Evidence for vocal signatures and voice-prints in a wild parrot

Simeon Q. Smeele[1,2,3], Juan Carlos Senar[4], Lucy M. Aplin[1,5,6,*], and Mary Brooke McElreath[1,2,*]

[1] *Cognitive & Cultural Ecology Research Group, Max Planck Institute of Animal Behavior, Radolfzell, Germany*
[2] *Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*
[3] *Department of Biology, University of Konstanz, Konstanz, Germany*
[4] *Museu de Ciències Naturals de Barcelona, Barcelona, Spain*
[5] *Department of Evolutionary Biology and Environmental Science, University of Zurich, Zurich, Switzerland*
[6] *Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, Australia*
[*] *Joint senior author*

January 2023

ORCID: SQS, 0000-0003-1001-6615; JCS, 0000-0001-9955-3892; LMA, 0000-0001-5367-826X; MBM, 0000-0003-3206-5485

Subject areas: behaviour

Keywords: vocal signature, monk parakeet, parrots, Bayesian statistics, individual recognition

Author for correspondence: Simeon Q. Smeele, ssmeele@ab.mpg.de

Abstract

In humans, identity is partly encoded in a voice-print that is carried across multiple vocalisations. Other species of vocal learners also signal vocal identity in calls, for example as shown in the contact call of parrots. However, it remains unclear to what extent other call types in parrots are individually distinct, and whether there is an analogous voice-print across calls. Here we test if an individual signal is also present in other call types, how stable this signal is, and if parrots exhibit voice-prints across call types. We recorded 5599 vocalisations from 229 individually-marked monk parakeets (*Myiopsitta monachus*) over a two year period in Barcelona, Spain. We find evidence for an individual signal in two out of five call types. We further show that while contact calls are individually distinct, they are more variable within individuals than previously assumed, changing over short time scales. Finally, we provide evidence for voice-prints across multiple call types, with a discriminant function being able to predict caller identity across call types. This suggests that monk parakeets may be able to use vocal cues to recognise conspecifics, even across vocalisation types and without necessarily needing active vocal signals of identity.

Introduction

Individual recognition and signalling of individual identity can play an important role in social interactions and decision-making. Examples of how individuals can benefit from individual recognition are wide-ranging, and include helping relatives (Russell and Hatchwell 2001), remembering reliable cooperators (Boesch 1994) and strategically directing aggression (Hobson, Mønster, and DeDeo 2021). For the individual that is recognised, signalling identity is beneficial if the benefits associated with incurring affiliative behaviour outweigh potential costs associated with misidentification (Johnstone 1997). While it sometimes also pays to hide identity (Tibbetts and Dale 2007; Johnstone 1997; Carlson, Kelly, and Couzin 2020), in most cases, the benefits of broadcasting identity likely outweigh the potential costs. In fission-fusion societies, for instance, signalling identity may allow individuals to preferentially reassociate with a subset of the population when confronted with a large number of potential interaction partners (Kummer 2017; Aureli, Schaffner, and Schino 2022). Early human societies were fission-fusion based and likely heavily dependent on cooperation between individuals (Migliano and Vinicius 2022); perhaps not surprisingly, the human face has evolved to allow for maximum individual distinctiveness (Sheehan and Nachman 2014).

Across species, individual identity has been found to be conveyed through multiple potential sensory modalities, including olfactory, acoustic or visual cues. For example, several social wasps display distinctive facial features (Tibbetts 2004). However, while visual or olfactory distinctiveness is useful during close interactions, they are likely less effective across longer distances or in low visibility environments such as tropical forests or turbid waters. Vocal signals are much better suited for these situations, and vocal broadcasting of identity has been found across a wide range of taxonomic groups, ranging from finches (Mundinger 1970) to dolphins (Janik and Sayigh 2013). These species often have one call type that is very stereotyped within individuals, with enough structural complexity to allow for many unique variants. For example, bottle-nosed dolphins (*Tursiops truncatus*) produce a very stereotyped signature whistle when out of visual contact, where the individual signature is encoded in the frequency modulation pattern, or in other words how the frequency goes up and down (Janik, Sayigh, and Wells 2006). Individuals predominantly produce 'their' signature whistle, and the duration combined with the frequency modulation allows for many unique patterns. Frequency modulation patterns are also much more resistant to attenuation than other features and can therefore be recognised over larger distances.

While a single vocal signal to broadcast identity is useful, individuals will often produce multiple call types and could benefit from being recognised across these calls. Three potential solutions to the need to be recognised in multiple call types are possible. The first is making each call type individually distinct. Such a strategy has been shown in a variety of bird species (Elie and Theunissen 2018; Charrier et al. 2001; Mäkelin et al. 2021), bats (Prat, Taub, and Yovel 2016) and some primate species (Keenan et al. 2020; Bouchet et al. 2012; Salmi, Hammerschmidt, and Doran-Sheehy 2014). However, maintaining multiple signals of identity is cognitively demanding for signallers and receivers to remember; consequently, this strategy is likely constrained to species with either very small vocal repertoires or small group sizes. The second solution is to combine a single identity call with the other call types in a sequence (Rauber, Kranstauber, and Manser 2020). The cognitive demands of this strategy are much lower, and if flexibly deployed, it potentially allows individuals to signal identity in contexts where recognition is beneficial and hide identity in other contexts. However, it increases the complexity and potential cost of vocal production, as all individually distinct vocalisations now involve at least two elements. The third solution is to evolve a recognisable voice-print across call types. This can be achieved via the specific morphology of the vocal production organ, leaving a unique and recognisable cue on all vocalisations that is consistent within individuals across call types but variable across individuals. This last solution is well suited for species that continuously modify the vocalisations they produce. It should be noted that such a voice-print differs from a vocal signal in that it is likely not actively produced, but is a by-product of the vocal tract. To distinguish between these types of vocal signals, throughout this study we use the term 'individual signal' to denote actively produced uniqueness within call types and 'voice-print' to denote the emergent individual signal resulting from vocal tract morphology.

The best known example of such a voice-print is in humans. Humans have a complex communication system with an almost unlimited number of sounds that can be produced, rendering it unfeasible to include identity calls in combination with secondary utterances. Yet despite this flexible production, the human vocal tract leaves an individually distinct cue in the timbre of the voice, allowing speakers to be recognised across most utterances (Mathias, Kriegstein, et al. 2014). Yet potential for such a voice-print in non-human animals has received surprisingly little attention. Thus far, voice-prints have only been shown in the mating calls of red deer stags (*Cervus elaphus*), where Reby et al. (2006) used mel frequency cepstral coefficients combined with a hidden Markov chain model to find that 63% of roars and barks could be correctly assigned to seven individuals. Notably, this study used relatively few call types and individuals of a fixed repertoire species. To our knowledge there has been no study showing voice-prints across call types in a non-human vocal learner with a large and flexible vocal repertoire. This is despite the fact that these species would benefit most from such an individual vocal recognition mechanism.

Parrots are open-ended vocal production learners that often exhibit large and flexible vocal repertoires (Bradbury and Balsby 2016; Wright and Dahlin 2018). In this group, most research focus has been on contact calls. These contact call are likely socially learned in early stages of development (Berg, Delgado, Cortopassi, et al. 2012; Teixeira et al. 2021) and are generally assumed to broadcast identity (Wright 1996; Smith-Vidaurre, Araya-Salas, and Wright 2020). Some species appear to actively modify their contact call over periods of weeks to converge with pairs or with flock mates (Dahlin et al. 2014; Scarl and Bradbury 2009), and there is even evidence for rapid convergence within vocal exchanges (Wright, Hara, et al. 2015; Balsby and Bradbury 2009; Balsby, Momberg, and Dabelsteen 2012). Despite this flexibility, some species have a stable individual signal in their contact call, at least within the time period of focus (Thomsen, Balsby, and Dabelsteen 2013; Smith-Vidaurre, Araya-Salas, and Wright 2020; Berg, Delgado, Okawa, et al. 2011). Other species have a stable group level signature in their call that also appears to persist over long periods of time. For example yellow-naped amazons, *Amazona auropalliata*, have dialects that are virtually unchanged throughout a period of 11 years in some locations (Wright, Dahlin, and Salinas-Melgoza 2008). However, it is not known how much of an individual signal exists in call types other than the contact call for adult parrots (but see Wein et al. (2021)), whether this is stable over time, or even if vocal distinctiveness carries across call types as a voice-print.

In our study we addressed these questions in monk parakeets (*Myiopsitta monachus*), a communal nesting parrot with a large native and invasive range. Monk parakeets are popular pets with good vocal imitative abilities and like all parrots, are life-long vocal learners. Their contact calls have been extensively studied (Martella and Bucher 1990; Buhrman-Deever, Rappaport, and Bradbury 2007; Smith-Vidaurre, Araya-Salas, and Wright 2020; Smith-Vidaurre, Perez-Marrufo, and Wright 2021; Smeele et al. 2022), with these studies suggesting that monk parakeet contact calls contain an individual signal (Smith-Vidaurre, Araya-Salas, and Wright 2020). In their invasive range, they also appear to exhibit geographically distinct dialects in contact calls (Buhrman-Deever, Rappaport, and Bradbury 2007; Smeele et al. 2022), although this is much less pronounced in their native range (Smith-Vidaurre, Araya-Salas, and Wright 2020). However, it should be noted that no study has recorded vocalisations from a large set of individually-marked birds, or extended this analysis to other call types.

We recorded 229 wild, individually-marked monk parakeet in Barcelona, Spain over a period of two months across two consecutive years, and we manually categorised calls into 12 call types, five of which were used for further analysis. First, we measured similarity between calls within the same call type and analysed the results with a Bayesian multilevel model to test how much individual signal exists in the most common monk parakeet call types and how stable these signals are over time. Second, we tested how much individual signal exists across call types by training the model on one set of call types and predicting on another set of call types. Based on previous work we predicted high levels of individual signal in contact calls and lower levels in other call types. Additionally, we predicted a stable signal over a month long period with reduction in similarity over years. Finally, if monk parakeets exhibit a voice-print in their vocalisations, we predicted that calls could be assigned to individuals across call types.

Methods

Study System

We studied monk parakeets (*Myiopsitta monachus*) in Parc de la Ciutadella and surrounding areas in Barcelona, Spain, where they have been reported as an invasive species since the late 1970s (Batllori and Nos 1985). Parc de la Ciutadella, Promenade Passeig de Lluís Companys and Zoo de Barcelona form a continuous habitat of grass and asphalt with multiple tree species in which monk parakeets nest and forage. They build complex stick nests in trees and other structures, often building new nest chambers on top of already existing nest structures (Eberhard 1998), creating colonies of birds living in close proximity.

Since May 2002, adults and juveniles have been regularly captured and marked using a walk-in trap on Museu de Ciències Naturals de Barcelona, while fledglings have been marked directly at their nests (Senar, Carrillo-Ortiz, and Arroyo 2012). Birds are ringed with unique leg-bands and fitted with neck collars displaying unique combinations of letters and digits. This effort has resulted in over 3,000 ringed birds since May 2002, of which 300-400 are recaptured/sighted each year. In November 2021, to increase the number of marked birds in the population for this study, we captured and tagged an additional 59 adults and juveniles at their nests, trapping individuals at night with hand nets. All birds were ringed with special permission EPI 7/2015 (01529/1498/2015) from Direcció General del Medi Natural i Biodiversitat, Generalitat de Catalunya, and with authorization to JCS for animal handling for research purposes from Servei de Protecció de la Fauna, Flora i Animal de Companyia (001501-0402.2009).

Data Collection

Vocalisations were recorded from marked individuals in two years between 27.10.20 - 19.11.20 and 31.10.21 - 30.11.21 using a Sennheiser K6/ME67 shotgun microphone and Sony PCM D100 recorder from a distance ranging between one and 20 meters. The IDs and behaviours of focal individuals, the behaviours of close-by individuals and the general contexts of the vocalisations were verbally annotated. Some recordings were also videotaped and IDs were transcribed afterwards.

In addition, we mapped all nests in the recording area using Gaia GPS on several Android cellphones. Errors were manually corrected to less than 10 meters. In order to determine nest occupancy, we monitored nests multiple times throughout the day until an individual was observed inside the nest at least three times. Individuals were assigned to a nest entry if they were seen at least once inside one of the nest entrances. If they were sighted at multiple nests, they were assigned to the nest where they were most often sighted. If no birds were observed at a nest, we continued to monitor the nest daily for the duration of the recording period.

Data Processing

All calls with fundamental frequencies clearly distinguishable from background noise and with no overlapping sounds were selected in Raven Lite (K. Lisa Yang Center for Conservation Bioacoustics 2016). Calls were then manually assigned to 12 broad call types based on structural similarity, five of which were used for further analysis. These were: 1) contact call - a frequency modulated call with at least three infliction points, 2) *tja* call - a tonal call with a single rising frequency modulation, 3) *trruup* call - a combination of amplitude modulated introduction (similar to alarm calls) with a tonal ending (similar to the *tja* call), 4) alarm - an amplitude modulate call with at least four ’notes’ and clear harmonics, predominantly used in distress situations, and 5) growl - an amplitude modulate call with at least four ’notes’ and no clear harmonics, predominantly used in social interactions (see supplemental materials for example spectrograms of each call type).

We used four methods to measure similarity between calls: dynamic time-warping (DTW, (Giorgino 2009)), spectrographic cross correlation (SPCC, (Clark, Marler, and Beeman 1987)), spectrographic analysis (SPECAN, specified in the supplemental materials) and mel frequency cepstral coeffienct cross correlation (MF4C, specified in the supplemental materials). We present the results of SPCC in the main text, since SPCC could be run on all call types, is the most used method in previous work and other methods gave similar results. The results of all other methods are presented in the supplementary materials. SPCC consists of sliding two spectrograms over each other and calculating the sum of the difference between each pixel per sliding window. The distance at maximal overlap, or in other words the minimal distance, between calls is then used as a measure of acoustic distance between two calls (see Figure 1A for a schematic overview). We implemented our own function for SPCC to control for as much background noise as possible (see supplemental materials for details).

Statistical Analysis

The first aim of this study was to determine whether call types contained an individual signal. Three of our methods (DTW, SPCC and MF4C) produce similarity matrices rather than single or multiple measures per call. The analysis of such a matrix is challenging, since most conventional models are designed for multivariate data sets. To estimate similarity between calls coming from the same individual compared to calls coming from different individuals, we used a Bayesian model that is structurally similar to the social relationships model (Kenny and La Voie 1984). Data points are weights of dyadic relationships, or in our case the acoustic distance between two calls. This model allows for the control of all important potential confounds including repeated and unbalanced sampling per individual, per recording and repeated comparisons per call (see supplemental materials for the full model definition). To visualise similarity between calls coming from the same versus different individuals, we computed the posterior contrast between the predicted acoustic distance between calls from two different individuals and between two different recordings of the same individual. To visualise similarity between calls from the same recording session, we computed the posterior contrast between calls from two different individuals and compared that to posterior contrasts between calls from the same individual and same recording.

The second aim was to test how stable the individual signals were across time. We tested this across three scales: within a recording, across days and across years. We only used acoustic distances between calls from the same individual. We then modeled the acoustic distance as a function of time separating the two calls with a Bayesian multilevel model (see supplemental materials for the full model definition). For the first model we included time on the log-scale. For the latter two models we only included acoustic distance between calls coming from different recordings and time was measured as days between recordings and same of different year respectively.

Third, to assess how recognisable individuals were across call types we ran multiple permuted discriminant function analyses (pDFAs) on the mel frequency cepstral coefficients (MFCC) summary statistics (mean and standard deviation). We chose to write our own function to run pDFA so we could choose vocalisations from different recordings for the training and test sets, balance these data sets and compare the resulting scores to scores from a randomised data set. This function was based on the work done by (Mundry and Sommer 2007). To test how reliable pDFAs could score individual identity within a call type, we first trained and tested a pDFA on contact calls. To test how much information was available across broad call type categories we trained a pDFA on amplitude modulated calls and then tested on tonal calls and vice versa (see Figure 3A for a schematic overview). For all pDFAs we report the distribution of differences between the trained score and the randomised score. To test if the model learned features related to sex or background noise we reran the procedure on calls from a single sex from Promenade Passeig de Lluís Companys, which is generally more noisy and also reran the procedure where labels were restricted to be randomised within location (Promenade Passeig de Lluís Companys and Parc de la Ciutadella). Throughout the text we use pDFA to refer to a full set of permuted discriminant function analyses, DFA to refer to a single run of discriminant function analyses and LDA to refer to the linear discriminant function itself.

Results

In total, we recorded 5599 calls across 229 individually marked birds over the two years of data collection, 3242 in year 1, and 2357 in year 2. Our manually sorting lead to 3203 contact calls, 185 *tja* calls, 265 *trruup* calls, 249 alarm calls and 364 growls. We then asked whether the five call types were individually distinctive. As expected from previous studies (Smith-Vidaurre, Araya-Salas, and Wright 2020), we found a reliable individual signal for the contact call (see Figure 1B). The *trruup* call contained an equally strong individual signal. There was some evidence for an individual signal in alarm calls, but this was relatively weaker, with the contrast overlapping with zero. Finally, for the *tja* and growls there was no evidence for an individual signal.

Additionally, we found evidence in all call types for short-term temporal variability, with calls from the same recording sounding more similar than calls coming from two different recordings. For all calls other than the growl there was also an increase in acoustic distance with time throughout a recording (see Figure 1B). In other words, calls coming right after each other were more similar than calls spaced further apart in the recording. For the *trruup* call, alarm call and growl acoustic distance also increased with days between recordings. However, at the largest time scale this temporal variability disappeared, with individual signal stable between years and calls not more similar within year than across (see supplemental materials).

To visualise if birds gradually changed their calls throughout a recording or gave clustered sequences of very similar calls, we ran principle coordinate analysis on a subset of the recordings and visualised the calls in sequence (see Figure 2). For most recordings and call types, calls coming from the same ten second period (same colour) were spaced more closely than calls coming from different ten second periods. For some recordings calls were slightly clustered, but there was no clear pattern of directional change or clustering.

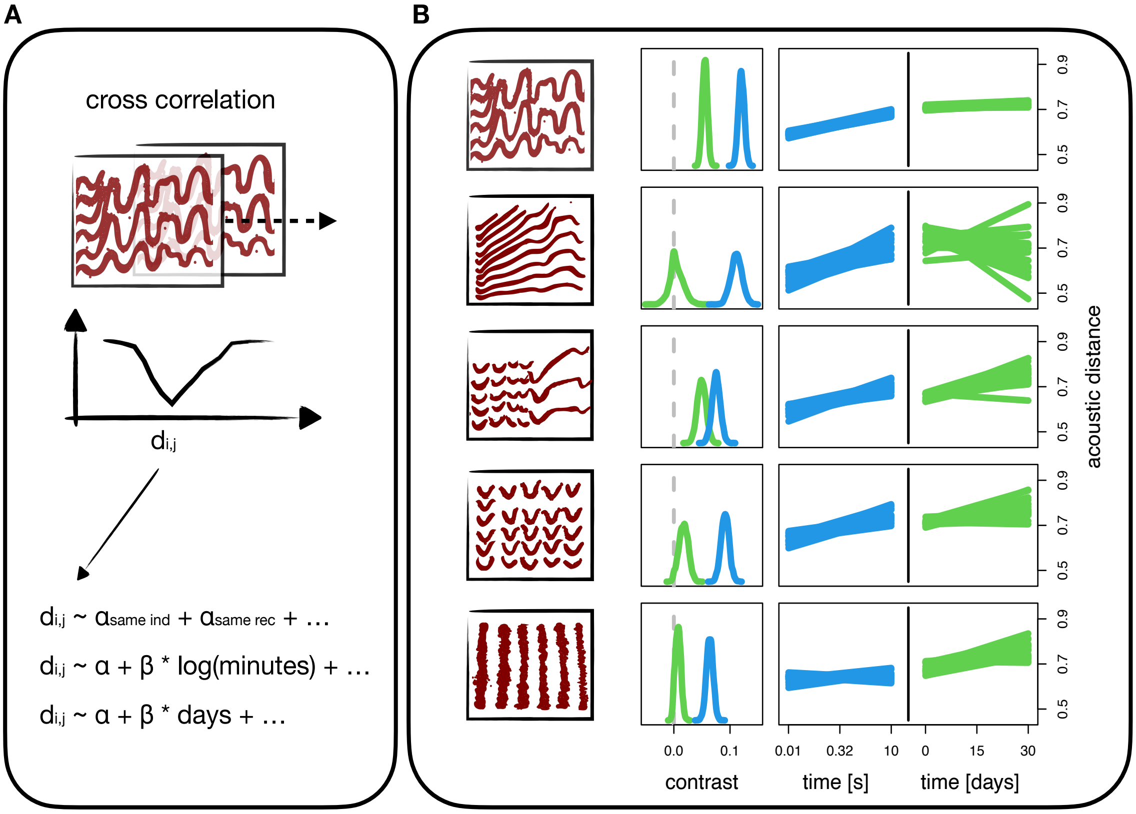

Figure 1: *Workflow and results for spectrographic cross correlation. (A) Schematic overview of the analysis pipeline. (B) Model results for contact calls,* tja *calls,* trruup *calls, alarm calls and growls (ttb). Blue density plots are the posterior contrast between the similarity of calls from different individuals versus the same individual and same recording. Green density plots are the posterior contrast between the similarity of calls from different individuals vs from the same individual but different recordings. Blue lines are 20 samples from the posterior prediction of acoustic distance throughout time within a recording. Green lines are 20 samples from the posterior prediction of acoustic distance throughout days between a recording.*

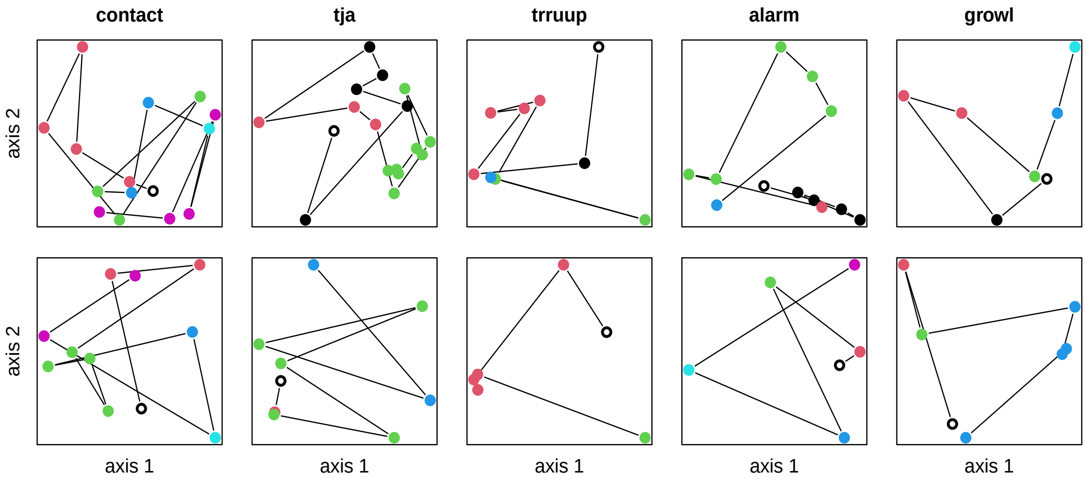

Figure 2: *Principle coordinate (PCO) representation of call similarity over ten recordings. Dots represent individual calls from two recordings (rows) per call type (columns). PCO space was constructed based on the spectrographic cross correlation distance matrix between calls within that recording. Dots are coloured based on ten second intervals. Dots are connected with lines to show order starting at the black with white dot. Calls close in time (same colour) are generally clustered, but are otherwise distributed randomly. In other words, calls are more similar when close in time, but birds do not slowly change the call over a longer time-span.*

We then used multiple permuted discriminant function analyses (pDFAs) on the mel frequency cepstral coefficients (MFCC) summary statistics (mean and standard deviation pDFAs to test whether DFAs trained on a subset of calls were able to successfully predict caller identity when presented with new calls. First, and as expected, results from the pDFAs further added to the evidence that contact calls contained an individual signal, with the trained DFA was on average 35% more successful in predicting identity than a randomised DFA (for the full distribution see supplemental materials). However, we also found evidence that calls contain general individualised features that were maintained across call types. A pDFA with amplitude modulated calls as training data and tonal calls as testing data or vice versa achieved a score of 15% and 9% more successful respectively than random (see Figure 3B for the results of a pDFA trained on amplitude modulated calls and tested on tonal calls). The trained DFA outperformed the random DFA in all iterations of the model.

While we did our best to select calls with no overlapping features or background noise, it is possible that our analysis was still detecting features that were more likely to occur in calls of particular individuals. Alternatively, individuals might have called in characteristic way in particular locations, creating a false signal in the data. To remove these potential biases, we re-ran our analysis within a single sex and location. In this case, only the pDFA trained on tonal calls and tested on amplitude modulated calls from females performed better than random. As this might be an effect of the greatly reduced dataset, we then re-ran our analysis with the full dataset, but restricting randomisation to only within location. In this case, the trained pDFA performed much better than chance, but overlap with zero increased to 6-7% (see supplemental materials for the full results).

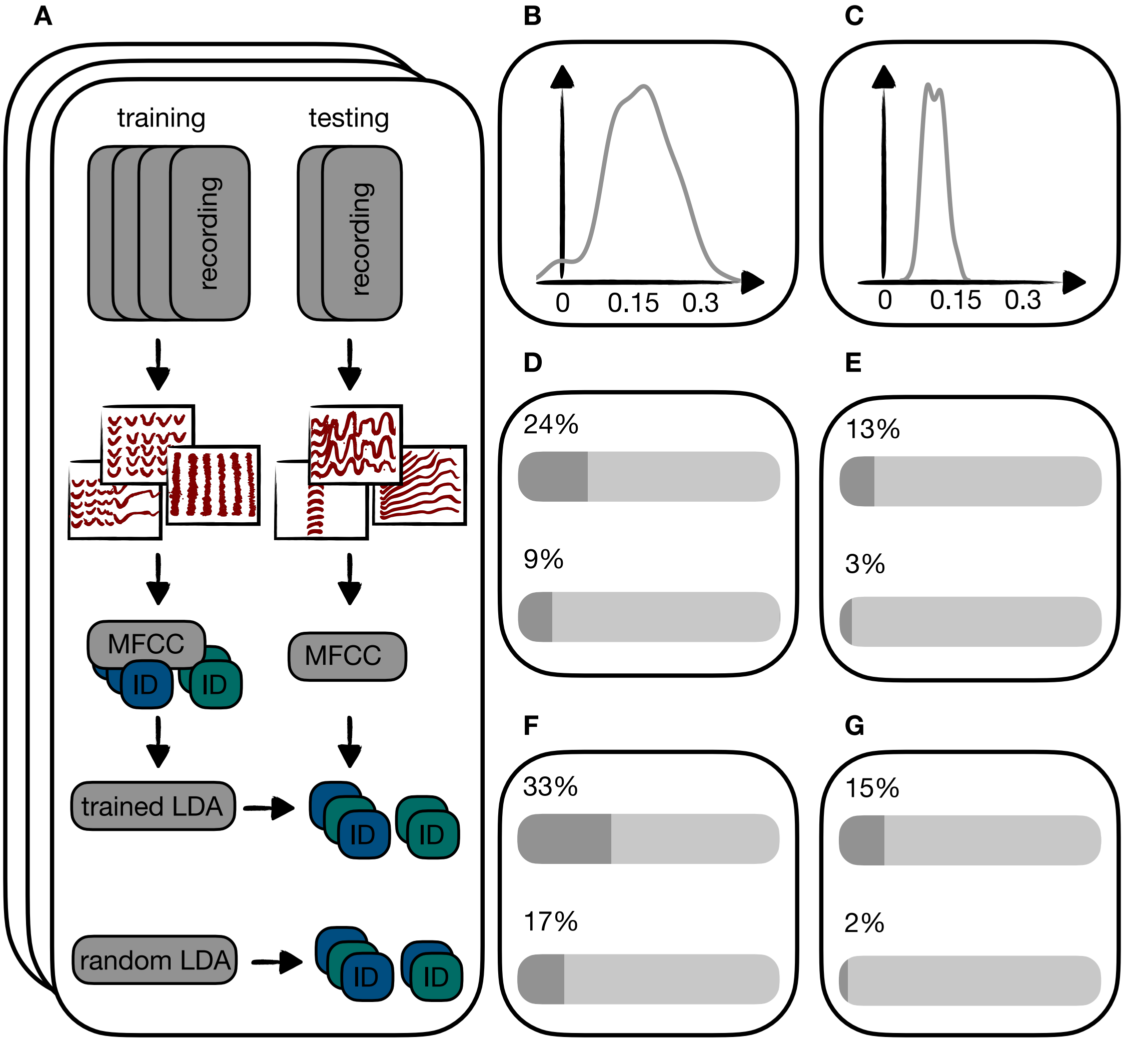

Figure 3: *Permuted linear discriminant function analysis across call types. (A) Workflow: for each iteration recordings are split in a training and a testing set. A linear discriminant classifier is trained on mel frequency cepstral coefficients from amplitude modulated calls from the training set. The correct classification percentage is calculated for the tonal calls from the testing set. (B-C) Density plot of the difference between the trained and randomised score for the pDFA trained on amplitude modulated calls and tested on tonal calls (B); and trained and tested on all calls (C). (D-G) Average correct classification percentage for the trained and randomised linear discriminant function trained on amplitude modulated calls and tested on tonal calls (D,F); and trained and tested on all calls (E,G); permuted within location (F,G).*

Discussion

Many species are likely to benefit from individual recognition. In many species of birds, this is thought to most likely occur through individually distinct vocalisations. Yet how this is achieved in species with open-ended vocal production learning, and in parrots in particular has been understudied (but see Thomsen, Balsby, and Dabelsteen (2013)). By recording vocalisations in individually-marked wild monk parakeets across one month and over two years, we reveal multiple insights into the vocal production of this parrot. First, we show that multiple call types given by monk parakeets contain an individual signal, and this signal is relatively stable over time, persisting within and between years (see Figure 1 and supplemental materials). Second, we show that despite this individual signal calls are not stereotyped, but are highly variable over short time scales. Within the same recording calls are generally more similar than calls from different recordings, and even within a recording calls close in time are more similar. Third, we tested if individual identity was distinguishable across call types. We used mel frequency cepstral coefficients (MFCC), training a permuted discriminant function (pDFA) on one set of call types and testing on another set of call types, doing so across recordings to make sure background noise could not be ’learned’ by the model. Our results suggest monk parakeets have a voice-print that exists across structurally different call types, although the strength of evidence varied across call types and analyses. To our knowledge this is the first evidence for the detection of voice-prints in a non-human vocal learner.

The ability to recognise individuals from their vocalisations should be highly advantageous in species with social systems like monk parakeets, where individuals may encounter many potential association partners during fission-fusion foraging dynamics. Previous studies have demonstrated individual signals in the contact calls of monk parakeets (Smith-Vidaurre, Araya-Salas, and Wright 2020), as well as in contact calls from other parrot species (Thomsen, Balsby, and Dabelsteen 2013; Balsby and Adams 2011; Farabaugh, Linzenbold, and Dooling 1994; Berg, Delgado, Okawa, et al. 2011). However, like many parrots monk parakeets have a large and variable vocal repertoire, where individuals might benefit from individual recognition across call types, rather than in a single identity call. In support of this prediction, we found that three of five tested call types in monk parakeets contained some evidence for an individual signal. We found no evidence for individual distinctiveness in the growl or *tja* calls. The structure and length of a call are likely to affect its ability to convey individualized information. The *tja* call is too short to allow for many unique variants and is often used in combination with other calls. Similarly, the growl is often used in close-range social interactions and has no tonal structure.

We proposed three hypotheses for how a vocal recognition system could be achieved in monk parakeets. First, individuals could utilize individual signals in several call types, unique to each call. Second, individuals could utilize a single unique signal that is added to the vocal sequences of multiple call types. Third, each individual could have a set of vocal features that are shared across all their calls, i.e., a voice-print. While our results provide evidence for an individual signal added to some call types, calls were also highly variable. Overall, our results best support the third hypothesis, that monk parakeets possess a voice-print that exists across call types, with a shared set of structural features that make them individually recognisable. While we cannot exclude the possibility that each call type also contains an individual signal, we find stronger evidence that a voice-print is present. If parrots can learn to recognise individuals based on a voice-print shared across calls, such a generalised mechanism relaxes the pressure to produce structural components in each call. This allows calls to include other signals (e.g. group identity) and reduces memory burden on the receiver significantly. A distinct and recognisable voice-print could be a particularly useful strategy used to manage individual recognition for open-ended vocal learners living in complex but cohesive social groups. There is a good reason to expect voice-prints to be present in parrots. Unlike song birds, that produce their vocalisation using two relatively independent syringeal sound sources, parrots have only one sound source and modulate their vocalisations using trachea, tongue and beak. This is very similar to how humans produce the sounds that make up words (Nottebohm 1976; Ohms et al. 2012; Larsen and Goller 2002; Beckers, Nelson, and Suthers 2004; DIANE K Patterson and Pepperberg 1994; Warren, Dianne K Patterson, and Pepperberg 1996; Bottoni, Masin, and Lenti-Boero 2009; Brittan-Powell et al. 1997). This modulation or filtering by the vocal tract allows for many more individual-specific features to arise and make a voice-print more recognisable.

Along with these main results, we found a high degree of variability within calls, with calls spaced ten minutes apart much less similar than calls spaced a second apart. It is unlikely that variation in background noise played a role in producing this result, since dynamic time warping was performed on manually cleaned fundamental frequency traces and obtained similar results (see supplemental materials). A more plausible explanation is that individuals are not capable of reproducing exactly the same call after too much time has elapsed. In other words they might have a general template but only produce identical calls when they are uninterrupted. It is also possible that monk parakeets modify their call based on the context, audience, or emotional state. For example, some variants might be used in a foraging context where a partner is present while others are given in isolation. Another possibility is that monk parakeets actively modify their contact call to match other individuals in their group, similar to the rapid convergence found in orange-fronted conures (Balsby and Bradbury 2009). If this is the case, we would expect a sequence of calls to vary depending on whom an individual is directing their call towards and the size of the audience. This would also suggest that individuals in larger groups should exhibit more variable calls. Both of these scenarios remain to be studied in more depth.

The fact that individuals are so variable in their calls raises a methodological problem for dialect studies on unmarked populations. When recording in the wild, individuals canonly be monitored for short periods of time, it was rarely possible to record contact calls for more than 3-5 minutes in our study. Our results show that in this short period individuals are likely to exhibit a consistent individual signal, but this signal is less consistent across recordings. A common technique to exclude repeated sampling of individuals across recordings is to look for highly similar calls and exclude these (Buhrman-Deever, Rappaport, and Bradbury 2007; Smith-Vidaurre, Araya-Salas, and Wright 2020). However, this assumes one can reliably estimate how similar a call needs to be in order to classify it as the same individual. We show that this can not be reliably estimated from short-term recordings. Moreover, we show that determining which calls come from the same individual in a large sample is not realistic given the amount of individual variability in contact calls. Instead, we suggest estimating the probability of recording the same individual multiple times and using a sensitivity analysis to test if the detected dialect signal is likely to be a true signal, or if it could have been caused by pseudoreplication (see e.g., Smeele et al. (2022)). With this approach, researchers can reliably estimate dialect differences and begin to uncover how social learning biases generate dialect patterns.

Conclusion and Outlook

Despite decades of research, the ability of parrots to identify each other based on vocal recognition is still not well understood. Some species have clear group signals and dialects (Wright 1996), while others appear to have much more pronounced individual signal in their contact calls (Buhrman-Deever, Rappaport, and Bradbury 2007; Thomsen, Balsby, and Dabelsteen 2013; Smith-Vidaurre, Araya-Salas, and Wright 2020; Berg, Delgado, Okawa, et al. 2011). However, we know very little about the temporal stability of these individual signals or the extent to which signals exist in other call types. This study provides the first evidence for an individual voice-print across multiple call types in parrots. Additionally, it demonstrates significant individual variability in the contact call. Our findings suggest that the contact call is not unique in its ability to broadcast caller identity. Instead it appears that parrots have evolved the capacity to generate a voice-print similar to humans, allowing for individual recognition (Lavan et al. 2019). While our study provides evidence for detectable voice-prints in monk parakeets, further investigation is needed to establish whether parrots actively use voice-prints to recognize conspecifics. More generally, our results suggest that voice-prints could occur in other animal species, particularly those with complex social interactions. Future research should focus on understanding the prevalence of voice-prints in nature, as well as the dynamics driving the evolution of voice-prints in parrots and other species.

Data accessibility. Code and small data files will be publicly available on GitHub: https://github.com/simeonqs/Evidence_for_vocal_signatures_and_voice-prints_in_a_wild_parrot. The full repository including large data files will be made publicly available on Edmond upon acceptance.

Competing interests. All authors declare they have no competing interests.

Ethics. All monk parakeets were ringed and blood samples taken with special permission EPI 7/2015 (01529/1498/2015) from Direcció General del Medi Natural i Biodiversitat, Generalitat de Catalunya, following Catalan regional ethical guidelines for the handling of birds. JCS received authorization (001501-0402.2009) for animal handling for research purposes from Servei de Protecció de la Fauna, Flora i Animal de Companyia, according to Decree 214/1997/30.07.

Funding. SQS received funding from the International Max Planck Research School for Quantitative Behaviour, Ecology and Evolution. LMA was funded by a Max Planck Research Group Leader Fellowship. SQS and LMA were supported by the Centre for the Advanced Study of Collective Behaviour (CASCB), funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy (EXC 2117-422037984). JCS was supported by a research project from the Ministry of Science and Innovation (CGL-2020 PID2020-114907GB-C21).

Acknowledgements. We would like to thank Zoo Barcelona and Josep M. Alonso Farré for granting us access to the zoo grounds and showing us around. We would like to thank Andrés Manzanilla, Gustavo Alarcòn-Nieto, Mireia Fuertes Clavero, Alba Ortega-Segalerva and José G. Carrillo-Ortiz for all their effort during fieldwork. Finally, we would like to thank Francesca S. E. Dawson Pell, Daniel Redhead, Jack W. Bradbury, Susannah Buhrman-Deever, Grace Smith-Vidaurre, Timothy F. Wright, Roger Mundry, Mirjam Knörnschild and Elizabeth Hobson for valuable advice during the early stages of this project.

References

Aureli, Filippo, Colleen M Schaffner, and Gabriele Schino (2022). “Variation in communicative complexity in relation to social structure and organization in non-human primates”. In: *Philosophical Transactions of the Royal Society B* 377.1860, p. 20210306.

Balsby, Thorsten JS and Danielle M Adams (2011). “Vocal similarity and familiarity determine response to potential flockmates in orange-fronted conures (Psittacidae)”. In: *Animal Behaviour* 81.5, pp. 983–991.

Balsby, Thorsten JS and Jack W Bradbury (2009). “Vocal matching by orange-fronted conures (Aratinga canicularis)”. In: *Behavioural processes* 82.2, pp. 133–139.

Balsby, Thorsten JS, Jane Vestergaard Momberg, and Torben Dabelsteen (2012). “Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network”. In: *PLoS One* 7.11, e49747.

Batllori, X and R Nos (1985). “Presencia de la cotorrita gris (Myiopsitta monachus) y de la cotorrita de collar (Psittacula krameri) en el área metropolitana de Barcelona”. In: *Miscel· lània Zoològica*, pp. 407–411.

Beckers, Gabriël JL, Brian S Nelson, and Roderick A Suthers (2004). “Vocal-tract filtering by lingual articulation in a parrot”. In: *Current Biology* 14.17, pp. 1592–1597.

Berg, Karl S, Soraya Delgado, Kathryn A Cortopassi, et al. (2012). “Vertical transmission of learned signatures in a wild parrot”. In: *Proceedings of the Royal Society B: Biological Sciences* 279.1728, pp. 585–591.

Berg, Karl S, Soraya Delgado, Rae Okawa, et al. (2011). “Contact calls are used for individual mate recognition in free-ranging green-rumped parrotlets, Forpus passerinus”. In: *Animal Behaviour* 81.1, pp. 241–248.

Boesch, Christophe (1994). “Cooperative hunting in wild chimpanzees”. In: *Animal Behaviour* 48.3, pp. 653–667.

Bottoni, L, S Masin, and D Lenti-Boero (2009). “Vowel-like sound structure in an African Grey Parrot (Psittacus erithacus) vocal production”. In: *The Open Behavioral Science Journal* 3.1.

Bouchet, Hélène et al. (2012). “Acoustic variability and individual distinctiveness in the vocal repertoire of red-capped mangabeys (Cercocebus torquatus).” In: *Journal of Comparative Psychology* 126.1, p. 45.

Bradbury, Jack W and Thorsten JS Balsby (2016). “The functions of vocal learning in parrots”. In: *Behavioral Ecology and Sociobiology* 70.3, pp. 293–312.

Brittan-Powell, Elizabeth F et al. (1997). “Mechanisms of vocal production in budgerigars (Melopsittacus undulatus)”. In: *The Journal Of The Acoustical Society Of America* 101.1, pp. 578–589.

Buhrman-Deever, Susannah C, Amy R Rappaport, and Jack W Bradbury (2007). “Geographic variation in contact calls of feral North American populations of the Monk Parakeet”. In: *The Condor* 109.2, pp. 389–398.

Carlson, Nora V, E McKenna Kelly, and Iain Couzin (2020). “Individual vocal recognition across taxa: a review of the literature and a look into the future”. In: *Philosophical Transactions of the Royal Society B* 375.1802, p. 20190479.

Charrier, Isabelle et al. (2001). “Individual identity coding depends on call type in the South Polar skua Catharacta maccormicki”. In: *Polar Biology* 24.5, pp. 378–382.

Clark, Christopher W, Peter Marler, and Kim Beeman (1987). “Quantitative analysis of animal vocal phonology: an application to swamp sparrow song”. In: *Ethology* 76.2, pp. 101–115.

Dahlin, Christine R et al. (2014). “A test of multiple hypotheses for the function of call sharing in female budgerigars, Melopsittacus undulatus”. In: *Behavioral ecology and sociobiology* 68.1, pp. 145–161.

Eberhard, Jessica R (1998). “Breeding biology of the Monk Parakeet”. In: *The Wilson Bulletin*, pp. 463–473.

Elie, Julie E and Frédéric E Theunissen (2018). “Zebra finches identify individuals using vocal signatures unique to each call type”. In: *Nature communications* 9.1, pp. 1–11.

Farabaugh, Susan M, Alison Linzenbold, and Robert J Dooling (1994). “Vocal plasticity in budgerigars (Melopsittacus undulatus): evidence for social factors in the learning of contact calls.” In: *Journal of Comparative Psychology* 108.1, p. 81.

Giorgino, Toni (2009). “Computing and visualizing dynamic time warping alignments in R: the dtw package”. In: *Journal of statistical Software* 31, pp. 1–24.

Hobson, Elizabeth A, Dan Mønster, and Simon DeDeo (2021). “Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information”. In: *Proceedings of the National Academy of Sciences* 118.10.

Janik, Vincent M and Laela S Sayigh (2013). “Communication in bottlenose dolphins: 50 years of signature whistle research”. In: *Journal of Comparative Physiology A* 199.6, pp. 479–489.

Janik, Vincent M, Laela S Sayigh, and Randall S Wells (2006). “Signature whistle shape conveys identity information to bottlenose dolphins”. In: *Proceedings of the National Academy of Sciences* 103.21, pp. 8293–8297.

Johnstone, Rufus A (1997). “Recognition and the evolution of distinctive signatures: when does it pay to reveal identity?” In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264.1387, pp. 1547–1553.

K. Lisa Yang Center for Conservation Bioacoustics (2016). *Raven Lite: Interactive Sound Analysis Softarbare*. Version 2.0.1. URL: `https://ravensoundsoftware.com`.

Keenan, Sumir et al. (2020). “The reliability of individual vocal signature varies across the bonobo’s graded repertoire”. In: *Animal Behaviour* 169, pp. 9–21.

Kenny, David A and Lawrence La Voie (1984). “The social relations model”. In: *Advances in experimental social psychology*. Vol. 18. Elsevier, pp. 141–182.

Kummer, Hans (2017). *Primate societies: Group techniques of ecological adaptation*. Routledge.

Larsen, Ole Næsbye and Franz Goller (2002). “Direct observation of syringeal muscle function in songbirds and a parrot”. In: *Journal of Experimental Biology* 205.1, pp. 25–35.

Lavan, Nadine et al. (2019). “Flexible voices: Identity perception from variable vocal signals”. In: *Psychonomic bulletin & review* 26.1, pp. 90–102.

Mäkelin, Saara et al. (2021). “Vocal behaviour of the Great Cormorant Phalacrocorax carbo sinensis during the breeding season”. In: *Bird Study*, pp. 1–9.

Martella, Mónica B and Enrique H Bucher (1990). “Vocalizations of the monk parakeet”. In: *Bird Behavior* 8.2, pp. 101–110.

Mathias, Samuel R, Katharina von Kriegstein, et al. (2014). “How do we recognise who is speaking”. In: *Frontiers in bioscience* 6, pp. 92–109.

Migliano, Andrea Bamberg and Lucio Vinicius (2022). “The origins of human cumulative culture: from the foraging niche to collective intelligence”. In: *Philosophical Transactions of the Royal Society B* 377.1843, p. 20200317.

Mundinger, Paul C (1970). “Vocal imitation and individual recognition of finch calls”. In: *Science* 168.3930, pp. 480–482.

Mundry, Roger and Christina Sommer (2007). “Discriminant function analysis with nonindependent data: consequences and an alternative”. In: *Animal Behaviour* 74.4, pp. 965–976.

Nottebohm, Fernando (1976). “Phonation in the orange-winged Amazon parrot, Amazona amazonica”. In: *Journal of Comparative Physiology* 108.2, pp. 157–170.

Ohms, Verena R et al. (2012). “Vocal tract articulation revisited: the case of the monk parakeet”. In: *Journal of Experimental Biology* 215.1, pp. 85–92.

Patterson, DIANE K and Irene M Pepperberg (1994). “A comparative study of human and parrot phonation: acoustic and articulatory correlates of vowels”. In: *The Journal of the Acoustical Society of America* 96.2, pp. 634–648.

Prat, Yosef, Mor Taub, and Yossi Yovel (2016). “Everyday bat vocalizations contain information about emitter, addressee, context, and behavior”. In: *Scientific Reports* 6.1, pp. 1–10.

Rauber, Ramona, Bart Kranstauber, and Marta B Manser (2020). “Call order within vocal sequences of meerkats contains temporary contextual and individual information”. In: *BMC biology* 18.1, pp. 1–11.

Reby, David et al. (2006). "Cepstral coefficients and hidden Markov models reveal idiosyncratic voice characteristics in red deer (Cervus elaphus) stags". In: *The Journal of the Acoustical Society of America* 120.6, pp. 4080–4089.

Russell, Andrew F and Ben J Hatchwell (2001). "Experimental evidence for kin-biased helping in a cooperatively breeding vertebrate". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1481, pp. 2169–2174.

Salmi, Roberta, Kurt Hammerschmidt, and Diane M Doran-Sheehy (2014). "Individual distinctiveness in call types of wild western female gorillas". In: *Plos One* 9.7, e101940.

Scarl, Judith C and Jack W Bradbury (2009). "Rapid vocal convergence in an Australian cockatoo, the galah Eolophus roseicapillus". In: *Animal Behaviour* 77.5, pp. 1019–1026.

Senar, Juan Carlos, Jose Carrillo-Ortiz, and Lluisa Arroyo (2012). "Numbered neck collars for long-distance identification of parakeets". In: *Journal of Field Ornithology* 83.2, pp. 180–185.

Sheehan, Michael J and Michael W Nachman (2014). "Morphological and population genomic evidence that human faces have evolved to signal individual identity". In: *Nature communications* 5.1, pp. 1–10.

Smeele, Simeon Q. et al. (2022). "Multi-level analysis of monk parakeet vocalisations shows emergent dialects between cities in the European invasive range". In: *bioRxiv*. DOI: `10.1101/2022.10.12.511863`. eprint: `https://www.biorxiv.org/content/early/2022/10/16/2022.10.12.511863.full.pdf`. URL: `https://www.biorxiv.org/content/early/2022/10/16/2022.10.12.511863`.

Smith-Vidaurre, Grace, Marcelo Araya-Salas, and Timothy F Wright (2020). "Individual signatures outweigh social group identity in contact calls of a communally nesting parrot". In: *Behavioral Ecology* 31.2, pp. 448–458.

Smith-Vidaurre, Grace, Valeria Perez-Marrufo, and Timothy F Wright (2021). "Individual vocal signatures show reduced complexity following invasion". In: *Animal Behaviour* 179, pp. 15–39.

Teixeira, Daniella et al. (2021). "Vocal signals of ontogeny and fledging in nestling black-cockatoos: Implications for monitoring". In: *Bioacoustics*, pp. 1–18.

Thomsen, Heidi M, Thorsten JS Balsby, and Torben Dabelsteen (2013). "Individual variation in the contact calls of the monomorphic peach-fronted conure, Aratinga aurea, and its potential role in communication". In: *Bioacoustics* 22.3, pp. 215–227.

Tibbetts, Elizabeth A (2004). "Complex social behaviour can select for variability in visual features: a case study in Polistes wasps". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271.1551, pp. 1955–1960.

Tibbetts, Elizabeth A and James Dale (2007). "Individual recognition: it is good to be different". In: *Trends in ecology & evolution* 22.10, pp. 529–537.

Warren, Denice K, Dianne K Patterson, and Irene M Pepperberg (1996). "Mechanisms of American English vowel production in a grey parrot (Psittacus erithacus)". In: *The Auk* 113.1, pp. 41–58.

Wein, Amelia et al. (2021). “Vocal development in nestling kea parrots (Nestor notabilis)”. In: *Bioacoustics* 30.2, pp. 142–162.

Wright, Timothy F (1996). “Regional dialects in the contact call of a parrot”. In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263.1372, pp. 867–872.

Wright, Timothy F and Christine R Dahlin (2018). “Vocal dialects in parrots: patterns and processes of cultural evolution”. In: *Emu-Austral Ornithology* 118.1, pp. 50–66.

Wright, Timothy F, Christine R Dahlin, and Alejandro Salinas-Melgoza (2008). “Stability and change in vocal dialects of the yellow-naped amazon”. In: *Animal behaviour* 76.3, pp. 1017–1027.

Wright, Timothy F, Erina Hara, et al. (2015). “Extreme vocal plasticity in adult budgerigars: Analytical challenges, social significance, and underlying neurogenetic mechanisms”. In: *The Journal of the Acoustical Society of America* 138.3, pp. 1880–1880.