in the first study reported here we tried to describe the vocal repertoire of this species. In the second study, we compared the repertoires of two locally distant individuals to arrive at a rough estimate of

similarity between repertoires of locally distant grey parrot groups. We worked with a group of four captive African grey parrots and one other captive grey parrot who never had any contact with this small group. We established the repertoire of all four captive grey parrots using a combination of three acoustic analyses and compared the repertoire of one female parrot from this small group with the repertoire of a female grey parrot from a different laboratory. Identification of shared calls in the repertoires of these two females who had no contact could serve as an argument for the existence of calls allegedly produced by all grey parrots, possibly based on some innate component. A comparison of proportions of shared inter- and intra-group calls could be the first step towards assessing the role of social learning versus innate predispositions in the establishment of repertoire in this species.

Study 1: Repertoire mapping

In agreement with previous studies (e.g. May 2004, Giret 2012), we assumed that grey parrots' calls can be classified into discrete categories. The vocal repertoire of a species is mapped by visual-acoustic inspection of spectrograms by human categorizers (Janik 1999; Bloomfield et al. 2005). The main disadvantage of this approach is that the threshold values for categorizing vocalizations are not clearly defined (Janik 1999). Moreover, humans might well miss differences of importance to the animals and more generally, classification by human categorizers is inevitably biased by subjective perception. The biological meaning of majority of grey parrots' calls is unknown and moreover, it is uncertain to what extent grey parrots' acoustic perception and spontaneous tendency to categorize calls is similar to human acoustic perception and categorization. Nevertheless, it has been argued that the acoustic systems of parrots and humans are largely analogical (e.g. Bottoni et al. 2009; Patterson and Pepperberg 1998; Pepperberg 2010) (see also discussion). The presence of multiple categorizers, proposed by Janik (1999), may reduce the inherent subjectivity but inter-categorizer reliability could be influenced by several factors (Jones et al. 2001). For these reasons, several

computer-aided methods had also been recently applied to the mapping of animal vocal repertoire.

The first commonly used computer-aided approach is spectrographic cross-correlation method developed by Clark et al. (1987). Spectrographic cross-correlation produces a similarity matrix between pairs of signals in a database. Similarity is measured as the peak of correlation between spectrograms of two calls by sliding two spectrograms across one another on the time axis. Each pixel in one spectrogram is compared with the corresponding pixel in the second spectrogram and time shift between the two is gradually modified until the point of maximum similarity is determined and measured, which creates a correlation matrix (Clark et al. 1987; Charif et al. 2010). This process is not affected by subjective perceptions of human categorizers and moreover, instead of focusing on particular sound parameters, it compares the entire spectrogram visualizations of calls (Khanna et al. 1997). It does, however, suffer from other limitations. The main problem is that correlation values are sensitive to fast Fourier transform frame length used for spectrogram generation (Baker and Logue 2003; Cortopassi and Bradbury 2000; Khanna et al. 1997). Moreover, the efficiency of this method decreases with increasing background noise in the recorded calls (Cortopassi and Bradbury 2000; Khanna et al. 1997) and the method fails to detect small differences in the sound structure that could be significant (Khanna et al. 1997). And even more importantly, the abovementioned studies tested the efficiency of spectrographic cross-correlation on tonal or harmonic calls but sounds with atonal qualities were not tested. This method's performance with respect to atonal sounds is thus yet to be examined.

Another commonly used computer-aided approach is based on extracting parameters directly from the calls (e.g. amplitude, call duration, energy distribution, frequency bandwidth). These parameters typically represent mathematically well-defined frequency and temporal characteristics of the call. This method's main advantage is that it avoids categorizer's

subjectivity and it can be highly accurate (e.g. Pérez-Espinosa et al. 2015; Tchernichovski et al. 2000). On the other hand, it is limited by the range of feature sets used, which need not be applicable to a particular categorization (Giret et al. 2011). Parameters are extracted either manually or automatically using various commercial sound programs such as Avisoft-SASLab Pro [Raimund Specht, Berlin] or Raven Pro [Cornell Lab of Ornithology] custom programs such as the LMA, which focus on atonal sounds (developed by K. Hammerschmidt; see Schrader and Hammerschmidt 1997), or Sound-analysis R packages, such as Seewave (Sueur et al. 2008), EchoviewR (Harrison et al. 2015), or warbleR (Araya-Salas and Smith-Vidaurre 2017). Extracted parameters are then processed using a discriminant function analysis. This method has been applied to several vocal studies of primates (Hedwig et al. 2014; Fischer et al. 2001; Pfefferle et al. 2016; Price et al. 2014), horses (Policht et al. 2011), rhinos (Policht et al. 2008), humans (Jürgens et al. 2011), passerine birds (Naguib et al. 2001), and cockatoos (Zdenek et al. 2015) and proved to be highly useful. In the following, we call it the multiparametric statistical method.

Aside from these commonly used methods, Giret et al. (2011) proposed an approach known as feature generation method. This classification method combines human categorizers' involvement with automatic feature measurement. In particular, visual categorization is used as a training database subsequently used to train and test a classifier by a supervised classification technique (Giret et al. 2011). Feature generation outperforms methods based on standard features in studies on animal vocalization (Giret et al. 2011; Larranaga et al. 2015; Molnár et al. 2008). On the other hand, due to its poor performance on full repertoires, it has only been applied to a limited number of call types. Giret et al. (2011) suggest that the best approach is to combine the complex integrating capability of the human eye with the objective specificity of an automatic process.

In Study 1, we classified vocalizations of four captive bred grey parrots (*Psittacus erithacus*) from a Czech laboratory using visual-acoustic classification by human categorizers. Categorization determined by the visual-acoustic method provided the foundation for its verification by two computer-aided methods: (1) a comparison of cross-correlation coefficients and (2) a multiparametric statistical method using the discriminant function in statistical analysis. We expected the cross-correlation method to perform better in the classification of tonal and harmonic calls, while the multiparametric statistical approach, which works with LMA parameters (Schrader and Hammerschmidt 1997) and parameters measured in Raven Pro 1.4 (Cornell Laboratory of Ornithology), were expected to perform similarly well on both tonal (harmonic) and atonal calls.

Material and methods

Subjects

Four captive grey parrots participated in this study: Durosimo (male), Jaro (male), Jarina (female), and Juruba (female). Durosimo, Jaro, and Juruba were part of a group that arrived from the Congo as subadults. Jarina was also wild born, with an unknown place of origin. At the time of recording, Durosimo, Jaro, and Juruba were 12 years old (based on seller's information). Jarina's age could not be determined. Durosimo, Jaro, and Juruba had never been handled by humans and were kept in an indoors aviary together. Jarina was habituated to human presence and handling. She had previously participated in language-learning experiments (see Giret et al. 2010). At the time of the recording, she had been kept together with the other three parrots for several years.

At the time of recording, subjects were kept together in one room. Each individual was housed in a cage measuring 90x60x60cm. The cages were placed about a meter apart, so the

parrots had visual and acoustic contact with each other. They could also hear other grey parrots in an adjacent room and various bird species living outside.

Recordings

The parrots were recorded from June to November 2008 (on 23 and 27 June, 13 and 21 July, 4 and 11 August, 31 October, and 1 November). Recording was conducted using Sennheiser ME 67 microphone and an external USB preamplifier MOBILEPRE-USB in 44 kHz (16 bits) wav format. The recording position was at the wall opposite the cages, app. 2m away from them, thus ensuring equal distance from all cages. This way, amplitude could be taken into account in subsequent analyses.

Most recording days consisted of three blocks: a morning block, a noon block, and an evening block. The morning block started shortly after dawn (e.g., 4:30 am in the summer) and the evening block ended at dusk (e.g., 9 pm in the summer). The noon block took place during the day, when the parrots were less vocally active (e.g. 9 am to 11 am in the summer). We included the noon block mainly into the early recording days to make sure we would not miss any categories that might be performed only in particular situations, such as feeding. It turned out, however, that parrots vocalize at this time of the day minimally and no new categories emerged. Later in the process, we have therefore omitted the noon block. Each recording block started with habituating the parrots to the recording situation and human presence. An observer entered the room with a microphone and sat quietly on a chair at the wall, facing the cages with parrots. Such habituation session preceded a recording session because our previous experience had indicated that parrots tend to change (or completely stop) their vocalizations after a human enters their room. In the initial sessions, the habituation lasted for half an hour, later it was adjusted to the parrots' reaction.

Each parrot was recorded for 15 minutes per block (3x15 minutes per day). In total, each individual was recorded 23 times and we obtained 345 minutes of recording. As shown in Figure 1, this recording time seems sufficient to describe the repertoire of the entire grey parrot group. Nevertheless, the repertoires of individual parrots could possibly expand until reaching the group numbers with additional recording time. During recording, the calls of a targeted individual were verbally marked on the recording by the observer. Marks directly followed the calls they related to and were not extracted with the calls. Individual calls were then manually demarcated using a spectrogram view and extracted from recordings into separate wav files using SoundForge 8.0. Calls that overlapped with background noise or with other parrots' calls were not included in the following analysis.

Analyses of recordings

The visual-acoustic method (VA)

Spectrograms were generated in SoundForge 8.0 with sample rate 22 kHz at 16 bits using a 2048-point fast Fourier transform length, with Blackman-Harris smoothing window and 75% window overlap, frequency resolution 21.5 Hz, and time resolution 0.046s. In the visual-acoustic (VA) analysis, calls were sorted in categories that were based on acoustic perception and visual perception of the spectrograms. Categorization was executed qualitatively by several rounds of individual sorting of calls, followed by search for agreement and discussion of differences between categorizers. All categorizers were trained in behavioral observation and vocal analysis, had extensive personal knowledge of the birds, and had previously participated in parrot language training and other experiments with these subjects.

In the first step, three categorizers (A, B, and C) independently categorized calls they had previously recorded. They identified 22 (A), 67 (B), and 44 (C) categories. Their agreement was relatively low: categorizer B and C agreed on 35 categories, but only 2 of these matched

categories identified by categorizer A. In most cases, one category by categorizer A included calls sorted into two or more categories identified by categorizers B and C. Subsequently, only categorizers A and B worked on categorization. They discussed clusters of similar categories (or one of categorizer A's original categories) and sought ways in which finer categories – typically based on a previous categorization by categorizers B and C – could be distinguished within such clusters. Acoustic and visual spectrum characteristics used to distinguish the resulting categories included: tonality vs. noise, dominant frequency, spectrum bandwidth, intensity of harmonics, and frequency modulation (e.g. a rising or falling trend). When agreeing on a set of distinguishing characteristics, both categorizers independently sorted calls from a cluster into these proposed categories. If they reached at least 75% agreement of assigned calls, the category was accepted into final categorization. If not, a new category was proposed. A new category was adopted when at least three calls could be identified as belonging to it. Categories were tagged by letter “K” followed by a number representing the frequency of calls belonging to the category in question (so that, e.g., K1 was the most frequently used category). The final set of categories was further split in clusters of similar categories of calls based on categorizers' A and B consensus. These clusters were then used in the multiparametric statistical method (see below).

Spectrographic cross-correlation (SPCC)

From calls categorized by the visual-acoustic method, we extracted a sample for the spectrographic cross-correlation (SPCC) method. The sample consisted of ten randomly selected calls for each of the four parrots in each of the 70 categories defined during the process described above. If a parrot produced ten or less calls in a category, all were included. The final sample consisted of 1,516 calls. We performed batch correlations in Raven Pro 1.4 by comparing spectrograms of all calls from the individual categories.

Spectrograms were calculated with 512 samples, 3dBFilter Bandwidth 135 Hz, hop size 256 samples, and grid spacing 93.8 Hz, using a 512-point discrete Fourier transform length with a Hann smoothing window and 50% window overlap, frequency resolution 86.1 Hz, and time resolution 0.012s. For each category, we computed a “category correlation matrix”, that is, a correlation matrix of peak correlations within a category. Similarity of calls within a category (which we further refer to as “inner consistency”) was calculated as the average of all values in the category correlation matrix. We set the lower boundary of inner consistency – used as an indicator of sufficient similarity of calls within a category – to 0.2 because individual calls in 11 categories sounded very similar to human ear but had inner consistency only between 0.2 and 0.3. Subsequently, we employed the VA method to select for each category several other categories (“outgroups”) with similar acoustic and visual parameters and/or a likely overlap of calls categorized into the assessed and the outgroup category. For each assessed category–outgroup pair, we calculated an “outgroup correlation matrix”, i.e., a correlation matrix of peak correlations between calls from the assessed category and calls from its outgroup category. T-tests comparing correlation coefficients in the category correlation matrix with outgroup correlation matrix were used to test whether similarity of calls within a category is greater than the similarity of calls between this category and another category (outgroup). If inner consistency was not significantly higher than the outgroup correlation, we considered merging the assessed category with its outgroup.

The multiparametric statistical (MS) method

For the multiparametric statistical (MS) method, we used the same sample of calls as for the SPCC method. From the spectrograms, we measured a number of acoustical parameters using automatic algorithms. Spectrograms were processed using the same parameters as in the SPCC method in Raven Pro 1.4 (Cornell Lab of Ornithology). Eight parameters were

290 measured using semi-automatic measurements (manual demarcation of a call and an
 291 automatic computation) in Raven Pro 1.4. Demarcation of calls was based on visual
 292 detectability of the sound in a spectrogram. In Raven Pro 1.4, the parameters were: *High*
 293 *Frequency [Hz]* (the upper frequency boundary of call selection), *Center Frequency [Hz]* (the
 294 frequency which divides the call selection in two frequency intervals of equal energy), *Energy*
 295 *[dB]* (the total energy of call selection), *Interquartile Range Bandwidth [Hz]* (difference
 296 between the first and third quartile frequencies, where quartiles describe the distribution of
 297 energy of a call on the frequency axis), *Interquartile Range Duration [s]* (difference between
 298 the first and third quartile times, where quartiles describe the distribution of energy of a call
 299 on the time axis), *Length* (the number of frames contained in a call selection), *Max Power*
 300 *Frequency [Hz]* (the frequency at which Max Power occurs within the call selection), *Max*
 301 *Power[dB re 1]* (maximum power in the call selection) (Charif et al. 2010). Additionally, the
 302 spectrograms were analyzed with LMA 2008 software (Schrader and Hammerschmidt 1997).
 303 A set of 12 temporal and frequency parameters was used, many of which measure the
 304 statistical distribution of frequency amplitudes in the spectrum (DFA) or describe the first
 305 three dominant frequency bands (DFB). DFA values mark the frequency at which amplitude
 306 distribution in the spectrum reaches the first, second, and third quartile of total distribution.
 307 DFB are defined by amplitudes that exceed a given threshold in a consecutive number of
 308 frequency bins. We employed the following measurements: *q1mean [Hz]* (mean frequency of
 309 the first DFA), *q2mean [Hz]*, (mean frequency of the second DFA), *q3mean [Hz]* (mean
 310 frequency of the third DFA), *df1mean [Hz]* (mean frequency of the first DFB), *df1chfre*
 311 (number of changes between the original and floating average curve of the first DFB),
 312 *df1maloc* (location of the maximum frequency of first DFB), *df2mean [Hz]* (mean frequency
 313 of the second DFB), *df3 mean [Hz]* (mean frequency of the third DFB), *ranmean [Hz]* (mean
 314 frequency range), *pfmax [Hz]* (the highest of frequencies with peak power established in

consecutive time segments during a call), *pfmean* [Hz] (mean peak frequency), and *pfjump* [Hz] (the maximum difference of peak frequency in two consecutive time segments). Both LMA and Raven parameters were used because LMA parameters had been defined mainly for atonal calls (Schrader and Hammerschmidt 1997). They could thus be less well suited for tonal and harmonic calls, which is what Raven parameters should compensate for. Subsequently, we applied a forward stepwise discriminant function analysis (DF) to confirm the VA categorization. To validate the results of discriminant analysis, we implemented a cross-validation process based on leave-one-out method in IBM SPSS Statistics for Windows, Version 21.0. Since the DF can discriminate between a limited number of categories, it was applied to clusters of similar call categories established by the VA method (see Table 3). Clusters were tagged by letter “C” followed by a number representing the frequency of use of that cluster (so that C1 was the most frequently used cluster).

We compared the results of these methods. The SPCC method and the MS approach resulted in recommendations to merge several call categories. We compared the merging suggestions from the two methods and where both computer-aided methods proposed the same merger, the categories were merged. If the merger was proposed only by one computer-aided method, it was revised by a repeated inspection of VA features of the categories concerned.

Results

We obtained a total of 3,052 calls from our four subjects. Three subjects were similarly vocally active (Jaro 29.23% of all recorded calls, Juruba 28.15%, Jarina 26.34%), while Durosimo produced less vocalizations (16.32%).

Using the VA categorization, we identified 70 categories containing 3 to 238 calls (see Table 1). No individual used all categories (see Table 2a) but the 20 most frequent categories were detected in all parrots. Categories were classified into 18 clusters (see Table 3).

The SPCC method resulted in a reduction of categories to 47. Thirty-three VA-determined categories had a sufficiently high inner consistency (above 0.2) and correlations in the category matrix were significantly lower than correlations in all tested outgroup correlation matrices as measured by a t-test. The remaining 14 categories were formed by merging two (nine cases), three (three cases), or four (two cases) categories determined by the VA analysis. We decided to merge these categories because their outgroup correlations did not differ from their inner consistencies (when compared by t-tests) and/or they showed a low inner consistency (under 0.2). For t-tests results, see Table S1. Two categories with low inner consistency were discarded since no attempted merging with an outgroup led to inner consistency of such new category above 0.2 (see Table 1).

Similarly, the MS approach resulted in a reduction of the number of categories formed by the VA method. In 10 (out of 18) clusters, a merger of two or more categories led to more precise discrimination between categories (see Table 3). We considered a merger when over 25% of calls were classified into a different category. From the 70 categories established by the VA method, 46 were well established in the MS approach. Nine “new” categories were suggested by merging two (six cases), three (two cases), or four (one case) categories. Two categories were excessively scattered and discarded due to no suitable merging category. The total number of categories proposed after implementation of the MS approach was 55 (see Table 1).

In seven cases, there was an agreement between the SPCC and MS approach on which VA-established categories should be merged. In another seven cases, however, the merger was proposed only by one of the methods, which led to a re-inspection of the VA features of the

categories in question. Overall, 51 (72.9%) categories determined by the VA could be confirmed by at least one of the two computer-aided methods.

Discussion

The use of computer-aided methods resulted in a relatively extensive confirmation of categories established by the VA method. In particular, computer-aided methods helped verify the discrimination of VA categories and confirm similarity among calls falling into the same VA category (relative to calls classified into other VA categories).

The VA method allowed for the formation of a relatively high number of categories characterized by high acoustic and visual coherence but the differentiation threshold among categories was sensitive to the subjectively variable perception of each categorizer. The SPCC method was based on an objective mathematical comparison of calls but even here, a subjective element was present: for instance, thresholds had to be set by categorizers to define sufficient consistency of a category. The shortcomings of SPCC method had to do with distinctiveness: lack of distinction between some similar categories may have been due to SPCC's low sensitivity with respect to atonal and noisy calls that lacked clear frequency contours. The MS method was capable of distinguishing among relatively acoustically close categories of both noisy and harmonic calls thanks to its reliance on a combination of parameters suited to both atonal and tonal/harmonic calls. Nevertheless, while the DF has been efficiently used in studies with a lower number of categories (e.g. Hedwig et al. 2014, Price et al. 2014), differentiation among a higher number of categories using this method is problematic (Zdenek et al. 2015). We have therefore set the upper limit to ten categories and performed a discriminant analysis on clusters of similar categories. Generally speaking, unlike the VA method, computer-aided methods do not evaluate a sound as a whole. They usually

385 extract only part of information contained in a sound and that may be insufficient. Divergence
386 between computer-aided methods may be due to differences in parameters they take into
387 account when analyzing calls. Human categorizers are capable of assessing sounds
388 holistically. An ideal computerized method would, similarly, find all parameters that capture
389 the acoustic variability detected by the VA method. At present, neural network approaches
390 seem to be closest to this goal but they need extensive training with preprocessed data and the
391 resulting extracted categorization rules are hard to interpret.

392 Although we do not know to what extent grey parrots' intraspecific acoustic perception
393 system tends to sort calls into discrete categories or rather process them within a continuous
394 gradient or to what level of detail parrots themselves distinguish their calls, based on
395 following indirect indices we suggest that the perception categories identified by the VA
396 method are close to what parrots themselves may perceive as distinct. Despite lack of studies
397 on grey parrots' acoustic perception, it is reasonable to assume that to process sounds as
398 discrete categories of calls is these birds' natural ability and their acoustic perception operates
399 in similar ranges as in humans. Given that perception is needed for reproduction, this
400 assumption of similarity finds support in parrots' ability to accurately reproduce human
401 vocalizations, to discriminate between phones and to reuse learned strings of phones in
402 producing new words (Bottoni et al. 2009; Patterson and Pepperberg 1998; Pepperberg 2010).
403 Auditory studies on budgerigars (Okanoya and Dooling 1987) and orange-fronted conures
404 (Wright et al. 2003) suggest that there are both possible parallels and differences between
405 humans, parrots, and songbirds. For most songbirds and parrots tested, the regions of best
406 auditory sensitivity corresponded to the optimal range of human auditory sensitivity (2–5
407 kHz) (Dooling 1980; Wright et al. 2003). On the other hand, birds in general are better at
408 discriminating absolute pitch (Weisman et al. 2004). Playback experiments in grey parrots

would be required to evaluate their sound perception abilities and their similarities with and differences from human ones.

Our results suggest that a combination of the VA method with one or more computer-aided methods could be well suited to analysis of the vocal repertoire of grey parrots. This corresponds to the findings of Giret et al. (2011). Nevertheless, the computer-aided methods used in this study all turned out to have some weaknesses and each performed well only on part of the parrots' rich and complex vocal repertoire.

Study 2: Intergroup repertoire comparison

In Study 2, we used the repertoire established by the VA method to preserve the methodology employed in Giret et al. (2012), which we used as a comparative material. In particular, we used the typical calls of each category to compare the repertoire of two individuals. Typical calls (prototypes) for our parrot's repertoire were determined by the SPCC as calls with the highest average of peak correlations with every call in a given category. Prototypes for the other parrot's repertoire were identified in a similar manner. One could expect that by using typical calls, we would maximize the distinctiveness of calls from different categories and lower the likelihood of finding a similarity between two compared repertoires.

Subjects

We performed a comparison of repertoires of two female grey parrots, Zoé and Juruba. The grey parrot Zoé from a French laboratory had been hand-reared in captivity and purchased from breeders. She arrived at the French laboratory in September 2003 aged 3 months. Since her arrival at the laboratory, she has been in close contact with humans and was trained to referentially imitate French labels. She was housed with three other grey parrots in an indoor aviary. From the aviary, Zoé could hear domestic canaries and various noises.

For the comparison, we chose Juruba, a female grey parrot from the Czech laboratory who produced the largest number of categories of sounds. Zoé and Juruba had never seen nor heard each other and had no previous contact with any member of the other parrot's social group. See Study 1 for information on Juruba from the Czech laboratory.

Material and methods

Zoé had been recorded regularly since her arrival at the laboratory in September 2003 until March 2007, i.e., from the age of three months until four years of age. All recordings were conducted in 22 kHz (16 bits) wav format with high-quality microphones. For further details, see Giret et al. (2012). For details on recording Juruba, see above in Study 1. The sampling rate of the two sets of recordings was the same and spectrograms were created using identical parameter settings. Both laboratories collected their recordings indoors, leading to both sets of calls being similarly affected by resonance and relatively low noise. There was a significant difference between the two individuals' recording periods: the relatively short recording period in Juruba may have somewhat reduced the number of call categories identified for her (see Figure 1).

We selected one prototype for each call category in the repertoire of each parrot. The set of prototypes for Juruba included 62 calls (Figure 2). The set of prototypes for Zoé contained 100 calls (Figure 3), whereby none were imitations of the French labels. Comparison between the two sets of prototypes was performed by the VA method in a two-step procedure. In the first step, the repertoires were compared by two independent categorizers (B and D) who tried to match Juruba's and Zoé's prototype calls. The categorizers mostly chose at most two possible call counterparts but in six cases, one categorizer chose three or four possibilities. Agreement between these two main categorizers on either assigning the same counterpart to a

call or finding no counterpart was 51.6%. In the second step, we therefore recruited two additional trained categorizers (E and F), who re-evaluated the cases of disagreement between categorizers B and D and decided between the counterparts suggested by them. In 100% cases, expert discussion concerning the problematic calls led to an agreement between the two new categorizers and one of the primary categorizers.

Results

Thirty-two (51.6%) calls from Juruba's set of prototypes were matched to similar calls in Zoé's set of prototypes (Figure 4). Similarly, 29 (29%) of Zoé's prototype calls were matched to similar calls in Juruba's set. In five cases, Zoé's calls matched two counterparts from Juruba's repertoire. In contrast, the two females from the Czech laboratory, Juruba and Jarina, share on average 88.4% ($SD = 7.8$) of their repertoire and average similarity of repertoires between any two parrots from the Czech laboratory is 84.4% ($SD = 6.2$) (see Table 2b). From the 70 categories determined by the VA method, all four parrots from the Czech group performed 37 categories (52.9%) at least once, while three of the four parrots shared additional 19 categories (27.1%), 12 categories (17.1%) were shared by two individuals, and the two remaining categories (2.9%) were produced only by Jaro, the male. Juruba herself, in addition to sharing 37 categories (59.7%) with all three other parrots from her social group, shared 17 additional categories (27.9%) with two parrots, and eight categories (13.1%) with one parrot. Nineteen (59.4%) of the categories shared between Juruba and Zoé were also shared by all four individuals in Juruba's social group. Another eight categories (25.8%) of those which Juruba shared with Zoé were also shared between Juruba and two individuals from her group, while the remaining five categories (16.1%) were shared between Juruba and one individual from her group.

Discussion

We found a relatively large proportion of shared calls in the repertoires of two grey parrots who never had any vocal contact with each other nor any member of the other parrot's social group and who were surrounded by humans speaking different languages (Czech in Juruba's case, French in Zoé's). Over half and over a fourth of calls produced by parrots with the smaller and larger repertoire, respectively, were shared between the two. Giret et al. (2012) lists only three categories allegedly used by all grey parrots. The first shared call, described as "squawk", was identified by Cruickshank et al. (1993), Giret et al. (2012), and May (2004). It is similar to our K8 category. The call was detected either during excitement (Giret et al. 2012) or in preflight and agonistic situations (May 2004). The other two shared calls suggested by Giret et al. (2012) were emitted during distress and in agonistic interactions. Those calls were not found in Juruba's repertoire, but they do appear in the repertoire of other individuals kept in our laboratory (data not shown).

A comparison between the proportion of calls that Juruba shared with members of her group and with the unfamiliar conspecific, Zoé, gives us some idea about the role of social learning in the establishment of vocal repertoire in grey parrots. Nevertheless, interpretation of calls shared within a group as socially learned as well as assumption that (only) calls shared between two distant individuals are possibly based on some innate predispositions are to some extent tentative.

We observed a high proportion of calls shared within a group. It is likely that this proportion may have been even higher had the recording period been expanded, eventually approaching 100% of the group repertoire, as indicated by Figure 1. Nevertheless, individual call categories can be either socially learned, learned from environment, or "invented" (e.g. by vocal play) by an individual parrot. Interestingly, for Juruba, Jarina, and Durosimo, we observed no calls specific to individuals. Only the male, Jaro, produced two call types never

recorded in another group member. Still, it is possible that group members learned calls originally performed by Juruba or they may have developed the same calls independently, perhaps due to a shared innate predisposition (and for some reason, such as some characteristics of the acoustic environment, such calls were not recorded in Zoé although she inherited the same predisposition). One thus cannot conclusively claim that all calls shared by members of our group and not shared with Zoé were socially learned by Juruba. Secondly, call sharing could be the result of parrots' parallel learning from members of other species or from their inanimate environment.

At the same time, the relatively large proportion of calls shared by two isolated individuals indicates that there may be common call categories that have a high probability to appear in all grey parrot populations by all grey parrot populations. Nonetheless, it is most likely that the results we obtained are but an inaccurate estimate of similarity of vocalizations in distinct grey parrot populations. First of all, this is due to the subjective nature of our methods of repertoire establishment and comparison. Moreover, our recording plot (Figure 1) indicates that had we continued recording, Juruba's repertoire may have expanded until reaching the levels observed for the whole group numbers, which could lead to additional identifications of calls shared by the two females, Juruba and Zoé. In short, the larger proportion of unmatched calls in Zoé's repertoire in comparison to Juruba's repertoire could be simply due to the fact that Zoé had been recorded over several years or due to different thresholds for categories in each laboratory, which may have resulted in a more detailed categorization of Zoé's vocal sample. Moreover, our comparison focused on sorting calls into certain categories and did not deal with the fact that calls assigned to one category can still be similar to varying degrees, whereby parrots themselves might infer important information from the perception of a continual degree of similarity apart from identification of a call type.

General discussion

531 A comparison of vocal repertoires of two parrot individuals from different and separate
532 settings indicates the existence of a substantial overlap in the repertoires of distinct grey
533 parrot populations. No previous study identified such a large proportion of intergroup shared
534 calls within this parrot species that is assumed to have the capacity of open-ended vocal
535 learning (Giret et al. 2012). Nevertheless, Salinas-Melgoza and Wright (2012) presented some
536 evidence of possible limitations to vocal learning in adult yellow-naped amazons and it is
537 likely that some calls tend to reappear in vocally isolated groups based on inherited
538 predispositions, in a manner similar to that which has been reliably documented for songbirds
539 (Gardner et al. 2005, Fehér et al. 2009).

540 Similar vocal patterns in independent grey parrot populations with shared innate vocal
541 predispositions may have simply developed by vocal play, that is, without any specific
542 external acoustic input. Learning of species-typical songs by “invention”, i.e., learning not
543 obviously related to tutor song, has been described in songbirds (Beecher & Brenowitz, 2005).
544 In the case of parrots, acquisition of new calls by a related mechanism called “vocal play” has
545 been proposed (Pepperberg 1999). In human language development, vocal play allows
546 children to explore the potential of their vocal tract by making new phonemic combinations
547 from preexisting labels (Stark 1980). A form of vocal play called “monologue speech,” which
548 refers to active communication without social stimulation, that is, vocalization in solitude
549 (Kuczaj 1983), has been detected in the grey parrot Alex and it is assumed it accelerated his
550 learning (Pepperberg 1999, 2002). While this mechanism obviously bears similarity to the
551 “subsong” known in songbirds (Marler 1970), it may also work analogically as in humans
552 when we consider learning of a physiologically limited variety of syllables (Ohala 1983).

553 At the same time, there is no doubt that social learning does influence the vocal repertoire of
554 grey parrots. Our results confirm it: birds from the same group shared substantially more calls
555 than one of them, Juruba, shared with the unfamiliar conspecific, Zoé. It is yet unknown

whether in grey parrots, vocal learning in a social group focuses on mimicking new call types or modification of already familiar calls by, e.g., changing their pitch, modulation, addition of new elements, etc. A study on the dialects of amazon parrot seems to indicate that gradual modification of pre-existing calls is a likely option (Salinas-Melgoza & Wright 2012). Pair call convergence in budgerigars indicates that both options are possible, because males seem to imitate new calls from female repertoire, whereas females only slightly change their own contact call (Hille et al. 2000). If the latter mechanism is prevalent, we could expect to find the same call categories in different groups, although calls within a category would resemble each other more closely among members of one social group than between unfamiliar individuals. Future research into the level of similarity between calls of the same category could shed more light on this topic.

Generally speaking, the adaptive function of extensive vocal learning in parrots remains unclear. Nowicki and Searcy (2014) offer five possible explanations for the origin of vocal production learning in animals: 1) the vocal dialect hypothesis, which states that vocal learning supports geographic variation in vocal signals and that facilitates local assortative mating and promotes local adaptation, 2) the sexual selection hypothesis, which explains repertoire expansion as a response to mating preference for more complex vocalizations, 3) the information sharing hypothesis, which explains vocal learning by the need to share information among kin, 4) the environmental adaptation hypothesis, which emphasizes vocal adaptation to local habitat, and 5) the individual recognition hypothesis, which claims that vocal learning allows for the emergence of individually specific vocal signals. The vocal dialect hypothesis (1) and sexual selection hypothesis (2) are not a likely explanation in grey parrots because there is no evidence that specific mating calls and preferences for local mates belong among behavioral strategies found in this species. The information sharing hypothesis (3) is also unlikely to be the case, because information sharing among kin by complex vocal

communication has not been demonstrated in any nonhuman species. Both the environmental adaptation hypothesis (4) and individual recognition hypothesis (5) seem to be at least partly applicable to grey parrots. These hypotheses find support in grey parrots' extraordinary mimicking abilities both in the wild and in captivity (Cruickshank et al. 1993; Giret et al. 2012; May 2004; Pepperberg 1999) and their ability to recognize each other as individuals by vocal communication (Giret et al. 2009; May 2004). Bradbury and Balsby (2016), however, suggest that the most likely explanation for the evolution of vocal learning in parrots is their need for warbles or duets used in nest defense and in the acquisition and modulation of contact calls which mediate social dynamics. So far, the most direct support for this proposal comes from studies which show that male budgerigars in captivity mimic the calls of females during courtship (Hile et al. 2000). In future studies, it would be interesting to look at calls shared in grey parrot bonded pairs.

We suppose that our group of grey parrots may use vocal learning to converge their vocal repertoires and thus strengthen the bonds within their social group, because call convergence in groups has been demonstrated in various songbird and parrot species (Balsby and Bradbury 2009; Berg et al. 2012; Enggist-Dueblin and Pfister 2002; Hile and Striedter 2000; Sewall 2009). The bonding effect of vocal learning in a group can be partly responsible for the outstanding "linguistic" abilities of Alex, who was very close to his caregiver (Pepperberg 1999) and for the disappointing results of the Model/Rival method in the study by Giret et al. (2010), where the parrots bonded with their caregivers to a much smaller extent. It has also been shown that grey parrots have little interest in learning vocalizations from playback (Giret et al. 2010; Pepperberg and McLaughlin 1996). It seems they prefer vocal learning in interaction with a live partner who is part of their social group.

The system of vocal learning resting on a shared vocal base could be evolutionary adaptive here. When conspecifics with different life histories meet, a shared vocal base might help

606 them learn from each other. Such system could facilitate flexible convergence of repertoires
607 in mating pairs and assimilation of new members into established groups. For future research,
608 it would be interesting to verify the bonding function of social learning of vocal repertoire, for
609 instance by experimentally manipulating individuals' membership in a group.

610 To map the repertoire of grey parrots, we used a combination of three methods. Outcomes of
611 the VA method were for the most part confirmed by the SPCC or the MS method or both.
612 SPCC was capable of distinguishing finer differences among harmonic calls, while the MS
613 performed well on both harmonic and atonal calls because it relied on a combination of
614 parameters suited to both atonal and tonal/harmonic calls. The comparison method used in
615 Study 2 was similar to the VA method used in Study 1 but adapted to a different task. Both
616 methods started with independent classifications of calls by two or three trained categorizers.
617 This generated a relatively low agreement but one ought to keep in mind that the level of
618 chance agreement was extremely low. Initial classification was therefore followed by
619 discussions regarding disagreements among categorizers until a consensus was reached. In
620 Study 1, discussions focused on establishing characteristics that define a category. This led to
621 agreement on a finer classification of calls. In Study 2, two additional trained categorizers
622 were called in to find consensus with one of the primary categorizers. A combination of the
623 more subjective VA method conducted by human observers and computer-aided methods led
624 to increased reliability of the final categorization of grey parrot vocal repertoire. Various
625 clustering methods, especially fuzzy c-means clustering (Wadewitz et al. 2015), might bring
626 some additional advances to mapping of vocal repertoire and deal with the possibility that
627 grey parrot's acoustic perception could, to some extent, have the form of continuous acoustic
628 gradient processing.

629 In this study, we worked with a small group of individuals. Small numbers of subjects are
630 common in laboratory studies of parrot behavior and communication due to inherent technical

limitations. Due to small sample size, the findings presented here cannot be directly generalized to the species as a whole. They can, however, be viewed as a starting point for future replication studies on different individuals from distinct social groups.

This study demonstrated the extent of similarity among repertoires of individual grey parrots from different social groups. The logical next step towards a greater understanding of the structure and dynamics of great parrots' vocal communication would take the form of observation of how group repertoire evolves over time and how it may be affected by changes in a social group. In particular, it would be interesting to observe whether long-term cohabitation or bonding of grey parrots leads to an increase in the shared repertoire or not, i.e., whether individual variability of calls in a group or dyad remains unchanged. It would also be interesting to find out whether vocal conversion is based on the mimicking of new calls or modification of pre-acquired ones, and what role grey parrots' disposition to produce a certain set of call categories plays in these processes.

Ethical approval: “All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.”

References

- Araya-Salas M, Wright T (2013) Open-ended song learning in a hummingbird. *Biol Lett* <https://doi.org/10.1098/rsbl.2013.0625>
- Araya-Salas M, Smith-Vidaurre G (2017). warbleR: an r package to streamline analysis of animal acoustic signals. <http://dx.doi.org/10.1111/2041-210X.12624>

653 Baker MC, Logue DM (2003) Population differentiation in a complex bird sound: A
654 comparison of three bioacoustical analysis procedures. *Ethology* 109(3):223-242

655 Balsby TJ, Bradbury JW (2009) Vocal matching by orange-fronted conures (*Aratinga*
656 *canicularis*). *Behav Process* 82(2):133-139

657 Beecher MD, Brenowitz, EA (2005) Functional aspects of song learning in songbirds. *Trends*
658 *ecol evol*, 20(3):143-149

659 Berg KS, Delgado S, Cortopassi KA, Beissinger SR, Bradbury JW (2012) Vertical
660 transmission of learned signatures in a wild parrot. *Proc Biol Sci* 279(1728):585-591

661 Bloomfield LL, Phillmore LS, Weisman RG, Sturdy CB (2005) Note types and coding in
662 parid vocalizations. III: The chick-a-dee call of the Carolina chickadee (*Poecile carolinensis*).
663 *Canad J Zool* 83(6):820-833

664 Bottoni L, Masin S, Lenti-Boero D (2009) Vowel-like sound structure in an African Grey
665 Parrot (*Psittacus erithacus*) vocal production. *Open Behav Sci J* 3(1)

666 Bradbury JW (2003) Vocal communication in wild parrots. In: DeWaal FBM, Tyack PL (eds)
667 *Animal social complexity: intelligence, culture, and individualized societies*. Harvard
668 University Press, Cambridge, MA, pp 293-316

669 Bradbury JW, Balsby TJ (2016) The functions of vocal learning in parrots. *Behav Ecol*
670 *Sociobiol* 70(3):293-312

671 Bradbury JW, Cortopassi KA, Clemmons JR, Kroodsma D (2001) Geographical variation in
672 the contact calls of orange-fronted parakeets. *Auk* 118(4):958-972

673 Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. *Nature*
674 417(6886):351-358

675 Brittan-Powell EF, Dooling RJ, Farabaugh SM (1997) Vocal development in budgerigars
 676 (*Melopsittacus undulatus*): Contact calls. J Comp Psychol 111(3):226

 677 Center for Conservation Bioacoustics. (2011). Raven Pro: Interactive Sound Analysis
 678 Software (Version 1.4) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology
 679 Available from <http://ravensoundsoftware.com/>

 680 Charif, RA, LM Strickman, and AM Waack. 2010. Raven Pro 1.4 User's Manual. The Cornell
 681 Lab of Ornithology, Ithaca, NY

 682 Clark CW, Marler P, Beeman K (1987) Quantitative analysis of animal vocal phonology: An
 683 application to swamp sparrow song. Ethology 76(2):101-115

 684 Cortopassi KA, Bradbury JW (2000) The comparison of harmonically rich sounds using
 685 spectrographic cross-correlation and principal coordinates analysis. Bioacoustics 11(2):89-
 686 127

 687 Cruickshank AJ, Gautier JP, Chappuis C (1993) Vocal mimicry in wild African grey parrots
 688 *Psittacus erithacus*. Ibis 135(3):293-299

 689 Dahlin CR, Young AM, Cordier B, Mundry R, Wright TF (2014) A test of multiple
 690 hypotheses for the function of call sharing in female budgerigars, *Melopsittacus undulatus*.
 691 Behav Ecol Sociobiol 68(1):145-161

 692 De Araújo CB, Marcondes-Machado LO, Vielliard JM (2011) Vocal repertoire of the yellow-
 693 faced parrot (*Alipiopsitta xanthops*). Wilson J. Ornithol. 123(3):603-608

 694 Dooling RJ (1980) Behavior and Psychophysics of Hearing in Birds. In: Popper AN, Fay RR
 695 (eds) Comparative Studies of Hearing in Vertebrates. Proceedings in Life Sciences. Springer,
 696 New York, NY, pp 261-288

697 Enggist-Dueblin P, Pfister U (2002) Cultural transmission of vocalizations in ravens, *Corvus*
 698 *corax*. Anim Behav 64(6):831-841

699 Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-
 700 type song culture in the zebra finch. Nature 459(7246):564-568

701 Fischer J, Hammerschmidt K, Cheney DL, Seyfarth RM (2001) Acoustic features of female
 702 chacma baboon barks. Ethology 107(1):33-54

703 Gardner TJ, Naef F, Nottebohm F (2005) Freedom and rules: the acquisition and
 704 reprogramming of a bird's learned song. Science 308(5724):1046-1049

705 Giret N, Albert A, Nagle L, Kreutzer M, Bovet D (2012) Context-related vocalizations in
 706 African grey parrots (*Psittacus erithacus*). Acta Ethol 15(1):39-46

707 Giret N, Monbureau M, Kreutzer M, Bovet D (2009) Conspecific discrimination in an object-
 708 choice task in African grey parrots (*Psittacus erithacus*). Behav Process 82(1):75-77

709 Giret N, Péron F, Lindová J, Tichotová L, Nagle L, Kreutzer M, ... Bovet D (2010)
 710 Referential learning of French and Czech labels in African grey parrots (*Psittacus erithacus*):
 711 Different methods yield contrasting results. Behav Process 85(2):90-98

712 Giret N, Roy P, Albert A, Pachet F, Kreutzer M, Bovet D (2011) Finding good acoustic
 713 features for parrot vocalizations: The feature generation approach. J Acoust Soc Am
 714 129(2):1089-1099

715 Hall WS, Cookson KK, Heaton JT, Roberts T, Shea SD, Brauth SE (1997) Audio - Vocal
 716 Learning in Budgerigars. Ann New York Acad Sci 807(1):352-367

717 Harrison L-MK, Cox MJ, Skaret G, Harcourt R (2015) The R package EchoviewR for
 718 automated processing of active acoustic data using Echoview. Front. Mar. Sci. 2:15 doi:
 719 10.3389/fmars.2015.00015

720 Hedwig D, Hammerschmidt K, Mundry R, Robbins MM, Boesch C (2014) Acoustic structure
 721 and variation in mountain and western gorilla close calls: a syntactic approach. Behaviour
 722 151(8):1091-1120

723 Hile AG, Striedter GF (2000) Call convergence within groups of female budgerigars
 724 (*Melopsittacus undulatus*). Ethology 106(12):1105-1114

725 Hile AG, Plummer TK, Striedter GF (2000) Male vocal imitation produces call convergence
 726 during pair-bonding in budgerigars (*Melopsittacus undulatus*). Anim Behav 59:1209-1218

727 Janik VM (1999) Pitfalls in the categorization of behaviour: A comparison of dolphin whistle
 728 classification methods. Anim Behav 57(1):133-143

729 Janik VM (2014) Cetacean vocal learning and communication. Curr Opin Neurobiol 28:60-65

730 Janik VM, Slater PJ (2000) The different roles of social learning in vocal communication.
 731 Anim Behav 60(1):1-11

732 Jarvis ED (2004) Learned birdsong and the neurobiology of human language. Ann New York
 733 Acad Sci 1016:749

734 Jones AE, ten Cate C, Bijleveld CC (2001) The interobserver reliability of scoring sonograms
 735 by eye: A study on methods, illustrated on zebra finch songs. Anim Behav 62(4): 791-801

736 Jürgens R, Hammerschmidt K, Fischer J (2011) Authentic and play-acted vocal emotion
 737 expressions reveal acoustic differences. Front Psychol 2:180.

738 Khanna H, Gaunt SLL, McCallum DA (1997) Digital spectrographic cross-correlation: Tests
 739 of sensitivity. Bioacoustics 7(3): 209-234

740 Knörnschild M (2014) Vocal production learning in bats. Curr Opin Neurobiol 28:80-85

741 Kuczaj SA (1983) Crib speech and language play. Springer, New York, NY

742 Larranaga A, Bielza C, Pongrácz P, Faragó T, Bálint A, Larranaga P (2015) Comparing
 743 supervised learning methods for classifying sex, age, context and individual Mudi dogs from
 744 barking. *Anim Cogn* 18(2):405-421

 745 Marler P (1970) Birdsong and Speech Development: Could There Be Parallels? *Am Sci*
 746 58:669-673

 747 May DL (2004) The vocal repertoire of Grey parrots (*Psittacus erithacus*) living in the Congo
 748 Basin. Dissertation, University of Arizona

 749 Molnár C, Kaplan F, Roy P, Pachet F, Pongrácz P, Dóka A, Miklósi Á (2008) Classification
 750 of dog barks: A machine learning approach. *Anim Cogn* 11(3): 389-400

 751 Montes-Medina AC, Salinas-Melgoza A, Renton K (2016) Contextual flexibility in the vocal
 752 repertoire of an Amazon parrot. *Front Zool* 13:40

 753 Mooney R (2009) Neurobiology of song learning. *Curr Opin Neurobiol* 19(6):654-660

 754 Naguib M, Hammerschmidt K, Wirth J (2001) Microgeographic variation, habitat effects and
 755 individual signature cues in calls of chiffchaffs *Phylloscopus collybita canarensis*. *Ethology*
 756 107(4):341-355

 757 Nowicki S, Searcy WA (2014) The evolution of vocal learning. *Curr Opin Neurobiol* 28:48-
 758 53

 759 Ohala JJ (1983) The origin of sound patterns in vocal tract constraints. In: MacNeilage PF
 760 (ed) The production of speech. Springer, New York, NY, pp. 189-216

 761 Okanoya K, Dooling RJ (1987) Hearing in passerine and psittacine birds: A comparative
 762 study of absolute and masked auditory thresholds. *J Comp Psychol* 101(1):7-15

763 Patterson DK, Pepperberg IM (1998) Acoustic and articulatory correlates of stop consonants
 764 in a parrot and a human subject. *J Acoust Soc Am* 103(4):2197-2215

765 Pepperberg IM (1999) The Alex studies: Cognitive and communicative abilities of Grey
 766 parrots. Harvard University Press, Cambridge, MA

767 Pepperberg IM (2002) Allospecific referential speech acquisition in grey parrots (*Psittacus*
 768 *erithacus*): Evidence for multiple levels of avian vocal imitation. In: Dautenhahn K, Nehaniv
 769 CL (eds) Imitation in animals and artifacts. MIT Press, Cambridge, MA, pp 109-131

770 Pepperberg IM (2010). Vocal learning in Grey parrots: A brief review of perception,
 771 production, and cross-species comparisons. *Brain Lang* 115(1):81-91

772 Pepperberg IM, McLaughlin MA (1996) Effect of avian–human joint attention in allospecific
 773 vocal learning by grey parrots (*Psittacus erithacus*). *J Comp Psychol* 110(3):286-297

774 Pérez-Espinosa H, Pérez-Martínez JM, Durán-Reynoso JÁ, Reyes-Meza V (2015) Automatic
 775 classification of context in induced barking. *Research in Computing Science* 100:63-74

776 Pfefferle D, Hammerschmidt K, Mundry R, Ruiz-Lambides AV, Fischer J, Widdig A (2016)
 777 Does the structure of female rhesus macaque coo calls reflect relatedness and/or familiarity?
 778 *PloS one* 11(8): e0161133

779 Policht R, Karadžos A, Frynta D (2011). Comparative analysis of long-range calls in equid
 780 stallions (Equidae): Are acoustic parameters related to social organization? *Afr Zool* 46(1):18-
 781 26

782 Policht R, Tomášová K, Holečková D, Frynta D (2008) The vocal repertoire in Northern
 783 white rhinoceros *Ceratotherium simum cottoni* as recorded in the last surviving herd.
 784 *Bioacoustics* 18(1):69-96

785 Price T, Ndiaye O, Hammerschmidt K, Fischer J (2014) Limited geographic variation in the
 786 acoustic structure of and responses to adult male alarm barks of African green monkeys.
 787 Behav Ecol Sociobiol 68(5):815-825

788 Reichmuth C, Casey C 2014 Vocal learning in seals, sea lions, and walruses. Curr.Opin.
 789 Neurobiol. 28:66-71

790 Salinas-Melgoza A, Wright TF (2012) Evidence for vocal learning and limited dispersal as
 791 dual mechanisms for dialect maintenance in a parrot. PloS one 7(11):e48667

792 Sewall KB (2009) Limited adult vocal learning maintains call dialects but permits pair-
 793 distinctive calls in red crossbills. Anim Behav 77(5):1303-1311

794 Schrader L, Hammerschmidt K (1997) Computer-aided analysis of acoustic parameters in
 795 animal vocalisations: A multi-parametric approach. Bioacoustics 7(4):247-265

796 Schwing R, Parsons S, Nelson XJ (2012) Vocal repertoire of the New Zealand kea parrot
 797 *Nestor notabilis*. Curr. Zool 58(5):727-740

798 Stark RE (1980) Stages of speech development in the first year of life. In: Yeni-Komshian G,
 799 Kavanagh JF, Ferguson CA (eds), Child phonology, vol. 1. Academic Press, New York, NY,
 800 pp 73-92

801 Stoeger AS, Manger P (2014) Vocal learning in elephants: neural bases and adaptive context.
 802 Curr Opin Neurobiol 28:101-107

803 Sueur J, Aubin T, Simonis C (2008). Seewave: a free modular tool for sound analysis and
 804 synthesis. Bioacoustics, 18:213-226

805 <http://www.tandfonline.com/doi/abs/10.1080/09524622.2008.9753600>.Tchernichovski O,
 806 Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated
 807 measurement of song similarity. Anim Behav 59(6):1167-1176

808 Todt D, Geberzahn N (2003) Age-dependent effects of song exposure: song crystallization sets
 809 a boundary between fast and delayed vocal imitation. *Anim Behav* 65:971-979
 810 <https://doi.org/10.1006/anbe.2003.2127>
 811 Wadewitz P, Hammerschmidt K, Battaglia D, Witt A, Wolf F, Fischer J (2015)
 812 Characterizing vocal repertoires—Hard vs. soft classification approaches. *PloS one* 10(4)
 813 Wanker R, Sugama Y, Prinage S (2005) Vocal labelling of family members in spectacled
 814 parrotlets, *Forpus conspicillatus*. *Anim Behav* 70(1):111-118
 815 Weisman RG, Njegovan MG, Williams MT, Cohen JS, Sturdy CB (2004) A behavior analysis
 816 of absolute pitch: Sex, experience, and species. *Behav Process* 66(3):289-307
 817 Wright TF, Cortopassi KA, Bradbury JW, Dooling RJ (2003) Hearing and vocalizations in the
 818 orange-fronted conure (*Aratinga canicularis*). *J Comp Psychol* 117(1):87
 819 Zdenek CN, Heinsohn R, Langmore NE (2015) Vocal complexity in the palm cockatoo
 820 (*Probosciger aterrimus*). *Bioacoustics* 24(3):253-267
 821

822 **Table 1:** Call categories identified with the visual-acoustic (VA) method and confirmed by
823 computer-aided spectrographic cross-correlation (SPCC) and multiparametric statistical (MS)
824 method in a captive group of grey parrots

VA Category	Count	SPCC	MS	Inner consit. (SPCC)	Higher outgroup correlations (SPCC)	Final Category
K1	238	✓	✓	0.392		K1
K2	191	✓	✓	0.200		K2
K3	179	✓	m ^{K32}	0.210		K3
K4	156	m ^{K38,K63}	m ^{K38}	0.616		K4 + K38
K5	130	m ^{K28,K30,K58}	m ^{K30,K35,K39}	0.518	0.543 ^{K30}	K5 + K30
K6	107	m ^{K11}	✓	0.224		K6
K7	92	✓	✓	0.369		K7
K8	92	m ^{K40}	m ^{K40}	0.517		K8 + K40
K9	89	✓	✓	0.290		K9
K10	85	✓	✓	0.255		K10
K11	80	m ^{K6}	✓	<u>0.167</u>	0.182 ^{K6}	K11
K12	78	m ^{K45}	✓	0.374		K12
K13	76	✓	m ^{K22,K29}	0.292		K13
K14	70	✓	✓	0.223		K14
K15	57	✓	✓	0.659		K15
K16	54	✓	✓	0.490		K16
K17	53	✓	✓	0.337		K17
K18	51	✓	✓	0.433		K18
K19	50	m ^{K21,K24}	m ^{K21,K24}	0.437		K19 + K21 + K24
K20	50	✗	✗	<u>0.151</u>	0.170 ^{K63*}	✗
K21	48	m ^{K19,K24}	m ^{K19,K24}	0.221	0.234 ^{K14} 0.316 ^{K19}	K19 + K21 + K24
K22	45	m ^{K29}	m ^{K13,K29}	0.360		K22 + K29
K23	44	m ^{K67}	✓	0.635		K23
K24	43	m ^{K19,K21}	m ^{K19,K21}	0.338	0.390 ^{K19}	K19 + K21 + K24
K25	41	✓	✓	0.302		K25
K26	40	m ^{K56,K68,K69}	m ^{K68}	<u>0.165</u>	0.180 ^{K56} 0.190 ^{K68} 0.285 ^{K69}	K26 + K68
K27	38	m ^{K64}	m ^{K66}	0.488		K27
K28	38	m ^{K5,K30,K58}	✓	0.442	0.474 ^{K23} 0.561 ^{K30} 0.452 ^{K67}	K28
K29	38	m ^{K22}	m ^{K13,K22}	0.268	0.269 ^{K22}	K22 + K29
K30	35	m ^{K5,K28,K58}	m ^{K5,K35,K39}	0.465	0.543 ^{K5} 0.494 ^{K23} 0.494 ^{K28}	K5 + K30
K31	33	✓	✓	0.360		K31
K32	28	✓	m ^{K3}	0.367		K32
K33	28	✓	✓	0.317		K33
K34	28	m ^{K52}	✓	0.421		K34
K35	28	m ^{K39}	m ^{K5,K30,K39}	0.473		K35 + K39
K36	27	✓	✓	0.325		K36
K37	26	✓	✓	0.206		K37
K38	25	m ^{K4,K63}	m ^{K4}	0.421	0.498 ^{K4}	K4 + K38
K39	25	m ^{K35}	m ^{K5,K30,K35}	0.405	0.456 ^{K5} 0.440 ^{K23} 0.422 ^{K28} 0.414 ^{K30} 0.531 ^{K35}	K35 + K39
K40	24	m ^{K8}	m ^{K8}	0.369	0.420 ^{K8}	K8 + K40
K41	23	✓	✓	0.301		K41
K42	22	✓	✓	0.297		K42
K43	21	✓	✓	0.298		K43
K44	21	✓	✓	0.451		K44
K45	21	m ^{K12}	✓	<u>0.175</u>	0.212 ^{K12}	K45
K46	21	m ^{K51}	✓	0.488		K46
K47	20	✗	✗	<u>0.125</u>		✗
K48	19	✓	✓	0.281		K48
K49	18	✓	✓	0.630		K49
K50	17	✓	✓	0.546		K50
K51	17	m ^{K46}	✓	0.335	0.383 ^{K46}	K51
K52	15	m ^{K34}	✓	<u>0.156</u>	0.258 ^{K34} 0.243 ^{K36}	K52
K53	13	m ^{K57,K70}	✓	0.620	0.622 ^{K70}	K53
K54	13	✓	✓	0.419		K54
K55	12	✓	✓	0.409		K55
K56	12	m ^{K26,K68,K69}	✓	0.221	0.244 ^{K69}	K56
K57	11	m ^{K53,K70}	m ^{K70}	0.633	0.676 ^{K49}	K57 + K70
K58	11	m ^{K5,K28,K30}	✓	0.466	0.490 ^{K5}	K58
K59	11	✓	✓	0.481		K59
K60	9	✓	✓	0.453		K60
K61	9	✓	✓	0.507		K61
K62	9	✓	✓	0.313		K62
K63	8	m ^{K4,K38}	✓	0.363	0.438 ^{K38} 0.397 ^{K4}	K63
K64	8	m ^{K27}	✓	0.383	0.403 ^{K27}	K64
K65	7	✓	✓	0.469		K65
K66	6	✓	m ^{K27}	0.449		K66
K67	6	m ^{K23}	✓	0.305	0.397 ^{K5} 0.476 ^{K23} 0.414 ^{K28} 0.428 ^{K30} 0.348 ^{K35} 0.338 ^{K39}	K67
K68	5	m ^{K26,K56,K69}	m ^{K26}	0.220	0.404 ^{K69}	K26 + K68
K69	4	m ^{K26,K56,K68}	✓	0.578		K69
K70	3	m ^{K53,K57}	m ^{K57}	0.695		K57 + K70

825

- 826 Note:
- 827 ✓ ... category confirmed by the respective computer-aided method;
- 828 m ... category merged (with category or categories shown in the upper index);
- 829 ✕ ... category discarded;
- 830 VA category... name of a category determined by the VA method;
- 831 Count... the number of times the category was used by all parrots;
- 832 SPCC... confirmation, merging, or discarding indicated by the SPCC method;
- 833 MS... confirmation, merging or discarding indicated by the MS method;
- 834 Inner consistency (SPCC method) ... calculated as the average of values from the peak correlation matrix for a category, estimates the
- 835 similarity of calls within a category;
- 836 Higher outgroup correlations (SPCC method) ... cases when outgroup correlation was higher than inner consistency;
- 837 Final category ... categories proposed after performing both computer-aided methods.

838

839 **Table 2a:** The use of call categories by individual parrots

Parrot	Number	%
Durosimo	57	81.43%
Jarina	52	74.29%
Jaro	60	85.71%
Juruba	62	88.57%

- 840 Note: The number of categories used by a parrot (out of the 70 categories determined by the VA method).

841

842

843

Table 2b: Comparison of use of categories by individual parrots

	Durosimo	Jarina	Jaro	Juruba
Durosimo		77.19%	91.23%	87.72%
Jarina	84.62%		82.69%	94.23%
Jaro	86.67%	71.67%		85.00%
Juruba	81.97%	80.33%	83.61%	

Note: Percentage of shared categories between individual parrots determined by the VA method.

Table 3: Confirmative classification of calls into categories by the multiparametric statistical method.

Cluster	Category	Correct classification	Validated	Merged/ Discarded	Classification after merging	Validated after merging
C1	K1, K2, K15	91%	88%	-	-	-
C2	K7, K9, K13, K22, K29, K42, K55, K60, K61	76%	70%	m ^{K13, K22, K29}	85%	80%
C3	K5, K23, K28, K30, K35, K39, K67	57%	53%	m ^{K5, K30, K35, K39}	79%	76%
C4	K14, K18, K19, K21, K24	63%	56%	m ^{K19, K21, K24}	82%	81%
C5	K3, K32, K33, K48	74%	65%	m ^{K3, K32}	79%	76%
C6	K4, K20, K38, K63	68%	68%	m ^{K4, K38} ; d ^{K20}	86%	86%
C7	K12, K34, K36, K45, K47, K52	76%	70%	d ^{K47}	84%	80%
C8	K6, K11	77%	77%	-	-	-
C9	K8, K40	83%	79%	m ^{K8, K40}	94%	92%
C10	K26, K37, K44, K58, K68, K69	66%	65%	m ^{K26, K68}	73%	71%
C11	K16, K31, K59	83%	79%	-	-	-
C12	K25, K27, K64, K66	71%	58%	m ^{K27, K66}	77%	77%
C13	K10	78%	77%	-	-	-
C14	K17, K43	88%	84%	-	-	-
C15	K41, K50, K54, K62	94%	90%	-	-	-
C16	K49, K53, K57, K70	89%	82%	m ^{K57, K70}	91%	87%
C17	K46, K51, K65	97%	95%	-	-	-
C18	K58	92%	91%	-	-	-

Note: Clusters established by the VA method. Category column shows categories which form the cluster in question. Correct classification marks the percentage of calls correctly classified by the DF analysis into categories belonging to the cluster in question. Merged/discarded column shows merged or discarded categories. Clusters with at least one merged category were reanalyzed by the DF analysis.

855 **Fig. 1:** Recording Plot

856 Note: Marks denote recording sessions

857 **Fig. 2:** Juruba's repertoire

858 **Fig. 3:** Zoé's repertoire (Giret et al. 2012)

859 **Fig. 4:** Call types performed by both Juruba and Zoé

860 **Table S1:** T-test results

861 Note: Numbers are t-stats, those marked with * are statistically significant.

862