Vocal communication in zebra finches: a focused description of pair vocal activity

Dissertation

Fakultät für Biologie

Ludwig-Maximilians-Universität München

durchgeführt am

Max-Planck-Institut für Ornithologie

Seewiesen

vorgelegt von

Pietro Bruno D'Amelio

München, 2018

Erstgutachter: Prof. Dr. **Manfred Gahr**

Zweitgutachter: Prof. Dr. **Niels Dingemanse**

Eingereicht am: 06.03.2018

Tag der mündlichen Prüfung: 08.05.2018

Diese Dissertation wurde unter der Leitung von Prof. Dr. **Manfred Gahr** und Dr. **Andries ter Maat** angefertigt.

Table of Contents

Summary

Vocal communication is the primary mode of signaling in a wide variety of species and commonly plays a decisive role in reproduction and survival of both the sender and the receiver. In birds, vocal signals have evolved to be astonishingly diverse, with thousands of different vocalizations used for highly disparate functions. Some vocalizations are loud and broadcast to reach the greatest possible number of receivers; others are barely audible and directed to a specific individual, often a mate. The function of each element of a bird's vocal repertoire is not always simple to interpret and the same signal can have multiple functions. To understand the role and influence of each vocal signal of a species, and thus the principles of vocal communication, we need to precisely quantify not only each sender vocalization, but also the context of its emission.

In my thesis, I used the zebra Finch (*Taeniopygia guttata*) as a model species to study the vocal communication of birds. I focused on pair communication as the smallest functional unit of a group. I considered all the vocal signals emitted but I focused my attention on the calls: of which thousands are emitted daily. These vocalizations are unlearned, very soft (low amplitude), very short (>100 milliseconds) and very similar among different individuals. Therefore, my colleagues and I developed and evaluated a method to individually record small animals (**chapter 1**). I describe a device, miniaturized backpack microphones, able to record birds singularly and the tools needed to record several individuals synchronously. This tool can be employed in many experimental settings to quantify the vocal behavior of multiple individuals with only a transient effect on their behavior and capable of capturing the softest of their vocalizations. I first used this tool to verify that birds were able to tell who was calling (**chapter 2**). I found that even the short unlearned calls have an individual vocal signature and birds are particularly eager to answer their mate's vocalizations. Then, I studied the development of, and signals used in within pair communication **(chapter 3)**. I described in detail the timed vocal exchanges of paired zebra finches; I observed that their antiphonal calling resembled duets. Moreover, I found that pair communication develops along with pair formation, so that by the time the pair is formed the number of calls used to answer each other is similar between partners. Finally, I aimed to discover when the calling interactions were taking place to have insights into the functions of these short, soft, unlearned vocalizations of zebra finches. To do so, I assembled a simple method to track automatically the position of the birds while recording their vocalizations. I was then able to describe how the relative positions of partners within a pair influenced the probability of calling (**chapter 4**), which revealed the precise meaning of a specific vocalization. This experiment also shows the accuracy and precision of our behavioral quantification, marking an important step towards new methods to automatically generate ethograms. Altogether, my dissertation contributes to our understanding of bird vocalization by determining that the continuous calling of zebra finches is not just “a soft background hum”, but rather the foundation of organized vocal networks.

General Introduction

Vocal communication

Understanding communication is part of studying how animals make decisions, which can be a conscious process or not (Schmidt, Dall, & van Gils, 2010). Focusing on the vocal channel, I developed and used tools to accurately describe bird behavior, quantify their decisions, and infer communicative principles.

Vocal communication is a widespread form of conveying messages. It is rapid, works over short to medium distances, and does not leave tracks (Bradbury & Vehrencamp, 1998). It is used by many taxa across the animal kingdom, primarily as a signal to inform and modify the decision process of receivers. For human listeners, birds hold a special place among animals. This is due to several practical reasons besides the aesthetic and euphonious fascination: they are common around us, often extremely chatty, and some species are easy to raise and manipulate. What also makes birds special for us is that they share with humans the ability to learn their vocalizations (Barrington, 1773; Doupe & Kuhl, 1999). As a result, birds became the most common model to study vocal communication and learning (Brainard & Doupe, 2013). Vocal communication studies have mainly focused on compiling and understanding vocal repertoires (Fischer, Wadewitz, & Hammerschmidt, 2017) and, more recently, constructing vocally defined social networks (Gill, Goymann, Ter Maat, & Gahr, 2015). Vocal learning is the capability to modify the spectral and temporal features of vocalizations using a template, and is quite rare in the animal kingdom (Petkov & Jarvis, 2012). Vocal learning is a relatively rare trait even in birds; it has evolved only 2 or 3 times; in the order of hummingbirds (Apodiformes) (Baptista & Schuchmann, 1990) and in the progenitor of parrots (Psittaciformes) and passerines (Passeriformes) (Suh et al., 2011). The independent evolution of vocal learning in these two clades is also possible (Zhang et al., 2014). Thus, only a few of the circa 20 (depending on the different taxonomies) orders of birds learn vocalizations. However, all birds communicate vocally.

Vocalizations in birds can be divided in songs and calls. The distinction is far from clear cut, but a simple definition states that songs are longer and more spectrally modulated, whereas calls are shorter and have simpler spectro-temporal structure. Passeriformes are taxonomically divided in two clades Passeri (oscines) and Tyranni (suboscines) and only the former are able to learn vocalizations. For the clade Passeri, the songs are usually regarded as learned whereas the calls, with few exceptions, as are thought to be unlearned (Marler, 2004). It must be noted that for the vast majority of species of this clade we do not know whether calls are learned or innate. However, in the vocal learner species for which we have sufficient information, the majority of the repertoire mainly consists of unlearned vocalizations. Despite this overwhelming majority of unlearned vocalizations present in the class Aves most of the research has focused on learned songs. Songs are indeed a useful model for vocal learning given its analogies with human speech acquisition (Doupe & Kuhl, 1999). For example, in the focal species of the present dissertation, the zebra finch (*Taeniopigya guttata*), thousands of studies have tried to unravel every aspect of their song, from learning to usage during mate choice, and only an order of magnitude less have included their calls (web of science for “zebra finch* song*”= 2725, “zebra finch* call*”= 388). However, in an average day, a male zebra finch produces only few hundred songs but it emits several thousand calls (**chapter 3**), and females do not even sing. Hence, I suggest that in order to understand vocal communication, a deeper knowledge about unlearned vocalizations is necessary. This knowledge will also aid the understanding of vocal learning because, undoubtedly, the learned sounds have evolved upon unlearned ones.

Unlearned vocalizations, calls, are particularly interesting because they serve many disparate functions, whereas learned sounds, usually only have a few (Marler, 2004). Songs were commonly found to have two functions: mate attraction and territorial defense. Call functions range, not exhaustively, from mate attraction, to territory defense, alarm, recruitment, echolocation, contact, aggression and, most interestingly, calls can mediate social relationships (Gill et al., 2015).

The vocal repertoire

The first step to comprehend how a species communicates acoustically is to describe, classify, and quantify the acoustic parameters and spectral features of all vocal signals. The entire variety of vocal signals of a species is called vocal repertoire, but the definition of the categories, in many cases, is far from univocal. I will briefly describe what drives the evolution of different spectral shapes for different calls and how vocalizations are catalogued. Then I am going to argue that the study of the entire repertoire of signals and their functions is necessary to tackle general questions about vocal communications such as “why do birds vocalize so much?” or “what selective pressures might have favored learning?” or “ are vocal and social complexity related?”

Calls can be divided into two categories depending on how strongly the environment drives the evolution: the ones shaped by the “acoustic adaptation hypothesis”(Morton, 1975), and the ones that do not have strong constraints in their spectral content. The acoustic adaptation hypothesis states that the environment, influencing sound propagation, determines the usage and/or the spectral content of calls. A review on this topic found that vocalizations of anurans and mammals can mostly be explained by this framework whereas for birds, this does not always hold true (Ey & Fischer, 2009). In fact, in birds, only alarm and territorial calls are shaped by the biotic and abiotic environment (Boncoraglio & Saino, 2007; Hollén & Radford, 2009). Alarm calls, for instance, need to aid or avoid being located and territorial calls have to travel as far as possible. No relation with the environment, however, can explain the spectral shape of the soft contact calls used in mediating relationships, or breeding calls used to choose the nest site. The only possible explanation to fit these calls into the acoustic adaptation hypothesis is if the spectral content stimulates the brain of the receiver in a specific way their shape would still be a byproduct of their function (sensory bias; Ryan, 1990); but this remains to be proven for any bird call. This lack of strong constraints for spectral shape makes these types of calls (contact, breeding, etc.) interesting because their spectral content can be rapidly changed and modified by evolution or referential rules, and they can easily evolve to serve multiple functions.

Calls have also been divided into stereotyped and graded (Marler & Mitani, 1988). In graded calls the spectral structure of the different vocalizations fades one into another (continuous), whereas in stereotyped calls it is fixed (discrete). Continuous signals are usually assigned to continuous variables; in contrast, discrete signals are usually assigned to alternative conditions (Bradbury & Vehrencamp, 1998). Where to set the limit between two categories of graded calls is often not an easy task. As a result the real repertoire size is often under debate (e.g. for zebra finches see Elie & Theunissen, 2015) and much more information is necessary to delineate it (Fischer et al., 2017). In addition, many studies on the vocal repertoire composition are dependent on the technique used to set the limit between the different categories, and not based on the animal perception. Seldom have researchers asked the animals whether the categories that we describe were recognized.

The tools and experiments described in my theses provide both methods to investigate the entire vocal repertoire and examples of insights into the complexity reached when considering all vocalizations produced by birds.

Bird calls

Studying more call functions and investigating larger parts of bird repertoires has recently led to the discovery that unlearned bird calls are involved in several cognitively demanding tasks. For example, playback experiments on Japanese Tits (*Parus minor*) have unraveled that sequences of calls actually have syntax (i.e. the position in the sequence is important to convey the meaning) (Griesser, Wheatcroft, & Suzuki, 2018; Suzuki, Wheatcroft, & Griesser, 2016). In this case of call usage we find lexical syntax which, from a linguistic point of view, is much more interesting than the phonological syntax of songs where the sequence of the elements is important but does not convey a specific meaning. Pied Babblers (*Turdoides bicolor*) can combine calls to form a new message, also called combinational meaning (Engesser, Ridley, & Townsend, 2016). Not only can calls be combined, but even a single call type can be altered (e.g. changing the amplitude) and used with specific rate and loudness to obtain many functions from a small repertoire (Ręk, 2015). Unlearned calls have also been found to have unsuspected functions such as the “egg calls” of the zebra finch, mothers vocalize to the eggs to modify the development of embryos and prepare the offspring for specific environmental variables (Mariette & Buchanan, 2016). Spectral features of calls, despite being unlearned, can still evolve quickly, even faster than songs in certain contexts, implying that their flexibility allows them to serve new functions (Sturge, Omland, Price, & Lohr, 2016). Although most calls are unlearned, their usage can be learned; nevertheless not many studies have investigated this aspect. At the neural level, although unlearned calls can be produced using the brainstem only (Simpson & Vicario, 1990), it is interesting to note that telencephalic areas are also activated during their production (Benichov et al., 2016; Ter Maat, Trost, Sagunsky, Seltmann, & Gahr, 2014). This pattern of neural activity might be derived from learning of how and when to use calls.

One reason why unlearned calls have not been studied extensively, despite all the listed intriguing characteristics, is that it was technically difficult. Calls are short vocalizations often very similar between individuals hence is difficult to locate the source. In addition, calls can be loud like in the case of alarm and some contact calls; but more often they are of very low amplitude like in the case of contact calls used within a flock. Ideally, the function of calls should be studied in wild populations, in a natural context. Moreover to capture the entire repertoire, all the life history stages should be studied. However, this is not always feasible and certain techniques still cannot be applied in the wild. Under controlled laboratory conditions and we can try to extrapolate meanings of calls and then compare them to those produced in the wild. During my doctoral studies, I tried to understand how and why zebra finches chat so much. I did not reduce my investigation to specific vocalizations but instead always tried to consider the entire repertoire.

Zebra finch vocal repertoire and communication

The zebra finch is a monogamous Australian bird which is the model for research in many different fields (Griffith & Buchanan, 2010). Many characteristics make zebra finches suitable for vocal communication studies. Zebra finches are very vocal, producing thousands of calls and songs daily (**chapter 3**). They have two levels of vocal exchanges: one with the members of the group and one with their lifelong partner (Zann, 1994). The zebra finch repertoire is one of the most investigated, but it is still unclear how many different call types it includes (**chapter 4**). The research about zebra finch biology and vocalization started in the second half of the 20th century (Immelmann, 1962, 1968, Zann, 1975, 1996). This was the foundation for the repertoire description, which was found to be composed of the following call types: The loud and distinctive *Distance* call, emitted when an individual loses sight of its mate. Two types of affiliative contact calls, the *Stack* and the *Tet*, used for keeping contact and advertise take off. An aggressive call, the *Wsst* call, emitted to intimidate a rival. Several types of breeding calls, *Kackle*, *Ark*, *Whine* and *Copulation* calls, very soft and with few harmonics, used to choose the nest site and pre-copulation. One alarm call, the *Thuck*, warning the young to fly. Moreover, juveniles have specific call types for food solicitation and localization. However, these early studies lacked a precise and objective quantification of behavior, the context of the calls was qualitatively assigned, the position of other birds seldom considered, and only a few vocalizations per type were typically analyzed.

More quantitative studies on the usage of contact calls for antiphonal calling were performed by Blaich and colleagues (Blaich et al., 1996; Blaich, Steury, Pettengill, Mahoney, & Guha, 1996). They described the calling pattern of separated pairs, and found that zebra finches answer to each other in a timed fashion with long duration *distance* and short contact calls. Because of technical limitations, birds had to be separated to be recognized which is why the dynamics of the interaction were not considered.

In the following years, Vignal and collaborators made an important contribution to the study of the zebra finch repertoire. They discovered that the distance calls of males and females contain identity signatures and that the social environment is essential for their utterance (Vignal, Mathevon, & Mottin, 2004, 2008).

They also described that in the wild paired zebra finches duet at the nest using breeding and contact calls (Elie et al., 2010). Subsequently they also investigated the meaning of the nest duets in captivity, finding that they are important to mediate incubation bouts (Boucaud, Mariette, Villain, & Vignal, 2016). All this work helped to refine the knowledge about vocal communication in the zebra finch. However the labels for each call type had not yet been clearly established.

Elie and Theunissen tried to fill this gap with a new approach (Elie & Theunissen, 2015). They used a dataset of high quality recordings of a few thousand vocalizations emitted in a large aviary representing different conditions. They linked vocalizations with different spectral features to behavioral observations, relating specific contexts and vocal emission. They then applied a blind classifier to understand when the two are correlated to finally infer the signal function. Even their study, however, has strong limitations, as it only considers a small fraction of the vocalizations emitted daily and lacks dynamicity, because the behavioral context is not described with sufficient temporal resolution. In fact, the turn taking dynamics and the real time movement and relative position of other birds was not considered. Therefore some calls with different spectral features were assigned to the same behavioral context.

All the listed studies also have further limitations: despite the fact that birds produce thousands of calls daily, which might require a considerable effort, very few studies quantified the number of calls produced per day or even per hour (Elie, Soula, Mathevon, & Vignal, 2011; Gill et al., 2015). As a consequence it is still unclear why a few call types are repeated so frequently. Another unclear point about calls studies is that the receiver perspective has never been considered: we can only assume that a *Tet* has the same meaning for all the receivers and there is no coding convention shared between senders and receivers. Only by studying the vocal communication from the first encounter of two individuals for many different pairs of birds we can resolve this doubt. Furthermore, until recently, it was hypothesized that the most common call type, the stack call, was not directed to specific individuals; it was described as undirected continuous noise in which the other calls are embedded (Zann, 1996). However we now have cues that this is not the case (Gill et al., 2015; chapter 2 and 3). Until now, we did not even know whether receivers can extract information about the individual identity while listening the short, soft, unlearned calls. Individual vocal recognition is an important piece of information if we want to assure that birds know to whom they are signaling when they decide to do so.

In my thesis, we used continuous individualized and synchronized recordings, and integrate this vocal information real time with the surrounding social context to clarify the function of the different zebra finch calls. I first describe a new method to record soft vocalizations. Then, I studied whether calls contain information about individual identity. Next, I described how each call of the zebra finch repertoire is used within a pair, and finally, I have studied when the signals are emitted to infer their functions. I hope that my work on vocal communication in zebra finches will go beyond refining the description of the repertoire and its usage, I aim to use the information produced to begin to bridge behavioral mechanisms and evolutionary perspective.

Methodological challenges and how they were approached

One microphone for each individual

To know who is talking to whom and saying what within a group we need audio recordings that are both individualized, to recognize who is talking, and synchronized, to reconstruct temporal relationships. Different strategies have been attempted to fulfill these characteristics. Researchers isolated birds in single cages (Fernandez, Vignal, & Soula, 2017) or they tried to recognize the caller a posteriori using discrimination algorithms (Fernandez, Soula, Mariette, & Vignal, 2016; Suzuki et al., 2017). However, a strategy that is becoming more and more common is to deploy tools on the bird itself (Anisimov et al., 2014; **chapter 1**). In the latter case two different types of device have been employed, loggers and transmitters. Briefly, the fundamental difference is that loggers store data and generally need to be retrieved to download them; whereas transmitters send data continuously via radio frequencies, making the device lighter and more battery efficient compared to loggers. However to be received, the device needs to be within the range of the antenna. In my first chapter we show benefits and limitations of the use of miniaturized radio transmitters applied to small birds. Besides describing how the transmitter is constructed and how to process the resulting radio signal we focus on how fast birds recovered and how the battery exchange, necessary every two weeks, affects the behavior. We then tested the device in difficult conditions such as noise and the presence of other individuals comparing its performance with recordings from a general microphone. We did not limit the study to the description of the used tool and its effect but we also reviewed the current literature and highlighted the research questions that can be addressed with our new tool.

Automating and integrating scoring of birds position and vocal behavior

The quantification of behavior is becoming more and more objective (Anderson & Perona, 2014; Berman, Choi, Bialek, & Shaevitz, 2014). Tools are being developed to follow animals continuously in order to score automatically how they interact with the environment and each other (Gill et al., 2016; Kays, Crofoot, Jetz, & Wikelski, 2015; Sakamoto et al., 2009). The relative position between individuals is important to predict and understand vocal communication. Qualitatively we know that the type and/or frequency of calls changes depending on the distance between birds (Blaich et al., 1996), however no precise quantification has been attempted so far. Tracking birds' position continuously has always been challenging because it needs to be done in three dimensions. With a simple set of strategies we overcame this problem and built a real time automatic position detection system for two birds. We employed it to integrate birds' positions with their individualized vocal emission with a fine time resolution. This approach allowed us to investigate the entire zebra finch vocal repertoire and resolve uncertainties of its characterizations (**chapter 4**).

Vocal individual recognition

Individuality becomes necessary in groups (Pollard & Blumstein, 2011). Recognition and memorization of individuals is required for maintaining relationships (e.g. with the mate or a place in hierarchical societies). The study of individual recognition is a necessary step when investigating the complexity and the structure of social systems (Wiley, 2013).

Individual vocal recognition is the ability to identify an individual solely by its vocalizations. This capability is widespread among many mammalian and avian species, which can vocally recognize their partners, kin or group members (Janik & Sayigh, 2013; Lambrechts & Dhondt, 1995; Nakagawa & Waas, 2004). Specifically, individual vocal recognition allows real time identification over distance, often when vision, olfaction and tactile information would not work. Furthermore, it allows communication specifically to one individual, even in cocktail party situations (Bee & Micheyl, 2008). Despite the amount of research invested in the topic, two areas remain unclear: i) the level of specificity and multiplicity of recognition (Wiley, 2013) because often only one level of familiarity is tested, and ii) whether birds recognize the voice of specific individuals or rely only on “signatures” in the caller vocalization. This is because research that has focused on the most prominent vocalizations, and the entire vocal repertoire of a species is never been investigated for vocal recognition. For example, vocal individual recognition using short, usually unlearned, vocalizations in adult Passeriformes, although possibly widespread, has been investigated only in few species. American Goldfinches (*Spinus tristis*), Silvereye, (*Zosterops lateralis*) and Pinion Jay (*Gymnorhinus cyanocephalus*), are able to identify the partner through short calls (Berger & Ligon, 1977; Mundinger, 1970; Robertson, 1996). Superb Starling (*Lamprotornis superbus*), Apostlebirds (*Struthidea cinerea*), Large-billed crow (*Corvus macrorhynchos*), and Chestnut-crowned Babblers (*Pomatostomus ruficeps*) are able to recognize single individuals in a group (Crane et al., 2015; Keen, Meliza, & Rubenstein, 2013; Kondo, Izawa, & Watanabe, 2010; Warrington, McDonald, & Griffith, 2015). In a few cases, kin and group recognition has been successfully found (Sharp, McGowan, Wood, & Hatchwell, 2005). However, not even in these cases has the entire repertoire been investigated.

The ability of zebra finches to recognize individuals has been known for a long time. The first experiment to show that they are able to connect vocalizations to specific individuals was performed by Miller in 1979 (Miller, 1979). They asked whether in a classic T-maze experiment females were able to recognize the mate only by listening to songs. In the following years other researchers investigated whether the most loud and prominent of the zebra finches' vocalizations, the *distance* call, also contains individual identity (Menardy et al., 2012; Vignal et al., 2004, 2008). Vocal individual recognition was found in males, whose call is learned, and females, where is unlearned. Given that this call type is emitted when birds have lost visual contact and need to find their partner, the presence of an acoustic signature can be expected (Blaich et al., 1996). Less obvious is why zebra finch receivers would be able to detect individual identity in the rest of the call repertoire since they are emitted when birds are close to each other (Zann, 1996). One hypothesis is that recognition of specific individuals allows for fast-turn-taking communication in a crowd. **Chapter 2** investigates which vocalizations in the repertoire of both sexes contain individual identity by contrasting three level of familiarity (mate, group member, and unfamiliar); we aim to find out if they are able to recognize to whom the voice belongs.

Pair communication

Learning when to answer your partner is a skill that requires some practice to be mastered. Nowadays, it is trivial to say that pairs are not all the same in terms of fitness, but why that is the case is still partially unresolved. For decades researchers have tried to identify the traits of single individuals that make a pair successful (Maklakov & Arnqvist, 2009). Part of this research field examined what makes an individual successful in relation with the partner, the study of pair compatibility (Neff & Pitcher, 2005). Studies on pair compatibility have compared chosen pairs with forced ones measuring the resulted fitness (for details on different experimental designs see introduction of: Ihle, Kempenaers, & Forstmeier, 2015). As I have illustrated, vocal communication in birds is a fundamental channel of signal exchange. Communication within a pair can be part of mate choice and pair maintenance, and might differ between more or less compatible pairs. The vocalizations of birds during mate choice have been the focus of intense research, and the role and pattern of emission of song during pair choice and maintenance is extremely well studied (Adkins-Regan & Tomaszycki, 2006; Bolund, Schielzeth, & Forstmeier, 2012; Byers & Kroodsma, 2009; Riebel, 2009). However, in Passeriformes the function of other vocalizations during the same processes has been almost completely neglected. Lifetime monogamous species not only have to find a mate, but they also need to keep the formed bond for an extended time. Lifetime monogamy is quite rare among Passeriformes (Henderson, Hart, & Burke, 2000), but its nurture in other taxa often involves vocalizations, for example vocal exchanges between Albatross partners are rich and complex (Lequette & Jouventin, 1991) and monogamous parrot species partners modify their contact calls to match their spectral features (Hile, Plummer, & Striedter, 2000). In Passeriformes, the vocal mechanisms helping pair maintenance are neither well known nor investigated. The zebra finch is a good model to explore pair vocal exchanges and their role in mate choice and pair maintenance in lifelong monogamy. They have a privileged channel of communication between mates that involves different types of vocalizations (Gill et al., 2015). In addition, pairs tend to maintain the same partner even when given the possibility to re-pair in captivity (Ihle et al., 2015) and in the wild (Zann, 1994). Therefore, pair choice is supposed to be a crucial moment for zebra finches since they will most likely keep their partner for life. However the research about the traits that makes a bird successful with the opposite sex has resulted in much controversy in this species, with many characteristics for quality that have been proposed to be important for mate choice: redness of the beak, song complexity or loudness, etc. (Adkins-Regan, 2011; Collins, Hubbard, & Houtman, 1994; Riebel, 2009). From all these efforts the picture that is emerging, at least in captive populations, is that, despite a mild repeatability of the pair choice, there are no single traits that can forecast pair choice and different individuals have repeatable preference for different traits

(Forstmeier & Birkhead, 2004) without a general consensus and without a preference for quality (Wang, Forstmeier, & Kempenaers, 2017). It was found that zebra finch that were allowed to choose their partner have higher fitness independently from the quality of the single individuals (Ihle et al., 2015). Therefore, pair compatibility seems to be important for success; however, we still do not know much about what parameters describe such compatibility. To be able to find some of these variables we should study pairs in a holistic way, starting from the first encounter leading to pair formation. Nonetheless, we are still lacking a complete description of what happens regarding vocal behavior involving calls when pairs form. Very little is also known about the involvement of vocalizations during the nurturing of the relationship. Singing does not seem to play and important role for pair maintenance (Adkins-Regan & Tomaszycki, 2006) but calls might be involved. While behaviorally we are still lacking a definitive quantitative description, neurobiological mechanisms of pair bonding have been studied (Tomaszycki, Richardson, & Mann, 2016) and interestingly converging mechanisms of social relationships with mammals have been found (Tomaszycki & Atchley, 2017). A more precise understanding of the vocal behavior will help to refine the knowledge about neural correlates of pair bonding.

To describe the role of vocalizations during pair formation and its nurturing we recorded continuous video and audio for a week for new and established pairs (**chapter 3**). We then compared the vocal and social behavior of forming pairs with already established ones. We used individual audio recordings via backpack microphones, thereby substantially improving the temporal precision and overall accuracy of our measurements compared to previous recording methods. We are now in position to speculate about the role of unlearned vocalizations during pair formation and maintenance.

References

Adkins-Regan, E. (2011). Neuroendocrine contributions to sexual partner preference in birds. *Frontiers in Neuroendocrinology*, *32*(2), 155–163. https://doi.org/10.1016/J.YFRNE.2011.01.003

Adkins-Regan, E., & Tomaszycki, M. (2006). Is male song quality important in maintaining pair bonds? *Behaviour*, *143*(5), 549–567. https://doi.org/10.1163/156853906776759529

Anderson, D. J., & Perona, P. (2014). Toward a Science of Computational Ethology. *Neuron*, *84*(1), 18–31. https://doi.org/10.1016/J.NEURON.2014.09.005

Anisimov, V. N., Herbst, J. A., Abramchuk, A. N., Latanov, A. V, Hahnloser, R. H. R., & Vyssotski, A. L. (2014). Reconstruction of vocal interactions in a group of small songbirds. *Nature Methods*, *11*(11), 1135–7. Retrieved from http://dx.doi.org/10.1038/nmeth.3114

Baptista, L. F., & Schuchmann, K. -L. (1990). Song Learning in the Anna Hummingbird (Calypte anna). *Ethology*, *84*(1), 15–26. https://doi.org/10.1111/j.1439-0310.1990.tb00781.x

Barrington, D. (1773). Experiments and Observations on the Singing of Birds. *Philosophical Transactions of the Royal Society of London*, *63*, 249–291. https://doi.org/10.1098/rstl.1773.0031

Bee, M. A., & Micheyl, C. (2008). The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, *122*(3), 235–251. Retrieved from http://psycnet.apa.orgjournals/com/122/3/235

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, *26*(3), 309–318. https://doi.org/10.1016/j.cub.2015.12.037

Berger, L. R., & Ligon, J. D. (1977). Vocal communication and individual recognition in the pinon jay, Gymnorhinus cyanocephalus. *Animal Behaviour*, *25*, 567–584. https://doi.org/10.1016/0003-3472(77)90107-5

Berman, G. J., Choi, D. M., Bialek, W., & Shaevitz, J. W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of the Royal Society, Interface*, *11*(99), 20140672. https://doi.org/10.1098/rsif.2014.0672

Blaich, C. F., Norman, M., Syud, F. A., Benitez, G., Frost, J., Ravenscroft, J., … Ware, P. (1996). The use of distance calls to maintain pair contact in zebra finches (*Taeniopygia guttata*). *Bird Behavior*, *11*(1), 25–30. https://doi.org/10.3727/015613896791748889

Blaich, C. F., Steury, K. R., Pettengill, P., Mahoney, K. T., & Guha, A. (1996). Temporal patterns of contact call interactions in pair-bonded domestic zebra finches (Taeniopygia guttata). *Bird Behavior*, *11*(2), 59--69(11).

Bolund, E., Schielzeth, H., & Forstmeier, W. (2012). Singing activity stimulates partner reproductive investment rather than increasing paternity success in zebra finches. *Behavioral Ecology and Sociobiology*, *66*(6), 975–984. https://doi.org/10.1007/s00265-012-1346-z

Boncoraglio, G., & Saino, N. (2007). Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. *Functional Ecology*, *21*(1), 134–142. https://doi.org/10.1111/j.1365-2435.2006.01207.x

Boucaud, I. C. A., Mariette, M. M., Villain, A. S., & Vignal, C. (2016). Vocal negotiation over parental care? Acoustic communication at the nest predicts partners' incubation share. *Biological Journal of the Linnean Society*, *117*(2), 322–336. https://doi.org/10.1111/bij.12705

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*. Sinauer Associates Sunderland.

Brainard, M. S., & Doupe, A. J. (2013). Translating Birdsong: Songbirds as a Model for Basic and Applied Medical Research. *Annual Review of Neuroscience*, *36*(1), 489–517. https://doi.org/10.1146/annurev-neuro-060909-152826

Byers, B. E., & Kroodsma, D. E. (2009). Female mate choice and songbird song repertoires. *Animal Behaviour*, *77*(1), 13–22. https://doi.org/10.1016/J.ANBEHAV.2008.10.003

Collins, S. A., Hubbard, C., & Houtman, A. M. (1994). Female mate choice in the zebra finch ? The effect of male beak colour and male song. *Behavioral Ecology and Sociobiology*, *35*(1), 21–25. https://doi.org/10.1007/BF00167055

Crane, J. M. S., Pick, J. L., Tribe, A. J., Vincze, E., Hatchwell, B. J., & Russell, A. F. (2015). Chestnut-crowned babblers show affinity for calls of removed group members: A dual playback without expectancy violation. *Animal Behaviour*, *104*, 51–57. https://doi.org/10.1016/j.anbehav.2015.02.022

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, *22*(1), 567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Elie, J. E., Mariette, M. M., Soula, H. A., Griffith, S. C., Mathevon, N., & Vignal, C. (2010). Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal Behaviour*, *80*(4), 597–605. https://doi.org/10.1016/j.anbehav.2010.06.003

Elie, J. E., Soula, H. A., Mathevon, N., & Vignal, C. (2011). Dynamics of communal vocalizations in a social songbird, the zebra finch (Taeniopygia guttata). *The Journal of the Acoustical Society of America*, *129*(6), 4037. https://doi.org/10.1121/1.3570959

Elie, J. E., & Theunissen, F. E. (2015). The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition*, *19*(2), 285–315. https://doi.org/10.1007/s10071-015-0933-6

Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(21), 5976–81. https://doi.org/10.1073/pnas.1600970113

Ey, E., & Fischer, J. (2009). The "Acoustic Adaptation Hypothesis"—A Review of evidence from birds, anurans and mammals. *Bioacoustics*, *19*(1–2), 21–48. https://doi.org/10.1080/09524622.2009.9753613

Fernandez, M. S. A., Soula, H. A., Mariette, M. M., & Vignal, C. (2016). A new semi-automated method for assessing avian acoustic networks reveals that juvenile and adult zebra finches have separate calling networks. *Frontiers in Psychology*, *7*(NOV), 1816. https://doi.org/10.3389/fpsyg.2016.01816

Fernandez, M. S. A., Vignal, C., & Soula, H. A. (2017). Impact of group size and social composition on group vocal activity and acoustic network in a social songbird. *Animal Behaviour*, *127*, 163–178. https://doi.org/10.1016/j.anbehav.2017.03.013

Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour*, *134*, 229–237. https://doi.org/10.1016/j.anbehav.2016.06.012

Forstmeier, W., & Birkhead, T. R. (2004). Repeatability of mate choice in the zebra finch: consistency within and between females. *Animal Behaviour*, *68*(5), 1017–1028. https://doi.org/10.1016/j.anbehav.2004.02.007

Gill, L. F., D'Amelio, P. B., Adreani, N. M., Sagunsky, H., Gahr, M. C., & ter Maat, A. (2016). A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods in Ecology and Evolution*, *7*(11), 1349–1358. https://doi.org/10.1111/2041-210X.12610

Gill, L. F., Goymann, W., Ter Maat, A., & Gahr, M. (2015). Patterns of call communication between group-housed zebra finches change during the breeding cycle. *eLife*, *4*, e07770. https://doi.org/10.7554/eLife.07770

Griesser, M., Wheatcroft, D., & Suzuki, T. N. (2018, June 1). From bird calls to human language: exploring the evolutionary drivers of compositional syntax. *Current Opinion in Behavioral Sciences*. Elsevier. https://doi.org/10.1016/j.cobeha.2017.11.002

Griffith, S. C., & Buchanan, K. L. (2010). The Zebra Finch: the ultimate Australian supermodel. *Emu - Austral Ornithology*, *110*(3), v–xii. https://doi.org/10.1071/MUv110n3_ED

Henderson, I. G., Hart, P. J. B., & Burke, T. (2000). Strict monogamy in a semi-colonial passerine: the Jackdaw Corvus monedula. *Journal of Avian Biology*, *31*(2), 177–182. https://doi.org/10.1034/j.1600-048X.2000.310209.x

Hile, A. G., Plummer, T. K., & Striedter, G. F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Animal Behaviour*, *59*(6), 1209–1218. https://doi.org/10.1006/anbe.1999.1438

Hollén, L. I., & Radford, A. N. (2009). The development of alarm call behaviour in mammals and birds. *Animal Behaviour*, *78*(4), 791–800. https://doi.org/10.1016/J.ANBEHAV.2009.07.021

Ihle, M., Kempenaers, B., & Forstmeier, W. (2015). Fitness Benefits of Mate Choice for Compatibility in a Socially Monogamous Species. *PLoS Biology*, *13*(9), e1002248. https://doi.org/10.1371/journal.pbio.1002248

Immelmann, K. (1962). *Beiträge zu einer vergleichenden Biologie australischer Prachtfinken (Spermestidae). Zoologische Jahrbücher Abteilung für Systematik Okologie und Geographie der Tiere*. G. Fischer.

Immelmann, K. (1968). Zur biologischen Bedeutung des Estrildidengesanges. *Journal of Ornithology*, *109*(3), 284–299. https://doi.org/10.1007/BF01678374

Janik, V. M., & Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *Journal of Comparative Physiology A*, *199*(6), 479–489. https://doi.org/10.1007/s00359-013-0817-7

Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478. https://doi.org/10.1126/science.aaa2478

Keen, S. C., Meliza, C. D., & Rubenstein, D. R. (2013). Flight calls signal group and individual identity but not kinship in a cooperatively breeding bird. *Behavioral Ecology*, *24*(6), 1279–1285.

https://doi.org/10.1093/beheco/art062

Kondo, Izawa, & Watanabe. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. *Behaviour*, *147*(8), 1051–1072. https://doi.org/10.1163/000579510X505427

Lambrechts, M. M., & Dhondt, A. A. (1995). Individual Voice Discrimination in Birds. In *Current Ornithology* (pp. 115–139). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-1835-8_4

Lequette, B., & Jouventin, P. (1991). The dance of the wandering albatross II: Acoustic signals. *Emu*, *91*(3), 172–178. https://doi.org/10.1071/MU9910172

Maklakov, A. A., & Arnqvist, G. (2009). Testing for Direct and Indirect Effects of Mate Choice by Manipulating Female Choosiness. *Current Biology*, *19*(22), 1903–1906. https://doi.org/10.1016/J.CUB.2009.08.058

Mariette, M. M., & Buchanan, K. L. (2016). Prenatal acoustic communication programs offspring for high posthatching temperatures in a songbird. *Science*, *353*(6301), 812–814. https://doi.org/10.1126/science.aaf7049

Marler, P. (2004). Bird Calls: Their Potential for Behavioral Neurobiology. *Annals of the New York Academy of Sciences*, *1016*(1), 31–44. https://doi.org/10.1196/annals.1298.034

Marler, P., & Mitani, J. (1988). Vocal Communication in Primates and Birds: Parallels and Contrasts. In *Primate Vocal Communication* (pp. 3–14). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-73769-5_1

Menardy, F., Touiki, K., Dutrieux, G., Bozon, B., Vignal, C., Mathevon, N., & Del Negro, C. (2012). Social experience affects neuronal responses to male calls in adult female zebra finches. *The European Journal of Neuroscience*, *35*(8), 1322–36. https://doi.org/10.1111/j.1460-9568.2012.08047.x

Miller, D. B. (1979). The acoustic basis of mate recognition by female Zebra finches (Taeniopygia guttata). *Animal Behaviour*, *27*, 376–380. https://doi.org/10.1016/0003-3472(79)90172-6

Morton, E. S. (1975). Ecological Sources of Selection on Avian Sounds. *The American Naturalist*, *109*(965), 17–34. https://doi.org/10.1086/282971

Mundinger, P. C. (1970). Vocal imitation and individual recognition of finch calls. *Science (New York, N.Y.)*, *168*(3930), 480–2. https://doi.org/10.1126/science.168.3930.480

Nakagawa, S., & Waas, J. R. (2004). O sibling, where art thou? a review of avian sibling recognition with respect to the mammalian literature. *Biological Reviews*, *79*(1), 101–119. https://doi.org/10.1017/S1464793103006249

Neff, B. D., & Pitcher, T. E. (2005, November 19). Genetic quality and sexual selection: An integrated framework for good genes and compatible genes. *Molecular Ecology*. Blackwell Science Ltd. https://doi.org/10.1111/j.1365-294X.2004.02395.x

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*, 12. https://doi.org/10.3389/fnevo.2012.00012

Pollard, K. A., & Blumstein, D. T. (2011). Social Group Size Predicts the Evolution of Individuality. *Current Biology*, *21*(5), 413–417. https://doi.org/10.1016/J.CUB.2011.01.051

Ręk, P. (2015). High functional complexity despite an extremely small repertoire of calls in the Spotted Crake (Porzana porzana). *The Auk*, *132*(3), 613–623. https://doi.org/10.1642/AUK-14-267.1

Riebel, K. (2009). Chapter 6 Song and Female Mate Choice in Zebra Finches: A Review. *Advances in the Study of Behavior*, *40*, 197–238. https://doi.org/10.1016/S0065-3454(09)40006-8

Robertson, B. C. (1996). Vocal mate recognition in a monogamous, flock-forming bird, the silvereye, *Zosterops lateralis*. *Animal Behaviour*, *51*(2), 303–311. https://doi.org/10.1006/anbe.1996.0030

Ryan, M. J. (1990). Sexual selection, sensory systems and sensory exploitation. *Oxford Surveys in Evolutionary Biology*, *7*, 157–195.

Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F., & Wanless, S. (2009). Can Ethograms Be Automatically Generated Using Body Acceleration Data from Free-Ranging Birds? *PLoS ONE*, *4*(4), e5379. https://doi.org/10.1371/journal.pone.0005379

Schmidt, K. A., Dall, S. R. X., & van Gils, J. A. (2010). The ecology of information: an overview on the ecological significance of making informed decisions. *Oikos*, *119*(2), 304–316. https://doi.org/10.1111/j.1600-0706.2009.17573.x

Sharp, S. P., McGowan, A., Wood, M. J., & Hatchwell, B. J. (2005). Learned kin recognition cues in a social bird. *Nature*, *434*(7037), 1127–1130. https://doi.org/10.1038/nature03522

Simpson, H. B., & Vicario, D. S. (1990). Brain pathways for learned and unlearned vocalizations differ in zebra finches. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *10*(5), 1541–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2332796

Sturge, R. J., Omland, K. E., Price, J. J., & Lohr, B. (2016). Divergence in calls but not songs in the orchard oriole complex: Icterus spurius and I. fuertesi. *Journal of Avian Biology*, *47*(1), 109–120. https://doi.org/10.1111/jav.00595

Suh, A., Paus, M., Kiefmann, M., Churakov, G., Franke, F. A., Brosius, J., … Schmitz, J. (2011). Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nature Communications*, *2*(1), 443. https://doi.org/10.1038/ncomms1448

Suzuki, R., Matsubayashi, S., Saito, F., Murate, T., Masuda, T., Yamamoto, K., … Okuno, H. G. (2018). A spatiotemporal analysis of acoustic interactions between great reed warblers (Acrocephalus arundinaceus) using microphone arrays and robot audition software HARK. *Ecology and Evolution*, *8*(1), 812–825. https://doi.org/10.1002/ece3.3645

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, *7*, 10986. https://doi.org/10.1038/ncomms10986

Ter Maat, A., Trost, L., Sagunsky, H., Seltmann, S., & Gahr, M. (2014). Zebra finch mates use their forebrain song system in unlearned call communication. *PloS One*, *9*(10), e109334. https://doi.org/10.1371/journal.pone.0109334

Tomaszycki, M. L., & Atchley, D. (2017). Pairing Increases Activation of V1aR, but not OTR, in Auditory Regions of Zebra Finches: The Importance of Signal Modality in Nonapeptide-Social Behavior Relationships. *Integrative and Comparative Biology*, *57*(4), 878–890. https://doi.org/10.1093/icb/icx043

Tomaszycki, M. L., Richardson, K. K., & Mann, K. J. (2016). Sex and pairing status explain variations in the activation of nonapeptide receptors in song and motivation Regions. *Behavioral Neuroscience*, *130*(5), 479–489. https://doi.org/10.1037/bne0000159

Vignal, C., Mathevon, N., & Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature*, *430*(6998), 448–51. https://doi.org/10.1038/nature02645

Vignal, C., Mathevon, N., & Mottin, S. (2008). Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behavioural Processes*, *77*(2), 191–8. https://doi.org/10.1016/j.beproc.2007.09.003

Wang, D., Forstmeier, W., & Kempenaers, B. (2017). No mutual mate choice for quality in zebra finches: Time to question a widely held assumption. *Evolution*, *71*(11), 2661–2676. https://doi.org/10.1111/evo.13341

Warrington, M. H., McDonald, P. G., & Griffith, S. C. (2015). Within-group vocal differentiation of individuals in the cooperatively breeding apostlebird. *Behavioral Ecology*, *26*(2), 493–501. https://doi.org/10.1093/beheco/aru217

Wiley, R. H. (2013). Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biological Reviews*, *88*(1), 179–195. https://doi.org/10.1111/j.1469-185X.2012.00246.x

Zann, R. A. (1975). Inter- and Intraspecific Variation in the Calls of Three Species of Grassfinches of the Subgenus Poephila (Gould) (Estrildidae). *Zeitschrift Für Tierpsychologie*, *39*(1–5), 85–125. https://doi.org/10.1111/j.1439-0310.1975.tb00902.x

Zann, R. A. (1994). Reproduction in a Zebra Finch Colony in South-eastern Australia: the Significance of Monogamy, Precocial Breeding and Multiple Broods in a Highly Mobile Species. *Emu*, *94*(4), 285–299. Retrieved from https://doi.org/10.1071/MU9940285

Zann, R. A. (1996). *The zebra finch: a synthesis of field and laboratory studies*. *Oxford Univ, Press*. Retrieved from http://books.google.com/books?id=5KO6cZH0WbEC&pgis=1

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., … Froman, D. P. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, *346*(6215), 1311–1320. https://doi.org/10.1126/science.1251385

Chapter 1

A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution

Abstract

1. To understand both proximate and ultimate factors shaping vocal communication, it is fundamental to obtain reliable information of participating individuals on different levels: First, it is necessary to separate and assign the individuals' vocalisations. Secondly, the precise timing of vocal events needs to be retained. Thirdly, vocal behaviour should be recorded from undisturbed animals in meaningful settings. A growing number of studies used animal-attached microphones to tackle these issues, but the implications for the study species and the research question often receded into the background. Here, we aim to initiate a discussion about the limitations, possible applications and the broader potential of such methods.

2. Using lightweight wireless microphone backpacks (0,75 g including customised leg-loop harness) combined with multi-channel recording equipment, we captured vocal behaviour of small songbirds. We evaluated the effect of the devices at various levels, including an assessment of how vocal and locomotor activities were affected by initial device attachment and battery exchange. We compared our approach to existing studies and identified suitable research examples.

3. We acquired continuous vocalisation recordings of zebra finches, and unequivocally assigned them to inter- acting individuals, with system-based synchrony, irrespective of background noise. We found effects of initial backpack attachment and of battery replacement on vocal and locomotor activity, but they were minimised through the extended recording duration (ca. 16 days) that outlasted habituation effects (ca. 3 days).

4. This method provides the tools to integrate individual vocal communications into a group setting, while enabling animals to behave freely in undisturbed, structured and acoustically complex environments. By

minimising the effects on the animals, the behaviour under study, and ultimately on the research question, this approach will revolutionise the ability to capture individual-level vocalisations in a variety of communication contexts, opening up many new opportunities to address novel research questions.

Published as: Gill, L. F.*, D'Amelio, P. B.*, Adreani, N. M.*, Sagunsky, H., Gahr, M. C., & ter Maat, A. (2016). A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods in Ecology and Evolution*, *7*(11), 1349–1358. https://doi.org/10.1111/2041-210X.12610

*joint first authors

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2016 doi: 10.1111/2041-210X.12610

A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution

Lisa F. Gill[†], **Pietro B. D'Amelio**[†], **Nicolas M. Adreani***,[†], **Hannes Sagunsky, Manfred C. Gahr and Andries ter Maat**

Max Planck Institute for Ornithology, Eberhard-Gwinner-Str., 82319 Seewiesen, Germany

Summary

1. To understand both proximate and ultimate factors shaping vocal communication, it is fundamental to obtain reliable information of participating individuals on different levels: First, it is necessary to separate and assign the individuals' vocalisations. Secondly, the precise timing of vocal events needs to be retained. Thirdly, vocal behaviour should be recorded from undisturbed animals in meaningful settings. A growing number of studies used animal-attached microphones to tackle these issues, but the implications for the study species and the research question often receded into the background. Here, we aim to initiate a discussion about the limitations, possible applications and the broader potential of such methods.

2. Using lightweight wireless microphone backpacks (0·75 g including customised leg-loop harness) combined with multi-channel recording equipment, we captured vocal behaviour of small songbirds. We evaluated the effect of the devices at various levels, including an assessment of how vocal and locomotor activities were affected by initial device attachment and battery exchange. We compared our approach to existing studies and identified suitable research examples.

3. We acquired continuous vocalisation recordings of zebra finches, and unequivocally assigned them to interacting individuals, with system-based synchrony, irrespective of background noise. We found effects of initial backpack attachment and of battery replacement on vocal and locomotor activity, but they were minimised through the extended recording duration (ca. 16 days) that outlasted habituation effects (ca. 3 days).

4. This method provides the tools to integrate individual vocal communications into a group setting, while enabling animals to behave freely in undisturbed, structured and acoustically complex environments. By minimising the effects on the animals, the behaviour under study, and ultimately on the research question, this approach will revolutionise the ability to capture individual-level vocalisations in a variety of communication contexts, opening up many new opportunities to address novel research questions.

Key-words: backpack microphones, habituation curve, individual vocalisation recordings, short-term effects, telemetry

Introduction

Due to technological limitations in obtaining good quality vocalisation recordings from individual animals, many open questions remain about animal vocal communication, especially in naturalistic settings. A growing body of studies has recently come to use animal-borne devices on species ranging from whales (Johnson, Aguilar de Soto & Madsen 2009) to chipmunks (Couchoux *et al.* 2015), opening up new fields of communication research. Before this, it was challenging to record and assign vocalisations of focal animals that were behaving freely inside their natural habitats, or in the presence of multiple sound sources, such as vocalising conspecifics or other background noises (Otter, Chruszcz & Ratcliffe 1997; Brumm 2004; Furrer & Manser 2009; Clemmons & Howitz 2010; Bousquet, Sumpter & Manser 2011). For example, when studying highly synchronised vocal interactions in social contexts, like in duetting birds, the role of individuals has been very difficult (Thorpe *et al.* 1972; Mann, Dingess & Slater 2006) or even impossible (Elie *et al.* 2010) to investigate. Using modern remote sensing technology, such as microphone transmitters (TerMaat *et al.* 2014; Gill *et al.* 2015) or audio loggers (Johnson, Aguilar de Soto & Madsen 2009; Ilany *et al.* 2013; Anisimov *et al.* 2014; Goldbogen *et al.* 2014; Couchoux *et al.* 2015), vocalisations are recorded directly from the sound-producing animal and can be assigned unequivocally. It is however just as fundamental to retain the precise temporal information of vocal events (Brumm & Slater 2007). This has been achieved through *post hoc* synchronisation (Anisimov *et al.* 2014) or multi-channel recordings (TerMaat *et al.* 2014; Gill *et al.* 2015). Now that it is becoming increasingly feasible

---

*Correspondence author. E-mail: mnadreani@orn.mpg.de

[†]Joint first authors.,

to obtain data with individual information and temporal precision, mechanisms and function of vocal interactions can be investigated with individual-level resolution. In some cases, vocalisations could even be recorded while the animals covered large distances or moved freely in otherwise inaccessible natural habitats (Johnson, Aguilar de Soto & Madsen 2009; Cvikel *et al.* 2014; Couchoux *et al.* 2015). However, it has been shown that animal-attached devices may not only lead to short-term changes in movement patterns (Hooge 1991; Schregardus *et al.* 2006; Chipman *et al.* 2007; Anisimov *et al.* 2014), but may also have longer-term fitness consequences (Phillips, Xavier & Croxall 2003; Barron, Brawn & Weatherhead 2010; Arlt, Low & Pärt 2013). Such results, often coming from the field of movement ecology, suggest that the devices have the potential to influence exactly those behaviours that were studied (Ropert-Coudert & Wilson 2005). In vocal communication research, it would therefore be important to assess the effect of on-board devices on vocal behaviour (Anisimov *et al.* 2014), in addition to more general behavioural patterns such as locomotion. But due to the novelty of the field itself (Wilmers *et al.* 2015), and in view of new exciting discoveries, this and other methodological aspects often receded into the background. So far, most studies did not quantify any direct effects of device attachment on normal behaviour (Hiryu *et al.* 2008; Cvikel *et al.* 2014; TerMaat *et al.* 2014; Gill *et al.* 2015), or neglected to investigate related aspects, such as handling before each recording period (Anisimov *et al.* 2014). In addition, on-board devices may have specific technical requirements that impose further constraints on animal well-being [e.g. frequent handling for data retrieval and/or battery exchange; unstructured environments necessary for infrared-based logger synchronisation (Anisimov *et al.* 2014)]. Because most investigations did not address in detail the implications of a specific approach for the study animals (Ropert-Coudert & Wilson 2005; McIntyre 2015), it is not clear how this could affect the design and outcome of the respective study. Therefore, the current literature is also lacking a thorough discussion of the limitations, of possible applications and of the broader potential of such methods.

Here, we propose a flexible method that could overcome many of such limitations and describe in detail its technical specifications, set-up and workflow. By focusing on the method's direct and indirect consequences on the animals and the research questions involved, we discuss possible applications and the potential for understanding mechanisms and function of vocal communication. In vocal communication research, songbirds have often been used as model organisms. Recording birds with on-board devices represents an especial challenge, because flight and other behaviours can easily be impaired (Pennycuick *et al.* 2012; Vandenabeele *et al.* 2012; Anisimov *et al.* 2014). Therefore, we validate the applicability of on-board microphones for detailed investigations of animal vocal communication, using the example of small songbirds and an improved version of a previously published microphone transmitter (TerMaat *et al.* 2014). This latest version is currently the lightest on-board device for sound recordings (0·56 g, 0·75 g including harness). In our study, we evaluate (i) transient effects on vocal and locomotor behaviour of the initial backpack attachment and of handling associated with battery exchange on zebra finches (*Taeniopygia guttata*); (ii) the system's performance at accurately and selectively recording individual vocalisations, even in loud noise; (iii) as well as its system-based between-transmitter synchrony. In combination with an extended recording duration that outlasts habituation, these features allow tracking individual vocal behaviour, even of small animals, in acoustically and physically complex environments.

Materials and methods

BACKPACKS

The backpacks (Fig. 1) consisted of a circuit board including the transmitter (TerMaat *et al.* 2014) (12·3 × 5·5 × 4·5 mm), a miniature condenser microphone (FB-23359; Knowles, Itasca, IL, USA), as well as a battery (Zn Air, p10; Power One, Ellwangen, Germany) and an established leg-loop harness (Rappole & Tipton 1991). In total, backpack weight was 0·56 g without and 0·75 g including the harness (4·78% of an average 15·7 g zebra finch in our colony). To build the harness, we attached an elastic cord (1 mm diameter, Kordel elastic; Veno Hermann Veddeler GmbH, Bad Bentheim, Germany) to a custom-built silicon casing (Fig. 1c). During attachment, harness length was customised for each bird to decrease the impact (Barron, Brawn & Weatherhead 2010). An earlier study on an electrophysiology transmitter showed that the effects of anaesthesia on singing behaviour and locomotion exceeded the effects of handling (Schregardus *et al.* 2006). Therefore, and because backpack attachment was easy and fast (see Results, Video S1), birds were not anaesthetised.

WORKFLOW

Each transmitter had a specific non-overlapping transmission frequency (270–320 MHz) which was received by its own antenna (crossed Yagi directional antenna for 300 MHz; Winkler Spezialantennen, Annaberg, Germany) and radio receiver (AOR8600; AOR Ltd., Tokyo, Japan, modified to have an audio bandwidth of 12 kHz), and the signal was tracked by an oscilloscope (DSO-X 2004A; Agilent Technologies, Santa Clara, CA, USA). This allowed stable recordings of single channels. To digitise the acoustic signals, we used a multi-channel A/D converter (Fast Track Ultra 8R; M-Audio, Cumberland, RI, USA) connected to a PC. Each of the resulting digitised signals were recorded in parallel using multi-channel software (16-bit, 44 100 Hz; ASIO Rec, adapted by Markus Abels, MPIO Seewiesen). For a flow diagram, see Fig. 4.

ANIMALS

We used a total of 10 adult, parent-raised, domesticated zebra finches (9 males, 1 female) that were kept on a 12-h/12-h light/dark cycle with food and water *ad libitum*. Seven males (aged ca. 1 year) were used for evaluating the effects of backpack attachment and of battery exchange (see 'Effects of backpack attachment and battery exchange') and had been held previously in non-breeding mixed-sex group aviaries. The remaining three birds (2 males, 1 female, and ca. 6 months old) were used to assess the selectivity of sound recordings (see 'Selective sound recordings'). They had

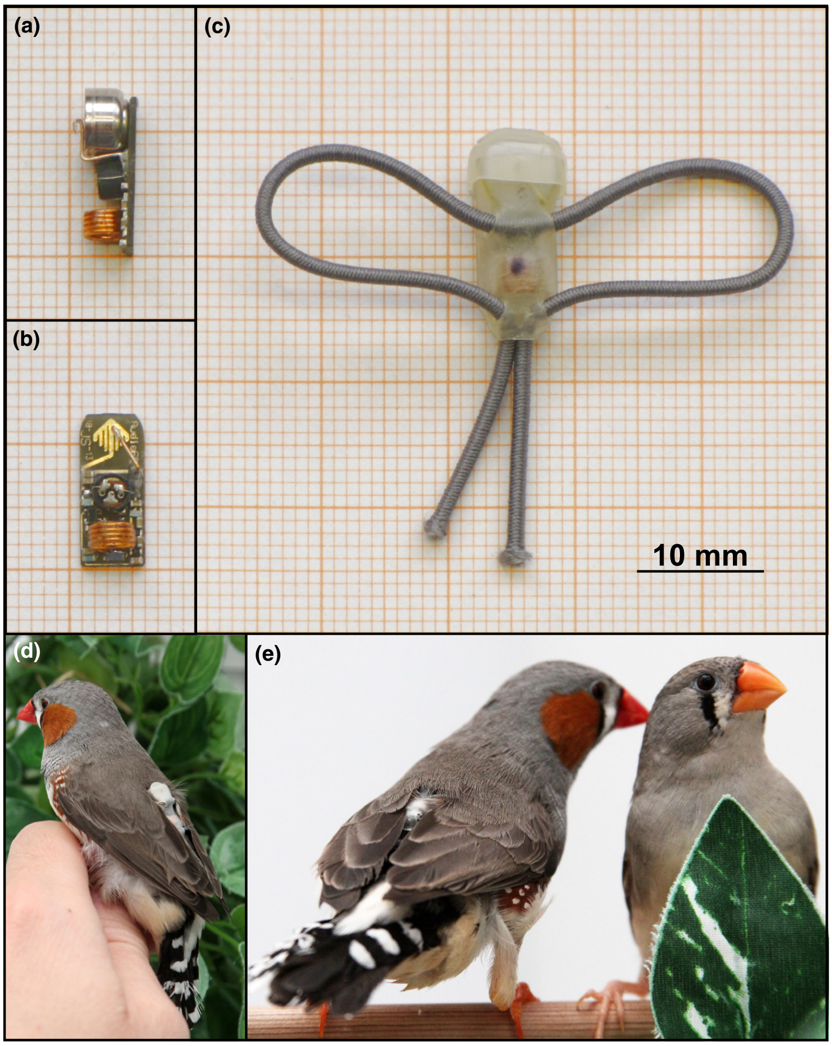


Fig. 1. Transmitter and leg-loop harness. Close-up pictures of microphone transmitter (a) lateral view, including battery; (b) top view, without battery; (c) silicon casing and leg-loop harness; (d) male with backpack exposed; (e) male in normal posture (backpack covered by feathers).

been housed together for 3 months after reaching sexual maturity (as part of a different experiment).

EXPERIMENTS

Effects of backpack attachment and battery exchange

To evaluate direct behavioural effects of backpack attachment on zebra finches, we measured individual calling and locomotor activity in a standardised environment. Because handling is often necessary during an ongoing experiment (e.g. to change batteries or download data), we also quantified changes in these behaviours due to battery exchange.

For this, seven male zebra finches were housed in individual cages (54 × 40 × 28 cm) inside custom-built sound-proof chambers (70 × 50 × 50 cm) equipped with a microphone (TC20; Earthworks, Milford, NH, USA) connected to the sound recording system mentioned above and a video camera (Handykam, Cornwall, UK). Sound and video recordings began as soon as the birds were placed in the sound chambers and continued throughout the day, for 20 days in total. The first 4 h after 'lights on' were used to analyse and track vocal and locomotor activity (data available at http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35). As a proxy for the birds' vocal activity, we used the number of calls recorded by the external microphones during the 4-h periods. Locomotor activity of the birds was measured through a quantification of automated motion-detected multi-channel videos (Surveillance System V8.5.0.0; Geovision, Wagenhofen, Germany). Whenever a bird changed its location in the cage, continuous recording was triggered. If there was no further movement of the bird within 5 s, the recording stopped. The duration of all the videos was extracted (MediaInfo, 0.7.71) and summed up using R (R Core Team 2015) for each bird on each morning and was subsequently used as a proxy for locomotor activity.

Isolating zebra finches impacts vocal activity (Perez *et al.* 2012), but on day 7, the calling activity of all birds reached a stable plateau (<15% day-to-day change for minimum 4 days; mean of all birds) which means this value could be used as baseline (i.e. habituation to cage). On the next morning (day 8), the birds were caught about 20 min before 'lights on' and were equipped with a microphone backpack dummy (same weight and external case). Subsequently, the birds were released back inside their cages, and the above-mentioned sound and video recording scheme were continued. Once vocal activity had again reached a stable plateau (day 15, see above), with no statistical difference to baseline levels (day 7, Fig. 2), we proceeded with the next step of the experiment. To investigate the effect of battery exchange, we caught the birds on day 16 about 20 min before 'lights on', and exchanged the backpack batteries. Subsequently, the birds were

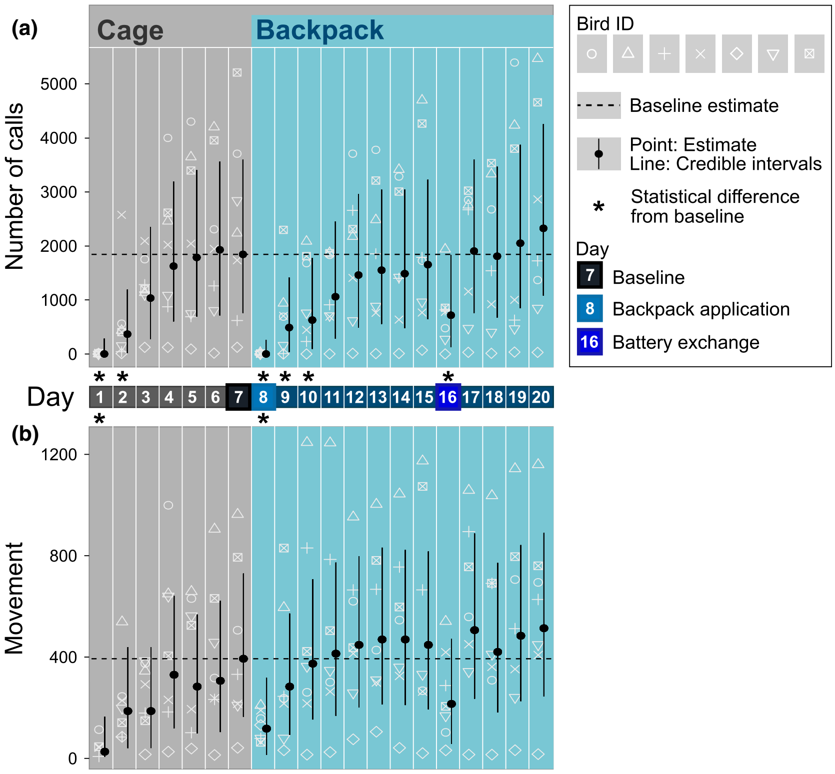


Fig. 2. Calling behaviour (a) and movement (b) before and after backpack attachment. (a) Number of calls and (b) amount of movement (= duration of motion-detected video recordings) for each day before (grey background) and after backpack attachment (turquoise background). Black points and vertical lines indicate Bayesian estimates and credible intervals (CrI). Dashed horizontal black lines show model estimate of calling activity (a) or of movement (b) on day 7 (= baseline). If credible intervals of 1 day do not overlap with this line in the same graph, there is a difference between the amount of calling (a) or of movement (b) compared to the baseline. Data points from different individuals are represented by different white empty shapes (Bird ID). See Tables S1 and S2 for exact fitted values and credible intervals. Data available at http://datadryad.org/review?doi = doi:10.5061/dryad.h8h35.

returned to their cages, and the recording procedures described above were resumed. The experiment ended on day 20 because bird activity levels had returned to baseline values and had not changed significantly during the four previous days.

Selective sound recordings

The aim of the next experiment was to evaluate the performance and selectivity of the backpack recordings in challenging acoustic environments, such as loud noise or the presence of conspecifics. We thus fitted 3 zebra finches (2 males; 1 female for vocal stimulation) with microphone backpacks, and subsequently housed them together in a cage (1 × 1 × 1 m) equipped with a microphone (TC20; Earthworks) and a loudspeaker (KENWOOD KFC-1761S; Kenwood Electronics, London, UK). After a 5-day habituation period, vocalisations were recorded via the backpack microphones as well as by the external microphone for the first 4 h after 'lights on', on two subsequent days, during different conditions. On day 1, we recorded the vocalisations without changing the set-up ('no-noise' condition). On day 2, we recorded the vocalisations during playback of loud, constant white noise ('noise' condition; 80 dB, measured at 1 m from the speaker using an SPL meter, HD600; Extech, Nashua, NH, USA).

Assigning focal vs. non-focal vocalisations. First, we evaluated the system's performance at assigning the vocalisations to the bird carrying the backpack ('focal bird') in the presence of vocalising conspecifics, in both noise conditions. For each male (**♂1, ♂2**), we identified the loudest vocalisation type by comparing waveform amplitudes. Due to the similarity of some song syllables and distance calls, we combined these two categories as 'loudest syllable type'. We counted how many syllables of the 'focal' and of the 'non-focal' individuals had been recorded in the 'focal' individual's backpack. Within each male's recording, we used spectral differences to assign syllables to the 'focal' or the 'non-focal' bird. Backpack vocalisation recordings of the 'focal' bird showed higher amplitudes (Fig. S1) and contained higher power in the low-frequency bands than the 'non-focal' birds (TerMaat *et al.* 2014; Gill *et al.* 2015) (Fig. 3). These vocalisations are individual-specific (each bird has a unique structure, Zann 1984), and therefore, we were able to validate the above described assignment based on spectral differences alone.

Performance of backpack vs. external microphone recordings. Next, to assess the system's efficiency at recording individual vocalisations despite loud external noise, we counted the song events recorded by the backpack and by an external microphone, in the two noise conditions. We assigned the songs to either of the two males based on spectro-temporal differences and used the most characteristic syllable as a proxy for song. Subsequently, we counted the number of song occurrences recorded by the external microphone and by the backpack microphone, for both males.

Synchronous sound recordings

Although the multichannel recording procedure results in synchronised audio recordings (system-based), we performed a simple playback experiment to demonstrate this empirically. Six audio transmitters were distributed equidistantly on a wooden platform, at 80 cm from a loudspeaker (KENWOOD KFC-1761S; Kenwood Electronics). During a period of 24 h, we automatically broadcasted a zebra finch call from the loudspeaker three times per hour at 75 dB (measured at 80 cm from the loudspeaker using an SPL meter, see above) while continuously recording sound via the backpacks. The six resulting sound files were analysed using custom software, following the paradigm described below (see Sound analyses). The timings of call onsets, as recorded by the different transmitters, were plotted against each other (e.g. Fig. 4b,c, data available at http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35).

Battery life

Theoretical battery life of the transmitters was calculated as the battery's rated capacity divided by the device's consumption of current (100 mAh/0·22 mA = 454 h, 18·9 days). For an empirical test, five transmitters were supplied with a new battery (1·45 V, see above) and placed in a room with constant humidity and temperature (45%, 24°C). Their signal reception was checked once a day. Battery life was noted as expired when there was no signal.

Transmission distance

Transmission distance was empirically tested using three devices. It was defined as the furthest point that allowed signal reception and was assessed in three different conditions: outside a building, without (i) and with (ii) an amplifier (Rohde and Schwartz, Berg am Laim, Germany, 20 dB rf amplifier), and inside a building (without amplification) (iii).

ANALYSES

Sound analyses

All sound files were processed and analysed semi-automatically using custom software, to 'cluster' (*k*-means) the individual vocalisations into calls and song (TerMaat *et al.* 2014; Gill *et al.* 2015). In the habituation experiment, we focused on the number of calls per individual (all call types), because, unlike song, calls were produced by the isolated birds in sufficient numbers for a sound statistical evaluation. In the backpack recordings, we measured the mean amplitudes of the background noise and of the loudest syllables coming from the focal and non-focal bird. All amplitude values were obtained by calculating the mean amplitude of the Hilbert transform of a given sound type using the package 'seewave' (Sueur, Aubin & Simonis 2008) in R (R Core Team 2015) (Fig. S1). We used 10 samples for noise and 15 each for focal and non-focal birds' vocalisations. All power spectra and sonograms (Figs 3 and 5) were drawn in R (R Core Team 2015) using the package 'seewave' (Sueur, Aubin & Simonis 2008) (Spectrum size 4096 points, FFT size 512,' hanning' window).

Vocal interaction analyses

For vocal interaction analyses, we used vocalisation recordings from the 2 males and 1 female in both noise conditions (see 'Selective sound recordings', 4 h per bird per day). We focused on the most frequent call, the 'stack' (TerMaat *et al.* 2014). As previously described (Anisimov *et al.* 2014; TerMaat *et al.* 2014; Gill *et al.* 2015), zebra finches are likely to 'reply' to another within 0·5 s. Therefore, we used the number of vocalisations emitted within this narrow time window as the number of 'answers'. Based on this, we drew vocal networks between three interacting individuals (Fig. 5b), with the direction and thickness of the arrows reflecting the direction and the number of answers (ranging from 76 to 645). To visualise in more detail the timing of vocalisations between each set of two birds (data available at http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35), we plotted peri-stimulus time histograms (Abeles 1982) (PSTHs, Figs 5b and S2) in addition to the arrows. In these PSTHs, vocalisation onsets of two birds were aligned and summed up in 20-ms bins. For plotting, all bin values were normalised by the overall maximum value of a single bin (150). The confidence intervals, indicated by horizontal red lines, are the result of a simulation of 1000 calls placed at random times in the recording (TerMaat *et al.* 2014).

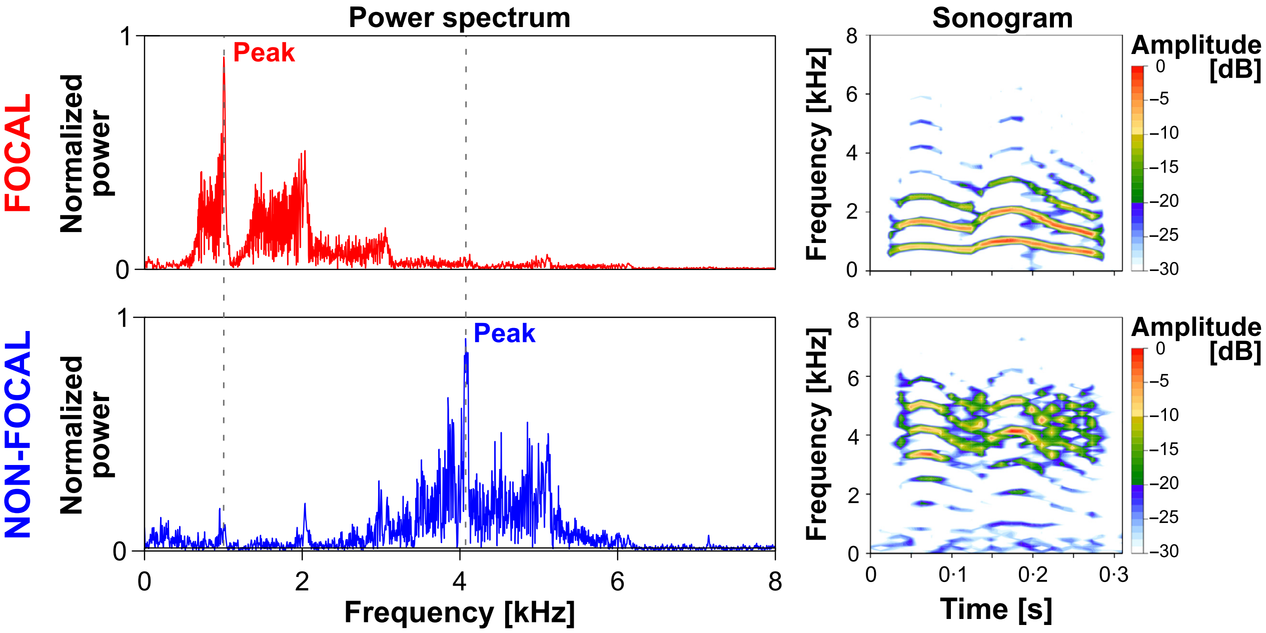


Fig. 3. Spectral analysis of the focal and non-focal recording of a distance call. Power spectra (left) and sonograms (right) of the same male distance call recorded separately by the 'focal' (top row, red) and by the 'non-focal' bird's backpack (bottom row, blue). Both recordings were normalised for spectral analysis. Note that the maximum-power frequency differed greatly between 'focal' and 'non-focal' recordings which allowed an unequivocal identification of the sound-emitting individual. Grey dashed lines indicate frequency of maximum power (peak). Also see Audio S2.

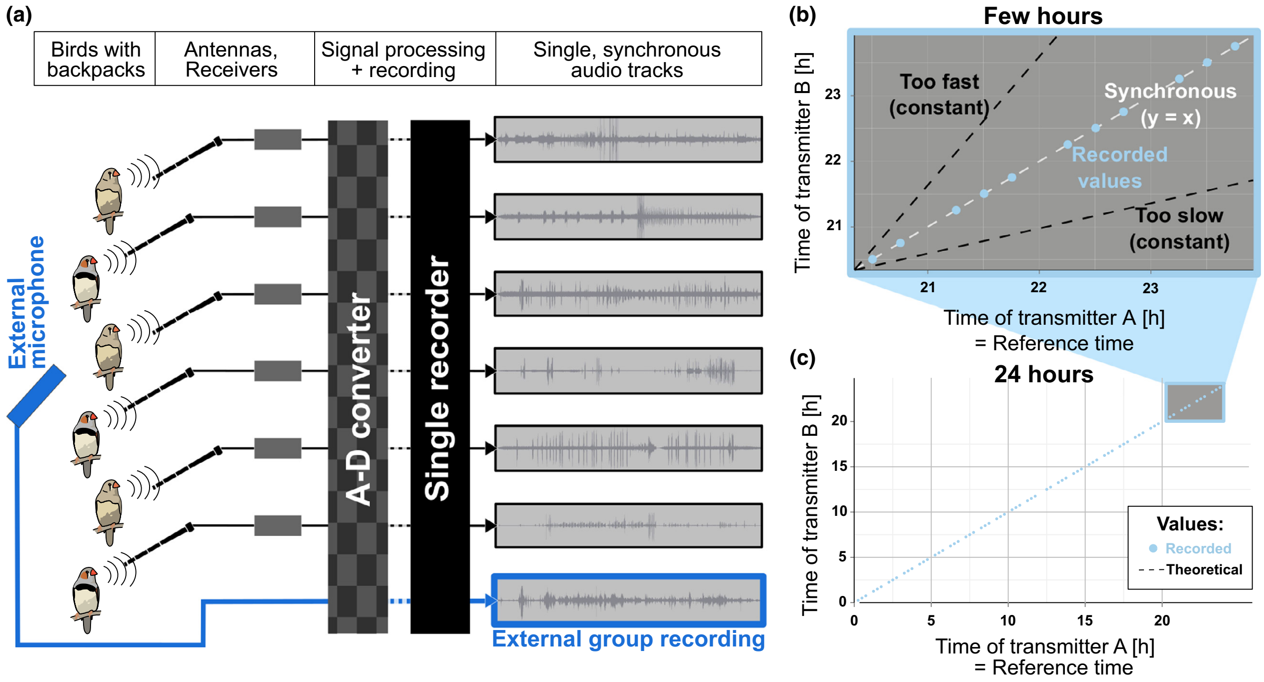


Fig. 4. Synchronous audio recordings: (a) Workflow, (b) 4-h and (c) 24-h test recording. (a) Workflow of recording process: Each group member carries a backpack that transmits acoustic signals via a unique transmission frequency. Signals are received by one antenna and receiver per individual. After digitisation ('A–D converter'), multi-channel software records all audio tracks simultaneously, in real-time. External microphones (blue) may be connected to the system, for example as reference. (b, c) Time [hours] of two transmitters plotted against each other (x-axis: transmitter A; y-axis: transmitter B). Dashed lines indicate theoretical examples of clock drift (black upper: B faster than A; black lower: B slower than A; white: zero clock drift). Light blue dots show real values from a few hours at the end (b) of a 24-h test recording (c). Data available at http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35.

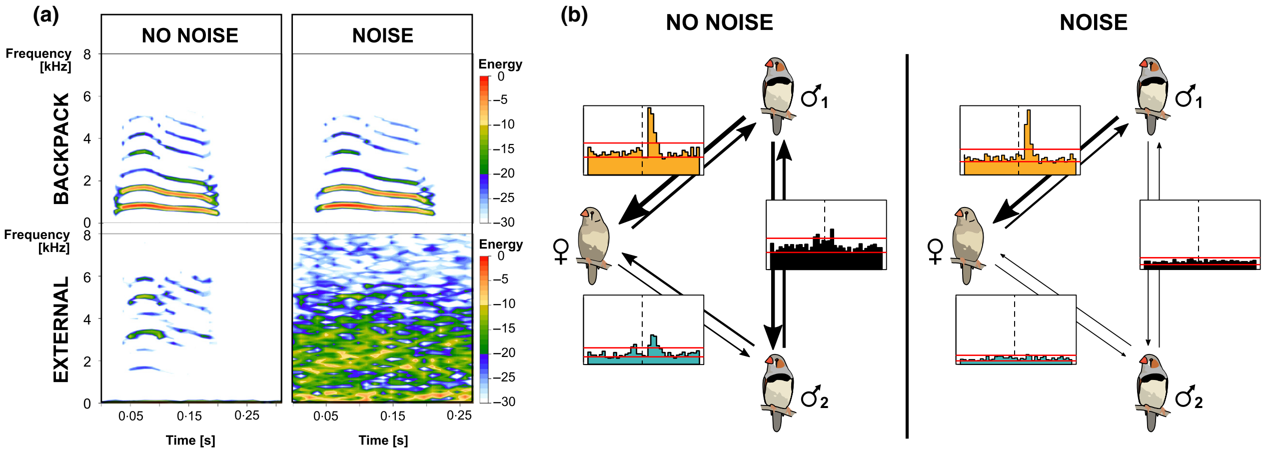


Fig. 5. Vocal communication in noise. (a) Sonograms of a loud syllable in the 'no-noise' (left) and 'noise' (right, 80 dB) condition. Exact same syllables were recorded by the backpack (top) and external microphone (bottom). Note that the noise completely masks the syllable in the external microphone (bottom right), but does not appear in the backpack recording (top right). (b) Vocal network of three birds (1 female, 2 males) recorded via backpack microphones in the 'no-noise' (left) and 'noise' (right) condition. Arrows indicate the direction and line thickness the amount of calls used to answer each individual (see Materials and methods). White rectangles represent peri-stimulus time histograms (see Materials and methods, Fig. S2, Audio S1).

Statistical Analyses

All statistical analyses were performed using R (R Core Team 2015). We modelled the number of calls and of video duration (proxy for locomotor activity, see above) using linear mixed models, specifically applying the function lmer [package 'lme4' (Bates *et al.* 2015)]. Both outcome variables, the number of calls and the video duration, were square root-transformed to approach a normal distribution of residuals. We used the day of the experiment as a categorical explanatory variable. For modelling the number of calls, we used individual id, the centred ($x_{\mathrm{centred}} = x - \mathrm{mean}(x)$) tarsus length and scorer id as random factors to account for between-individual differences in calling activity and size, as well as a possible observer bias, respectively. For the locomotor activity model, we used the same random factors, excluding scorer id (automated scoring). Models including all combinations of random factors were compared to each other using AIC scores

(Akaike 1970). Highest ranking models were the ones that included all random factors. These models were used for subsequent analyses of calling and locomotor activity. For both linear models, standard diagnostic plots were used to assess whether assumptions were met. For the model parameters, we used flat prior distributions. This means the credible intervals (CrI) were equal to confidence intervals obtained by frequentist methods, and sensitivity analyses of prior distributions are not required. To obtain the new set of parameters, we simulated 10000 values from the joint posterior distribution of the model parameters using the function ‘sim’ of the package ‘arm’. The 95% CrIs shown in Fig. 2 were calculated from the 2·5% and 97·5% quantiles of the simulated values as lower and upper limits, and we used the means as estimates (fitted values). We used the derived parameters from the posterior distributions to test specific hypotheses (e.g. differences from the baseline). To calculate the probability of a difference between activity levels on different days, we calculated the proportion of differing simulated values (e.g. number of calls on different days). We considered them to be statistically different from each other (significant) if the posterior probability was larger than 0·95. See Tables S1 and S2 for exact fitted values and CrI plotted in Fig. 2 (negative values set to 0).

Results

EFFECTS OF BACKPACK ATTACHMENT AND BATTERY EXCHANGE

Attaching the established and customised leg-loop harness (Rappole & Tipton 1991) lasted 53·7 ± 18·2 s (mean ± standard deviation, $n = 7$), and birds were able to fly immediately afterwards (Video S1). Handling duration while fitting the backpacks was thus comparable to, or even shorter than, regular bird banding – a standard procedure in ornithology.

Calling behaviour (Fig. 2a, Table S1) strongly decreased immediately after backpack attachment ($(\text{fit}_{\text{day}} - \text{fit}_{\text{baseline}})/\text{fit}_{\text{baseline}}$: by 99·97%; $P < 0{\cdot}001$) and gradually increased from the next day onwards. After 72 h, there was no statistical difference (see ‘Materials and methods’ for definition) between pre- and post-backpack calling activity ($P = 0{\cdot}102$; but remaining 42·5% decrease). Locomotor activity (Fig. 2b, Table S2) decreased immediately (by 70·16%; $P < 0{\cdot}001$), but returned to pre-backpack values after 24 h (27·96% decrease; $P = 0{\cdot}202$).

Capturing the birds and replacing the transmitter batteries (battery exchange: 5–10 s, total handling time including catching: <1 min) after birds had habituated to the backpacks significantly decreased calling only on the day of manipulation (by 61·05%; $P = 0{\cdot}024$, Fig. 2a; next day: 3·34% increase; $P = 0{\cdot}532$; all comparisons with baseline, see Fig. 2a, Table S1). Locomotor activity was not affected statistically but decreased by 45·31% ($P = 0{\cdot}079$, see Fig. 2b, Table S2). Thus, this disturbance was weaker and birds recovered faster than after initial backpack attachment. Additionally, they did not impede successful reproduction. Copulations, egg-laying and parental care were frequently observed (P.B.D’ Amelio & N.M. Adreani, unpublished data). Also, using an older version of the backpacks on previously unmated birds resulted in a mean clutch size of 4·55 eggs (unpublished data from Gill *et al.* 2015) which was similar to the mean in one of the institute’s breeding colonies (W. Forstmeier, pers. comm.).

SELECTIVE SOUND RECORDINGS IN NORMAL AND NOISY CONDITIONS

Bird song can be strongly directional (Brumm & Todt 2003; Brumm, Robertson & Nemeth 2011), and even without noise, the external microphone picked up fewer song events than the backpack microphones. As explained above (see Materials and methods ‘Assigning focal vs. non-focal vocalisations’), this was not due to an incorrect assignment of ‘focal’ birds’ vocalisations (Figs 3 and S1). Instead, this means that backpack recordings give a more accurate representation of vocalisation events than external microphones.

Focal vs. non-focal

Assuming that ‘focal’ birds’ syllables had higher overall amplitudes (Fig. S1) and contained more power in the low-frequency bands than the recordings of ‘non-focal’ birds (TerMaat *et al.* 2014; Gill *et al.* 2015; Fig. 3, Audio S2, see Materials and methods) resulted in 0–10% of ‘non-focal’ syllables recorded by the ‘focal’ backpacks during both noise conditions. Specifically, in the ‘no-noise’ condition, we counted 936 (♂**1**) and 777 syllables (♂**2**) in the ‘focal’, and 92 (♂**1**) and 4 (♂**2**) in the ‘non-focal’ backpack recordings (i.e. 9·8 and 0·5%, respectively). In the ‘noise’ condition, we found 1353 (♂**1**) and 1244 (♂**2**) syllables in the ‘focal’ and 88 (♂**1**) and 0 (♂**2**) in the ‘non-focal’ backpack recordings (9·3 and 0%, respectively). Using the individually distinct syllables for validation gave identical results. The differences in spectral features alone were enough to accurately assign the signaller: none of these syllables were assigned to an incorrect signaller. Therefore, there was no need for any additional device, for example accelerometers (Anisimov *et al.* 2014), to reliably assign the vocalisations.

Backpack vs. external

In the ‘no-noise’ condition, fewer events were detected by the external microphone than by the backpacks. Specifically, we found 492 (♂**1**) and 607 songs (♂**2**) in the external and 525 (♂**1**) and 655 (♂**2**) in the backpack recordings. During ‘noise’, songs were completely masked in the external recordings and were only detectable in the backpack recordings (950 (♂**1**) and 247 (♂**2**) songs, Fig. 5a, Audio S1). Thus, the on-board microphones were more accurate than conventional ones at collecting data from interacting birds across different environmental conditions. Therefore, it is possible to draw vocal networks and investigate vocal interactions with individual-level resolution, even in acoustically challenging conditions (Fig. 5b, Audio S1).

SYNCHRONOUS RECORDINGS, BATTERY LIFE AND TRANSMISSION DISTANCE

The set-up we describe here involved a single, multichannel recording device. This resulted in identical stimulus onsets in different recordings ($n = 6$) (Fig. 4, http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35), that is precise system-based temporal alignment of multiple audio tracks.

Transmission distance ($n = 3$) without using an amplifier ranged from 11·3 to 15·8 m inside and from 11·3 to 13·72 m outside. With amplification, higher values were obtained (25·1–26·7 m).

Battery life ($n = 5$) amounted to 16·2 ± 0·83 days in standardised conditions (see Materials and methods).

Discussion

Using our transmitters, we were able to record multiple vocalising individuals unequivocally, even in noise, without any additional identification technique (e.g. accelerometry in Anisimov *et al.* 2014). For many questions in vocal communication, it is also important to reconstruct the precise timing of vocal events (Brumm & Slater 2007; TerMaat *et al.* 2014; Gill *et al.* 2015). Our approach involved system-based synchrony of multiple audio tracks, thus eliminating clock drift. Also, our microphone backpacks were fast and easy to apply and allowed the birds to fly immediately afterwards (Video S1). From quantifying activity levels, we found that attaching backpacks affected not only locomotor, but also vocal behaviour, that is the behaviour which this methodology was intended to capture. Moreover, vocal activity was also affected by handling necessary for battery replacement or data retrieval (Anisimov *et al.* 2014; Couchoux *et al.* 2015) which, to our knowledge, has not been evaluated before. All measured behavioural effects were only temporary, with vocal and locomotor activity recovering after 72 and 24 h after backpack attachment, and vocal activity recovering within 24 h after battery exchange. Because stress levels are lower in non-isolated zebra finches (Perez *et al.* 2012), it is possible that more naturalistic experimental settings could result in even faster recovery rates.

An ideal way to study the ecology and evolution of vocal communication would be to capture all vocalisations and their drivers in natural settings. But complex social and physical contexts can be acoustically challenging (Brumm 2004; Clemmons & Howitz 2010; Bousquet, Sumpter & Manser 2011), especially when individual-level information is important. Compared to most data commonly collected via remote sensing, acoustic recordings require high sampling rates and thus have high storage and energy demands. Therefore, vocal communication has not been frequently studied using backpacks (Wilmers *et al.* 2015). Recent technology has made previously unexplored questions more accessible by providing data from individuals behaving in more biologically meaningful settings than ever before. Animal-borne recorders (loggers) have proven useful in gaining normalised sound recordings, irrespective of the distance towards any receiving base station (Johnson, Aguilar de Soto & Madsen 2009; Cvikel *et al.* 2014; Couchoux *et al.* 2015). However, they are subject to weight, storage and battery constraints. In addition, whenever multiple devices have internal clocks, their timing eventually drifts apart (clock drift). *Post hoc* synchronisation of audio tracks is possible by using additional signals, such as a master device (Levin *et al.* 2015), GPS (Walker *et al.* 2015) or infrared pulses (Anisimov *et al.* 2014) as an external reference (e.g. Fig. 4b). However, this requires a substantial amount of data processing and comes at a cost of precision. Moreover, such devices could affect potential research questions by increasing animal weight load or by imposing specific technical requirements, for example unstructured environments [necessary for infrared-based synchronisation (Anisimov *et al.* 2014)]. Telemetric devices (transmitters), on the other hand, do not store information on the animals and thus allow extended, continuous recording periods and system-based synchrony while maintaining small size and light weight (TerMaat *et al.* 2014; Gill *et al.* 2015). Therefore, they may not only have smaller direct effects on the animals (Wilson & McMahon 2006), but also allow longer continuous recordings that can bypass handling-related effects. However, their application can be restricted by transmission range. In our set-up, the signals were received up to a distance of 26 m which allows recording animals in large enclosures and even in the field under certain conditions, if an adequate power supply is provided. Specifically, if the position of an animal can be anticipated, it is possible to capture its vocalisations by placing the receiving antenna accordingly, for example close to the nest, roost, lek or colony. Or, if the species allows being approached (Radford & Ridley 2008; Bousquet, Sumpter & Manser 2011), it could even be possible to follow a set of vocalising individuals. Additionally, transmission distance and recording time could be increased if the study species could cope with an augmented weight.

In small animals, particularly in birds, devices need to be light weight and ergonomic to prevent constraints on locomotion and other behaviours (Hooge 1991; Barron, Brawn & Weatherhead 2010; Vandenabeele *et al.* 2012, 2014; Blackburn *et al.* 2016). So far, the effects on vocal behaviour in addition to locomotor activity were reported only once (Anisimov *et al.* 2014), and, compared to in our study, ceased after a longer period of time. This was probably due to a device with higher proportional weight and less ergonomic harness whose attachment required anaesthesia. Especially in the field, it is important to ensure that such devices do not hamper normal locomotion (Barron, Brawn & Weatherhead 2010; Blackburn *et al.* 2016). Our device is currently the lightest on-board microphone method and can be applied to animals as small as hummingbirds (ongoing study on black jacobins, *Florisuga fusca*, pers. comm. Monte & Gahr, 2016). However, deploying observation devices on living animals may influence various aspects of their lives (Hooge 1991; Ludynia *et al.* 2012; Anisimov *et al.* 2014), and weight is not the only property of a backpack that may affect them. Therefore, it is important to pay attention to unexpected complications that may arise. For example, during earlier stages of the device, we noted strong behavioural effects of the protruding antenna. In fact, in a previous study, the antenna had to be incorporated into the harness for females to ensure normal reproductive behaviours (Gill *et al.* 2015), which came at a cost of transmission quality. In the device's current version, this has been overcome by incorporating the antenna wire in the backpack (coil). Thus, our backpacks allow successful reproduction (pers. obs., see Results 'Effects of backpack attachment and battery exchange') and can be used in various complex physical and social settings.

In general, if recordings are made before animals have recovered from backpack attachment, there can be confounding effects in the outcome of an experiment. One way of overcoming this is through a thorough assessment and subsequent exclusion of any affected recordings, which is possible only when recording duration (up to 16 days in our study) exceeds the habituation period (up to 3 days in our study). Such long-term continuous recordings thus further extend the scope of scientific questions that can be addressed by enabling us to track within-individual changes in vocal behaviour of undisturbed animals, for example in response to changing contexts. For instance, it would be possible to investigate with minimal impact the vocal ontogeny of individuals, by following the development of all their vocalisations, even in complex social environments.

Because the multi-channel recording procedure results in perfectly synchronised tracks, it is not only possible to study spectro-temporal features of vocalisations, but also the temporal relationships between vocal events. This method would therefore be suitable for investigating vocal networks of highly vocal, group-living animals (see Fig. 5b). For example, the white-browed sparrow weaver (*Plocepasser mahali*) produces complex duets and chorus song that seem to be linked with social status and pair bond (Voigt, Gahr & Leitner 2006). However, their vocalisations are difficult to assign to specific individuals. Because the singing behaviour is predictable in time and place, this species seems a promising candidate for field studies (ongoing pilot project, pers. comm. Leitner & Voigt, 2016). But fine-scale temporal precision and synchrony are not only important for studying vocal interactions. They are also a prerequisite for other subjects of investigation, such as the coordination of high-speed multimodal courtship behaviours (Bostwick 2003; Ota, Gahr & Soma 2015), echolocation (Cvikel *et al.* 2014), or electrophysiology (TerMaat *et al.* 2014). Because the recording process allows connecting and recording further channels in parallel (e.g. spiking neurons, heart rate), it is possible to investigate the meaning of vocal signals by integrating them into a variety of contexts.

In conclusion, it is important to evaluate potential effects of a given methodology in order to minimise an influence on the study species and ultimately on the research question. Using our method, it is possible to quantify vocal behaviour with individual-level resolution in largely undisturbed subjects. Therefore, challenging aspects of vocal communication can be investigated in small animals, even in small birds, while they behave freely in structured, acoustically complex or changing environments, even in the field. Animal-attached microphones thus allow collecting unbiased vocal information from individuals in previously unexplored situations ranging from large groups to loud anthropogenic noise, which in turn will advance our understanding of mechanisms and evolution of animal communication.

Ethics statement

All housing and experimental procedures were approved by the Government of Upper Bavaria (55.2-1-54-2532-25-09 and 55.2-1-2532-181-12).

Acknowledgements

We are grateful to Lisa Trost and Markus Abels for technical support and advice. We thank Damien Farine and Rudolf Alexander Steinbrecht for useful comments and suggestions. We also thank the IMPRS research school for training and support.

Data accessibility

All processed data used in this publication are provided within the text and in the data repository dryad (http://dx.doi.org/10.5061/dryad.h8h35) (Gill *et al.* 2016).

Author contributions

L.F.G.*, P.B.D.* and M.N.A.* share the first authorship and the order in which the names appear was chosen at random. L.F.G.*, P.B.D.* and M.N.A.* carried out conception and study design. L.F.G.*, P.B.D.* and M.N.A.* performed experiments and analysed results. A.T. and M.C.G. advised on the analyses and on the interpretation of the results. A.T. and M.C.G. conceptualised the wireless microphones and drove their development. H.S provided key technical expertise. L.F.G.*, P.B.D.* and M.N.A.* wrote the manuscript with support from all authors.

Conflict of interest

The authors declare no competing financial interests.

References

Abeles, M. (1982) Quantification, smoothing, and confidence limits for single-units' histograms. *Journal of Neuroscience Methods*, **5**, 317–325.

Akaike, H. (1970) Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, **22**, 203–217.

Anisimov, V.N., Herbst, J.A., Abramchuk, A.N., Latanov, A.V., Hahnloser, R.H.R. & Vyssotski, A.L. (2014) Reconstruction of vocal interactions in a group of small songbirds. *Nature Methods*, **11**, 1135–1137.

Arlt, D., Low, M. & Pärt, T. (2013) Effect of geolocators on migration and subsequent breeding performance of a long-distance passerine migrant. *PLoS One*, **8**, e82316.

Barron, D.G., Brawn, J.D. & Weatherhead, P.J. (2010) Meta-analysis of transmitter effects on avian behaviour and ecology: meta-analysis of avian transmitter effects. *Methods in Ecology and Evolution*, **1**, 180–187.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using "lme4.". *Journal of Statistical Software*, **67**, 1–48.

Blackburn, E., Burgess, M., Freeman, B., Risely, A., Izang, A., Ivande, S., Hewson, C. & Cresswell, W. (2016) An experimental evaluation of the effects of geolocator design and attachment method on between-year survival on Whinchats *Saxicola rubetra*. *Journal of Avian Biology*, **47**, 1–10.

Bostwick, K.S. (2003) High-speed video analysis of wing-snapping in two manakin clades (Pipridae: Aves). *Journal of Experimental Biology*, **206**, 3693–3706.

Bousquet, C.A.H., Sumpter, D.J.T. & Manser, M.B. (2011) Moving calls: a vocal mechanism underlying quorum decisions in cohesive groups. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1482–1488.

Brumm, H. (2004) The impact of environmental noise on song amplitude in a territorial bird. *Journal of Animal Ecology*, **73**, 434–440.

Brumm, H., Robertson, K.A. & Nemeth, E. (2011) Singing direction as a tool to investigate the function of birdsong: an experiment on sedge warblers. *Animal Behaviour*, **81**, 653–659.

Brumm, H. & Slater, P. (2007) Animal communication: timing counts. *Current Biology*, **17**, R521–R523.

Brumm, H. & Todt, D. (2003) Facing the rival: directional singing behaviour in nightingales. *Behaviour*, **140**, 43–53.

Chipman, E.D., McIntyre, N.E., Ray, J.D., Wallace, M.C. & Boal, C.W. (2007) Effects of radiotransmitter necklaces on behaviors of adult male western burrowing owls. *The Journal of Wildlife Management*, **71**, 1662–1668.

Clemmons, J. & Howitz, J.L. (2010) Development of early vocalizations and the Chick-a-dee Call in the Black-capped Chickadee, *Parus atricapillus*. *Ethology*, **86**, 203–223.

Couchoux, C., Aubert, M., Garant, D. & Réale, D. (2015) Spying on small wildlife sounds using affordable collar-mounted miniature microphones: an innovative method to record individual daylong vocalisations in chipmunks. *Scientific Reports*, **5**, 10118.

Cvikel, N., Levin, E., Hurme, E., Borissov, I., Boonman, A., Amichai, E. & Yovel, Y. (2014) On-board recordings reveal no jamming avoidance in wild bats. *Proceedings of the Royal Society of London B: Biological Sciences*, **282**, 20142274–20142274.

R Core Team, 2015 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

Elie, J.E., Mariette, M.M., Soula, H.A., Griffith, S.C., Mathevon, N. & Vignal, C. (2010) Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal Behaviour*, **80**, 597–605.

Furrer, R.D. & Manser, M.B. (2009) Banded mongoose recruitment calls convey information about risk and not stimulus type. *Animal Behaviour*, **78**, 195–201.

Gill, L.F., Goymann, W., Ter Maat, A. & Gahr, M. (2015) Patterns of call communication between group-housed zebra finches change during the breeding cycle. *eLife*, **4**.

Gill, L.F., D'Amelio, P.B., Adreani, N.M., Sagunsky, H., Gahr, M.C. & ter Maat, A. (2016) Data from: A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Dryad Digital Repository*, http://dx.doi.org/10.5061/dryad.h8h35

Goldbogen, J.A., Stimpert, A.K., DeRuiter, S.L., Calambokidis, J., Friedlaender, A.S., Schorr, G.S., Moretti, D.J., Tyack, P.L. & Southall, B.L. (2014) Using accelerometers to determine the calling behavior of tagged baleen whales. *Journal of Experimental Biology*, **217**, 2449–2455.

Hiryu, S., Shiori, Y., Hosokawa, T., Riquimaroux, H. & Watanabe, Y. (2008) On-board telemetry of emitted sounds from free-flying bats: compensation for velocity and distance stabilizes echo frequency and amplitude. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, **194**, 841–851.

Hooge, P.N. (1991) The effects of radio weight and harnesses on time budgets and movements of acorn woodpeckers (El Efecto del Peso de Radio Transmisores y Arneses en la Inversión del Tiempo y los Movimientos de Melanerpes formicivorus). *Journal of Field Ornithology*, **62**, 230–238.

Ilany, A., Barocas, A., Kam, M., Ilany, T. & Geffen, E. (2013) The energy cost of singing in wild rock hyrax males: evidence for an index signal. *Animal Behaviour*, **85**, 995–1001.

Johnson, M., Aguilar de Soto, N. & Madsen, P. (2009) Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: a review. *Marine Ecology Progress Series*, **395**, 55–73.

Levin, I.I., Zonana, D.M., Burt, J.M. & Safran, R.J. (2015) Performance of encounternet tags: field tests of miniaturized proximity loggers for use on small birds. *PLoS One*, **10**, e0137242.

Ludynia, K., Dehnhard, N., Poisbleau, M., Demongin, L., Masello, J.F. & Quillfeldt, P. (2012) Evaluating the impact of handling and logger attachment on foraging parameters and physiology in southern rockhopper penguins. *PLoS One*, **7**, e50429.

Mann, N.I., Dingess, K.A. & Slater, P.J. (2006) Antiphonal four-part synchronized chorusing in a Neotropical wren. *Biology Letters*, **2**, 1–4.

McIntyre, T. (2015) Animal telemetry: tagging effects. *Science*, **349**, 596–597.

Ota, N., Gahr, M. & Soma, M. (2015) Tap dancing birds: the multimodal mutual courtship display of males and females in a socially monogamous songbird. *Scientific Reports*, **5**, 16614.

Otter, K., Chruszcz, B. & Ratcliffe, L. (1997) Honest advertisement and song output during the dawn chorus of black-capped chickadees. *Behavioral Ecology*, **8**, 167–173.

Pennycuick, C.J., Fast, P.L.F., Ballerstädt, N. & Rattenborg, N. (2012) The effect of an external transmitter on the drag coefficient of a bird's body, and hence on migration range, and energy reserves after migration. *Journal of Ornithology*, **153**, 633–644.

Perez, E.C., Elie, J.E., Soulage, C.O., Soula, H.A., Mathevon, N. & Vignal, C. (2012) The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? *Hormones and Behavior*, **61**, 573–581.

Phillips, R.A., Xavier, J.C. & Croxall, J.P. (2003) Effects of satellite transmitters on albatrosses and petrels. *The Auk*, **120**, 1082.

Radford, A.N. & Ridley, A.R. (2008) Close calling regulates spacing between foraging competitors in the group-living pied babbler. *Animal Behaviour*, **75**, 519–527.

Rappole, J.H. & Tipton, A.R. (1991) New harness design for attachment of radio transmitters to small passerines. *Journal of Field Ornithology*, **62**, 335–337.

Ropert-Coudert, Y. & Wilson, R.P. (2005) Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*, **3**, 437.

Schregardus, D.S., Pieneman, A.W., Ter Maat, A., Jansen, R.F., Brouwer, T.J.F. & Gahr, M.L. (2006) A lightweight telemetry system for recording neuronal activity in freely behaving small animals. *Journal of Neuroscience Methods*, **155**, 62–71.

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

TerMaat, A., Trost, L., Sagunsky, H., Seltmann, S. & Gahr, M. (2014) Zebra finch mates use their forebrain song system in unlearned call communication. *PLoS One*, **9**, e109334.

Thorpe, W.H., Hall-Craggs, J., Hooker, B., Hooker, T. & Hutchinson, R., 1972. *Duetting and Antiphonal Song in Birds: Its Extent and Significance*, Behaviour Supplements. Brill, Leiden.

Vandenabeele, S.P., Shepard, E.L., Grogan, A. & Wilson, R.P. (2012) When three per cent may not be three per cent; device-equipped seabirds experience variable flight constraints. *Marine Biology*, **159**, 1–14.

Vandenabeele, S.P., Grundy, E., Friswell, M.I., Grogan, A., Votier, S.C. & Wilson, R.P. (2014) Excess baggage for birds: inappropriate placement of tags on gannets changes flight patterns. *PLoS One*, **9**, e92657.

Voigt, C., Gahr, M. & Leitner, S. (2006) Repertoire and structure of duet and solo songs in cooperatively breeding white-browed sparrow weavers. *Behaviour*, **143**, 159–182.

Walker, J.S., Jones, M.W., Laramee, R.S., Holton, M.D., Shepard, E.L., Williams, H.J. *et al.* (2015) Prying into the intimate secrets of animal lives; software beyond hardware for comprehensive annotation in "Daily Diary" tags. *Movement Ecology*, **3**, 1–16.

Wilmers, C.C., Nickel, B., Bryce, C.M., Smith, J.A., Wheat, R.E. & Yovovich, V. (2015) The golden age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology. *Ecology*, **96**, 1741–1753.

Wilson, R.P. & McMahon, C.R. (2006) Measuring devices on wild animals: what constitutes acceptable practice? *Frontiers in Ecology and the Environment*, **4**, 147–154.

Zann, R.A. (1984) Structural variation in the zebra finch distance call. *Zeitschrift für Tierpsychologie*, **66**, 328–345.

Received 7 March 2016; accepted 12 June 2016
Handling Editor: Robert Freckleton

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. Amplitudes of different sound events recorded by backpack.

Fig. S2. Interaction details from Fig. 5 (PSTH).

Audio S1. Sound files corresponding to Fig. 3a.

Audio S2. Sound files corresponding to Fig. S3.

Video S1. Backpack attachment (real time).

Table S1. Estimated parameters for calling activity.

Table S2. Estimated parameters for locomotor activity.

Supporting Materials

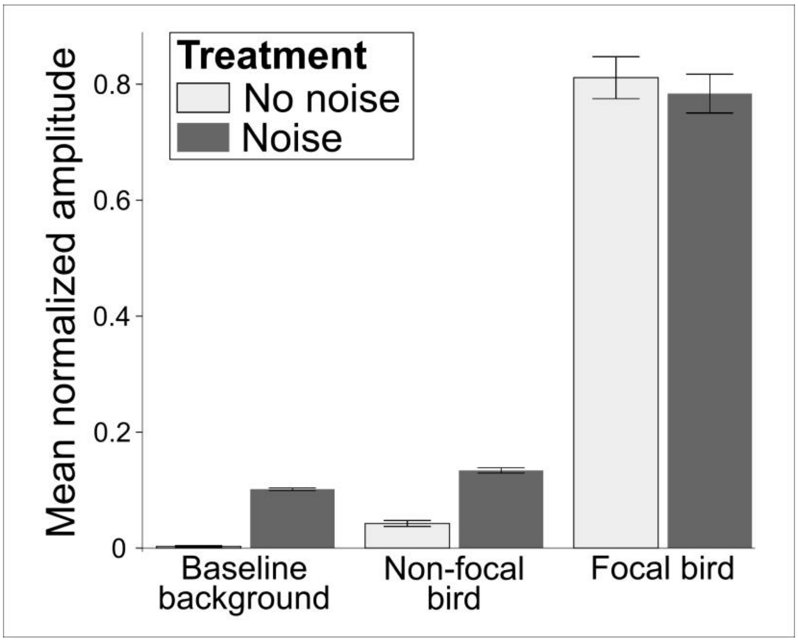


Figure S1: Amplitudes of different sound events recorded by backpack

Mean normalised amplitudes of backpack recordings during “no noise” (light grey bars) and “noise” condition (dark grey bars): Baseline background levels (n = 10) and loudest syllables recorded by the Non-focal (n = 15) and Focal bird’s backpack (n = 15). In both noise conditions, the syllables recorded by the Focal bird’s backpack were much higher in amplitude than those recorded by the Non-focal ones.

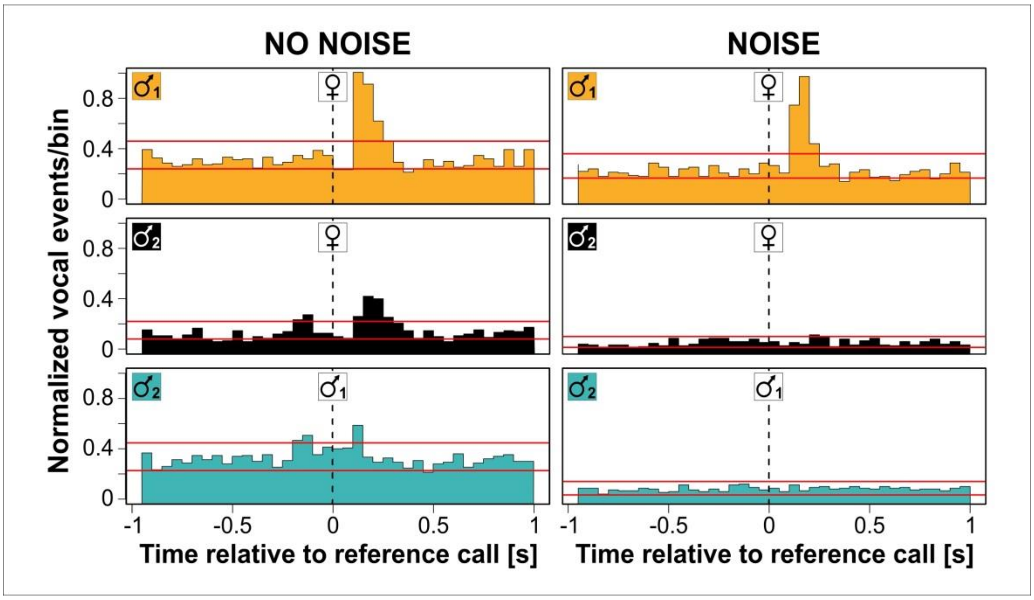


Figure S2: Interaction details from Fig. 5 (PSTH)

PSTHs of the stack-call interactions of three birds in the "no noise" (left) and "noise" (right) condition. In each PSTH, the vertical dashed line represents the onset of the call given by the reference bird (symbol inside white square), and the coloured area represents normalised binned vocalisation onsets of the other bird (symbol inside corresponding coloured square). Red horizontal lines represent the upper and lower confidence intervals (see Methods). Data available at http://datadryad.org/review?doi=doi:10.5061/dryad.h8h35.

Supplementary Video 1: Backpack attachment (real time)

Demonstration of how to attach a transmitter backpack on a zebra finch (real time).

Supplementary Audio 1: Sound files corresponding to Figure 3a

Two calls recorded by backpack and external microphones in different conditions in the following order: Call 1 recorded in “no noise” condition via backpack (i) and via external microphone (ii). Call 2 recorded in “no noise” condition via backpack (iii) and via external microphone (iv).

Supplementary Audio 2: Sound files corresponding to Figure S3

The identical call recorded by the caller’s backpack and by the backpack of the other individual (i.e. as “focal” versus “non-focal” call).

Table S1: Estimated parameters for calling activity

Day	Lower	Upper	Fit
1	0	273.8517	1.047623
2	11.27283	1211.867	368.2638
3	271.0442	2301.281	1036.209
4	591.1981	3146.983	1628.205
5	685.9842	3406.461	1788.999
6	768.2764	3552.053	1930.71
7	724.7044	3452.677	1846.059
8	0	265.655	0.484488
9	36.4427	1443.643	490.1048
10	78.97351	1662.567	628.777
11	260.5525	2349.662	1061.494
12	475.0668	2927.306	1463.353
13	558.8101	3026.157	1553.253
14	498.0024	2952.411	1488.713
15	599.7624	3164.497	1656.169
16	119.5749	1832.445	719.002
17	751.116	3583.635	1907.76
18	700.0914	3393.852	1811.839
19	853.4044	3746.917	2053.378
20	1020.361	4159.362	2328.865

Table S2: Estimated parameters for locomotor activity

Day	Lower	Upper	Fit
1	5.866373	161.356	26.50014
2	42.61392	430.3366	186.7778
3	40.98511	432.8327	186.7778
4	122.9156	641.5026	330.0278
5	94.61405	570.4926	283.3611
6	107.7157	603.1087	306.25
7	159.1418	721.5347	393.3611
8	13.499	324.9939	117.3611
9	94.55955	567.7036	283.3611
10	148.5686	691.938	373.7778
11	175.5833	742.969	413.4444
12	199.6481	792.922	448.0278
13	209.6277	833.1265	469.4444
14	210.0272	829.9171	469.4444
15	197.6279	791.9319	448.0278
16	56.94039	469.3772	215.1111
17	237.8173	879.1853	506.25
18	178.1898	763.7226	420.25
19	223.261	852.2261	484
20	242.4351	882.3769	513.7778

Chapter 2

Individual recognition of opposite sex vocalizations in the zebra finch

Abstract

Individual vocal recognition plays an important role in the social lives of many vocally active species. In group-living songbirds the most common vocalizations during communal interactions are low-intensity, soft, unlearned calls. Being able to tell individuals apart solely from a short call would allow a sender to choose a specific group member to address, resulting in the possibility to form complex communication networks. However, little research has yet been carried out to discover whether soft calls contain individual identity. In this study, males and females of zebra finch pairs were tested with six vocalization types - four different soft calls, the distance call and the male song - to investigate whether they are able to distinguish individuals of the opposite sex. For both sexes, we provide the first evidence of individual vocal recognition for a zebra finch soft unlearned call. Moreover, while controlling for habituation and testing for repeatability of the findings, we quantify the effects of hitherto little studied variables such as partners' vocal exchange previous to the experiment, spectral content of playback calls and quality of the answers. We suggest that zebra finches can recognize individuals via soft vocalizations, therefore allowing complex directed communication within vocalizing flocks.

Published as: D'Amelio, P. B., Klumb, M., Adreani, M. N., Gahr, M. L., & ter Maat, A. (2017). Individual recognition of opposite sex vocalizations in the zebra finch. *Scientific Reports*, *7*(1), 5579. https://doi.org/10.1038/s41598-017-05982-x

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Individual recognition of opposite sex vocalizations in the zebra finch

Pietro B. D'Amelio, Milena Klumb, Mauricio N. Adreani, Manfred L. Gahr & Andries ter Maat

Received: 9 March 2017
Accepted: 20 June 2017

Individual vocal recognition plays an important role in the social lives of many vocally active species. In group-living songbirds the most common vocalizations during communal interactions are low-intensity, soft, unlearned calls. Being able to tell individuals apart solely from a short call would allow a sender to choose a specific group member to address, resulting in the possibility to form complex communication networks. However, little research has yet been carried out to discover whether soft calls contain individual identity. In this study, males and females of zebra finch pairs were tested with six vocalization types - four different soft calls, the distance call and the male song - to investigate whether they are able to distinguish individuals of the opposite sex. For both sexes, we provide the first evidence of individual vocal recognition for a zebra finch soft unlearned call. Moreover, while controlling for habituation and testing for repeatability of the findings, we quantify the effects of hitherto little studied variables such as partners' vocal exchange previous to the experiment, spectral content of playback calls and quality of the answers. We suggest that zebra finches can recognize individuals via soft vocalizations, therefore allowing complex directed communication within vocalizing flocks.

Within a social context, such as a cocktail party, to recognize your spouse's voice might be of survival importance. The capacity to unambiguously identify a conspecific solely by its vocalizations, called "individual vocal recognition", allows for fast turn-taking communication in a crowd[1]. Individual vocal recognition is widespread in the animal kingdom and many mammalian and avian species are able to vocally recognize their partners, kin or group members[2–6]. Previous research on this topic shares a common feature: only the most prominent or common vocalization types of each species were examined. However, several species are known to produce a large set of different acoustic signals, allowing individuals to flexibly convey information in various contexts using a large vocal repertoire. Therefore, to understand the communication processes in a given species the whole of its vocal repertoire must be explored.

Zebra finches are group living, socially monogamous, highly vocal birds[7]. In the wild, in non-breeding condition, they usually form small groups of which the pair is the central unit[8]. In addition to the male song, wild zebra finches use up to eleven distinct call types[7], a repertoire which can also be observed in captive birds[9–12]. Zebra finches are able to coordinate their responses to different call types within fractions of a second[11]. However, it remains unknown how such fast communication is achieved within a group. Specifically, it is not known whether birds can rely on auditory stimuli alone or whether they need other cues to recognize a caller. In a social situation it is predominantly the partner that answers most call types (produced by the mate)[11] but previous research has only tested the males' song, distance calls of adults and begging calls of chicks for recognition of individual identity[13–17], without clarifying how soft calls are responded to in large groups. Zebra finch songs[13] and distance calls[7, 16, 18] can indeed be individually recognized. In playback experiments comparing distance calls of the mate and a familiar individual, adult males and females showed a different vocal response towards of the individuals depending on their familiarity[14, 15]. Building on this playback approach, we tested whether six types of zebra finch vocalizations contained information regarding individual identity: four different unlearned soft calls™stack, kackle, tet, and hat), the distance call and male song. The function of the soft calls is currently being refined, with the tet and stack are social calls often used for within pair communication[10–12], the kackle is a breeding call used at the nest[19], and the function of the hat call is not yet clear, but it is probably an alarm call previously named tuck or thuk[7, 12]. In our experiment, test subjects were presented with playback of three opposite-sex birds differing in their familiarity: (1) their partner, (2) a familiar and (3) an unfamiliar individual. The experiment was replicated on 2 successive days to establish the repeatability of birds' responses. We predicted that if birds recognized individuals vocally, they would answer differentially based on the level of familiarity. We compared the type, number and latency of response to playbacks. Wireless backpack microphones[20] allowed us to unambiguously record single birds during

Max Planck Institute for Ornithology, Department Gahr - Behavioural Neurobiology, Seewiesen, 82319, Germany. Correspondence and requests for materials should be addressed to P.B.D. (email: pie.damelio@gmail.com)

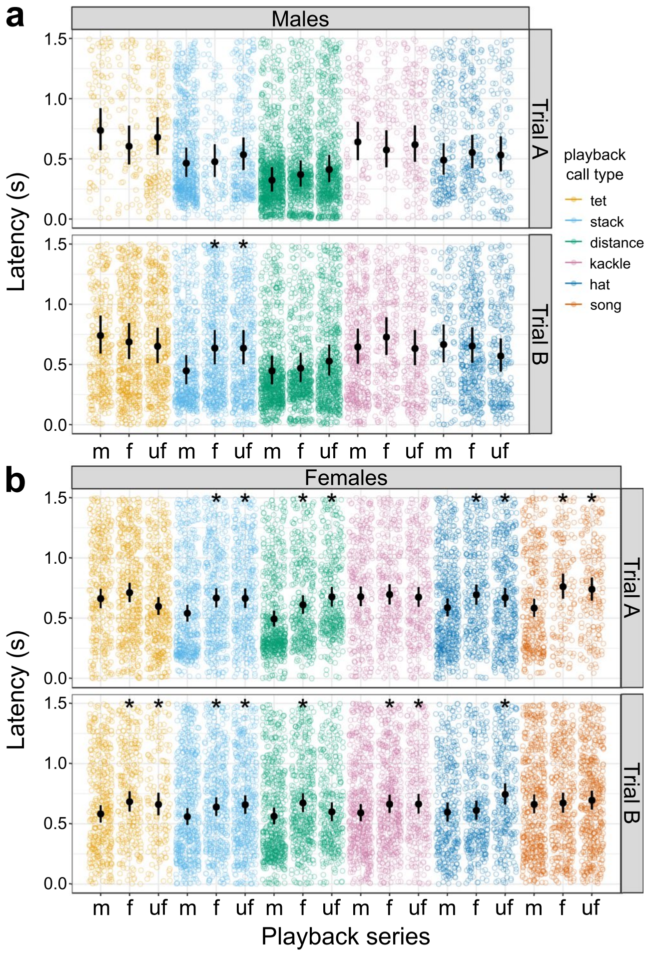


Figure 1. Latency during playback. Latency to the first answering call for different playback series (analysed time interval: 0–1.5 s after the onset of the playback stimulus). Colours represent the type of playback call broadcast; dots indicate individual calls (raw data). Familiarity categories: m = mate of the focal bird; f = familiar individual; uf = unfamiliar individual. For males (**a**) and females (**b**) in both trials the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Significant differences between mate and the other familiarity categories are marked by black asterisks.

the experiment and during couples' normal calling interactions prior to experimentation. Pairs are known to differ in the quality of their bond (fitness)[21] and calling patterns[10, 22], and these characteristics might be related[11]. Therefore, we also examined whether the strength of the relationship before the experiment had an influence on the number of answers during the playback experiment. Finally, we explored whether differences in spectral features of the playback stimuli influenced the observed response. If soft calls could be used to recognize individuals, they would allow birds to address specific individuals in vocalizing flocks.

Results

Call rate throughout the experiment.

First we checked whether the birds habituated to our playback design. We found no consistent decrease in the vocal activity during the experiment (see Supplementary Fig. S1). There were no systematic variations observed in the number of calls; birds continued to use a similar rate of calls throughout the experiment. There was no statistical difference in the number of calls of the first and the last bin for males (p = 0.48) and females (p = 0.10) suggesting no habituation to the experimental design.

Responses to playback: latency.

Playback is an established method to test whether animal vocalizations contain individual identity, and we expected that different levels of familiarity between caller and receiver elicit different vocal responses.

First, we examined whether the latency of response to the playback calls differed by familiarity. In males we found a single significant difference (Fig. 1a, see Supplementary Table S1): individuals answered faster to stack calls of their mate than to those of non-mates during the second trial (trial B, mean ± SD, mate: 0.52 s ± 0.37 s, familiar: 0.65 s ± 0.42 s, unfamiliar: 0.64 s ± 0.41 s; differences between fitted values, i.e. effect size: stack-m – stack-f 0.19 s faster, p = 0.0018; stack-m – stack-uf 0.19 s faster, p = 0.0018). The difference between the two trials was due to a slower response to both familiar and unfamiliar calls in trial A (compare: trial A, familiar:

0.55 s ± 0.39 s, unfamiliar: 0.59 ± 0.41); in contrast, the latency to respond to the mate's playback remained similar in both trials (trial A, mate: 0.55 s ± 0.39 s).

For all female call types we found at least one statistically significant difference for one of the trials (Fig. 1b, see Supplementary Table S2). When answering stack calls, females responded faster to their mate than to familiar and unfamiliar birds in both trials (mean ± SD, mate: 0.59 ± 0.40 (A), 0.65 ± 0.37 (B); familiar: 0.70 ± 0.39 (A), 0.69 ± 0.39 (B); unfamiliar: 0.72 ± 0.39 (A), 0.67 ± 0.38 (B); effect size for each trial, stack-m – stack-f $p = 0.0004$ (A), $p = 0.0142$ (B); stack-m – stack-uf = $p = 0.0005$ (A), $p = 0.0036$ (B)).

Results regarding the song and remaining four calls were equivocal between trials. However, all statistically significant differences in each trial always went towards the expected direction: the focal individual responded faster to the partner than to a familiar or unfamiliar male.

For distance calls, answers to the mates were always faster than to those of familiar and unfamiliar males in both trials (mean ± SD, mate: 0.51 ± 0.35 (A), 0.58 ± 0.38 (B); familiar: 0.64 ± 0.37 (A), 0.68 ± 0.37 (B); unfamiliar: 0.69 ± 0.35 (A), 0.64 ± 0.36 (B)). The latencies to unfamiliar calls in the second trial were the only ones that were not significantly slower than the answers to a mate's calls (effect size, distance-m – distance-f = 0.12 s, $p = 0.0004$ (A), 0.11 s, $p < 0.0001$ (B); distance-m – distance-uf = 0.18 s, $p <= 0.016$ (A), 0.04 s, $p = 0.16$ (B)).

For hat calls, the answers to the mate were also faster than those to familiar and unfamiliar individuals (mean ± SD, mate: 0.62 ± 0.39 (A), 0.63 ± 0.39 (B); familiar: 0.76 ± 0.40 (A), 0.65 ± 0.38 (B); unfamiliar: 0.71 ± 0.41 (A), 0.76 ± 0.41 (B)). All but one comparison were statistically significant (effect size, hat-m – hat-f = 0.11 s, $p = 0.0032$ (A), 0.01 s, $p = 0.3497$ (B); hat-m – hat-uf = 0.08 s, $p = 0.0187$ (A), 0.15 s, $p < 0.0001$ (B)).

With kackle calls, responses to the mate were significantly faster than to both familiar and unfamiliar calls only during the second trial (mean ± SD, mate: 0.75 ± 0.41 (A), 0.67 ± 0.40 (B); familiar: 0.75 ± 0.39 (A), 0.70 ± 0.38 (B); unfamiliar: 0.74 ± 0.40 (A), 0.74 ± 0.40 (B); effect size, kackle-m – kackle-f = 0.02 s, $p = 0.3322$ (A), 0.07 s, $p = 0.0256$ (B); kackle-m – kackle-uf = 0 s, $p = 0.5333$ (A), 0.07 s, $p = 0.0235$ (B)).

The answers to songs of the mate were also significantly faster than those to familiar and unfamiliar songs for the first but not for the second trial, despite trends in the expected direction (mean ± SD, mate: 0.66 ± 0.41 (A), 0.68 ± 0.42 (B); familiar: 0.79 ± 0.41 (A), 0.70 ± 0.42 (B); unfamiliar: 0.78 ± 0.38 (A), 0.71 ± 0.40 (B); effect size, song-m – song-f = 0.18 s, $p < 0.0001$ (A), 0.01 s, $p = 0.3855$ (B); song-m – song-uf = 0.16, s $p = 0.0004$ (A), 0.03 s, $p = 0.2033$ (B)).

For tet calls, there was a significantly faster response to the mate compared with the other two levels of familiarity during trial B (mean ± SD, mate: 0.76 ± 0.41 (A), 0.67 ± 0.40 (B); familiar: 0.75 ± 0.41 (A), 0.77 ± 0.39 (B); unfamiliar: 0.65 ± 0.39 (A), 0.75 ± 0.41 (B); effect size, tet-m – tet-f = 0.05 s, $p = 0.116$ (A), 0.10 s, $p = 0.0048$ (B); tet-m – tet-uf = −0.06 s, $p = 0.9474$ (A), 0.08 s, $p = 0.0196$ (B)).

Taken together, our results demonstrate that females can identify individual identity upon hearing males' stack calls. Although inconclusive, our results suggest that partners can be distinguished from individuals of differing familiarity also by the remaining call types.

Response to playback: number of calls per playback series. Second, we considered the total number of calls emitted during each playback series. Overall, the number of answers did not differ between the two trials (probability trial B > trial A; $p = 0.60$ for females and $p = 0.75$ for males).

In males, during trial A, the number of answers elicited by stack calls depended strongly on their familiarity. Male zebra finches responded with a higher number of calls to the stack calls of their pair-bonded females (mean ± SD of number of calls, 171.8 ± 152.9) than to those of a familiar (29.7 ± 22.1; probability stack-m < stack-f = 0.0003) or unfamiliar female (69 ± 52.6, probability stack- m < stack- uf = 0.0064; Fig. 2a). No other comparison reached statistical significance in males. In females, the number of answers showed no strong relationship with the level of familiarity of the caller (Fig. 2b). These results suggest that males can identify individual identity by hearing females' stack calls.

Type of answering call. Males and females responded to playbacks using different call types. Females predominantly responded with stack calls (69.2% of the total answers), followed by distance calls (22.4%) (see Supplementary Fig. S2). In males we found a similar pattern, the majority of their responses being stack call (50.1%), followed by distance calls (17.6%) and hat calls (11.8%) (see Supplementary Fig. S3). Not only the number of calls or their latency, but also the quality of the answer (type of call used to answer) might differ when responding to different familiarity levels. However, we did not find any such differences (see Supplementary Tables S3 and S4 for females and males, respectively). Thus, we conclude that the level of familiarity did not influence the type of response to playback.

Habituation within playback series. Because the specificity of the answers for the different familiarity levels could change during the experiment due to behavioural habituation, we compared results obtained during the first and last 30 playback calls of each series with those of the complete series (300 calls). Regarding the number of calls, in males, we found only two statistical differences between the two datasets for both first and last playback calls (see Supplementary Figs S4a and S5a) which were in the expected direction, there were more calls in response to the mate compared with the other familiarity levels. In females, we also found only three differences between response to the first 30 playback calls versus the complete dataset, and only one difference between response to the last 30 playback calls and the complete dataset (see Supplementary Figs S4b and S5b). Thus also for females we found limited differences between datasets and all in the expected direction with more answers to mates. Interestingly, considering only the first playback calls in the first trial we found more answers to mates compared to both of the other levels of familiarity when answering male songs.

There were very few differences in response to experimental conditions over time when considering the latency of response to playback when comparing the full dataset to the first 30 calls (see Supplementary Fig. S6

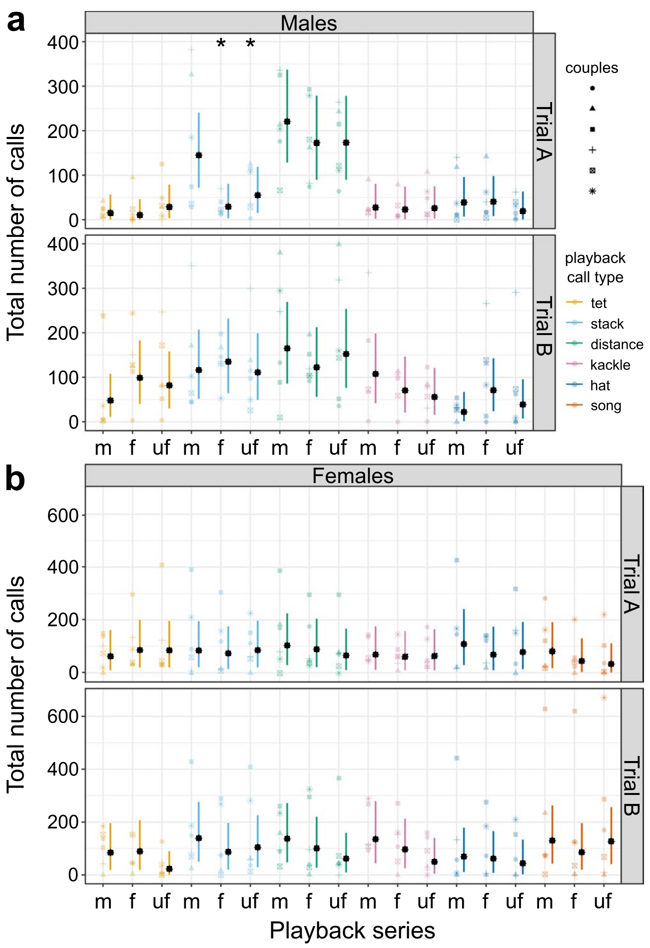


Figure 2. Number of calls during playback. Number of answering calls that focal individuals emitted during the different playback series. Raw data for males (**a**) and females (**a**) and both trials (symbols indicate responses of individual birds) and the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Colours represent the type of playback call broadcast. Familiarity categories are indicated by letters: m = mate of the focal bird; f = familiar individual; uf = unfamiliar individual. Significant differences between mate and the other familiarity categories are marked by black asterisks.

and Tables S5 and S6) and to the last 30 calls (see Supplementary Fig. S7 and Tables S7 and S8). Only when considering the 30 call subsets, males showed significant differences in the direction opposite to our expectations in kackle and hat, i.e. slower answers to the mate than to non-pair-bonded females. This might be due to the rarity of elicited answers and the absence of a real pattern leading to false positive results.

For females, for which we showed that latency was very important (Fig. 1b), there was only one case out of 24 for each first and last 30 playback calls dataset where the direction of the difference opposed our expectations. The majority of differences between latency to respond to the mate versus other familiarities observed in the full dataset were also present in the subset data, especially in the first trial (A) (see Supplementary Figs S6b and S7b); a change in response to experimental treatment over time occurred during the last 30 playback calls of the second trial (B) were most of the differences with the full dataset are concentrated (see Supplementary Table 8). These results confirm that the quality of the answer did not change much throughout the experiment and that longer playback series produced more reliable results.

Relation between calling behaviour during baseline and experiment. We investigated whether the calling relationship of mates before the experiment influenced the number of answers during the experiment. We first asked which call combinations showed repeatable patterns during the baseline period and, therefore, might result in a predictable answering rate during the playback experiment (see Supplementary Table S9). We therefore considered only the stack-stack exchanges whose percentage of answers was consistent during baseline recording for males (repeatability = 0.995 + 0.004) and females (repeatability = 0.83 + 0.13). Subsequently, we correlated the percentage of answers during the playback experiment (separately for trial A and B) with the percentage of answers during the baseline (mean of both days). There was no relationship between baseline and experiment in females (trial A p = 0.7905, trial B p = 0.4784) but we detected a negative relationship in males (trial A p = 0.0325, trial B p = 0.0523) (see Supplementary Fig. S8). The relationship between response during baseline and experimental periods in males demonstrates that vocal performance during the experiment is related to preceding vocal relationships.

How the variability of call type spectral features is related to the variation in conspecific response. Finally, we identified the most individually distinct call types and whether the individual variability of each call type was a good predictor of the conspecific response (quantified as the number of calls and latency to respond). We aim to demonstrate that what we interpret as recognition is not a by-product of easier discrimination. According to our analysis the hat was the most distinct call type for both males and females. The stack call, although unambiguously recognised by both males and females, was not the most distinct call type (see Supplementary Fig. S9a). Furthermore, the magnitude of the response (response to the mate – response to the familiar/unfamiliar) was not correlated with the index of individual distinctiveness (see Supplementary Fig. S9b). Because the most individually distinct call types were not the ones recognised best, we demonstrated that there was recognition beyond discrimination.

Discussion

Females and males unambiguously showed individual recognition of stack calls produced by the opposite sex. We found a differential response to distinct familiarity levels, both in the number and timing of the response. Females only used timing to demonstrate recognition: they vocalized at similar rates when responding to the playback of different familiarity levels, but responded more quickly to the calls of their mate versus non-mates. Intriguingly, males used multiple strategies in different trials to demonstrate recognition: a higher number of calls in response to mates during the first and a shorter latency to respond during the second. Furthermore, in females, consistent differences between answers to their mate and those to at least one other level of familiarity were detected for two call types (hat and distance call). Only latency and number of calls regardless of type differed between familiarity levels.

Until recently, soft calls have been considered "a background hum in which other calls are embedded (…), not directed at specific individuals and do not stimulate specific replies"[7]. In contrast to this view, growing evidence demonstrates that soft calls are indeed directed at individuals and can elicit specific replies[10, 11, 22]. In group contexts, addressing specific subjects is a prerequisite for effective communication and we now provide the missing link explaining how this is achieved: we show that soft calls can be assigned to individuals and that the latency of an answer can provide specific information. For stack calls, we estimated a difference between the answer to the mate and familiar or unfamiliar individuals of approximately 188 ms for males and 125 ms for females. This delay is roughly similar to the mean latency of calls used as replies and double the length of stack calls[10]; therefore, this response gap can be biologically relevant and directly used within a communicating group. Individual vocal recognition using contact calls has rarely been investigated in Passeriformes although different functions have been proposed: for example, Large-billed Crows (*Corvus macrorhynchos*) can recognize strangers' loud calls[23], Long-tailed bushtits (*Aegithalos caudatus*) kin's contact calls[6], Chestnut-crowned Babblers (*Pomatostomus ruficeps*) group members' contact calls[24] and Silvereye (*Zosterops lateralis*) mates' calls[25]. However, in most of the studies it was unclear whether these calls were learned and a large proportion of the typical repertoire remained untested. Therefore, more research is needed to identify common themes in the evolution of recognition of soft vocalizations and to establish whether addressing specific individuals in a vocalizing group is common among Passeriformes.

Females tended to respond differently to their partner's vocalizations in most call categories tested, whereas males' answers only differed when responding to stack calls. We cannot yet explain this result, but differential discrimination abilities[26] and sex-specific roles in the communication process have been proposed[27]. Despite several lines of evidence indicating that females might be able to recognize individual identity from all soft call types, we could only confirm this for hat and stack calls. This may be due to the differing functions of specific call types. For example, recognition of the hat call might be important in identifying the alarming bird, a colony member or an external individual, whereas recognizing the stack call may serve to maintain vocal contact with a mate. Tet and kackle calls, in contrast, are part of the private communication occurring at the nest where other individuals are not present and thus, individual recognition may be less important[12, 19]. Breeding calls, such as the kackle, become more common once a couple is nesting[11]. The stack call, on the other hand, is one of the most common call types in non-breeding groups[10, 11], a situation which resembles the context of our playback experiment, which might explain why stack calls were promptly recognized in our study.

Notably, our results entail a process of comprehension learning of the stack call, which is important because comprehension learning is a prerequisite for the evolutionary origin of vocal learning[10, 28]. Soft calls are generally used when partners are in close proximity of each other – and therefore see each other[7, 12, 19]; hence cues other than acoustic modalities are available to facilitate individual recognition. Therefore, identifying individuals is indeed not the sole intention of these vocalization types; the encoded identity can be used in communication between specific individuals in a group. Our results suggest that birds are integrating information about call type and call identity to tailor vocalizations and provide the correct answer type and time. Previous observations and multiple independent lines of evidence led us to postulate that our findings agree with the hypothesis that vocal learning is driven by social complexity[29]. Learning acoustic parameters is a precursor for any subsequent learned modification of the spectral features of a vocalization[30]. Moreover, soft calls are encoded in a high order telencephalic nucleus of the motor pathway[10], which is fundamental for the control of learned vocalizations, and may facilitate coordination of communication[31, 32]. Therefore, although vocal recognition is present in many vocal non-learners[5] and comprehension learning may just be a prerequisite rather than a driving force, we suggest that unlearned calls in vocal learners might provide a model to better understand the origin of vocal learning capabilities.

Until now, only part of the repertoire of the zebra finch had been tested for individual recognition. As for many other Passeriformes[33], song has repeatedly been shown to contain individual characteristics that can be used for identification[7, 27, 34]. In most cases this has been proven in simultaneous choice tests[13, 35, 36]. In contrast, higher vocal response towards the partner's song was only reported once and exclusively when vocalizations emitted towards the speaker were taken into account[36]. Our results regarding song are equivocal; females showed a differential response during their first trial, confirming previously published results, but this response did not hold during the second trial, which could partially be explained by habituation to the experimental design for this vocalization type. In addition to song, several studies investigated the distance call of both sexes, attempting to assess whether these vocalizations contain individual information and whether this information is used for identification[13–15, 37]. In our experiment, we found sex-dependent behavioural responses to recordings of distance calls from individuals of different familiarity. Female zebra finches displayed more and faster responses towards their mate's calls than towards familiar and unfamiliar distance calls, thereby confirming previously reported discrimination capabilities[14]. Conversely, in males, no such differences were found, despite trends in the expected direction. Our results thus diverge from previous studies, which demonstrated a significantly higher number of answers during the replay of the mate's than of familiar distance calls. However, in previous studies only distance calls were regarded as answers[15] or only considered the neuronal response in males' high-order auditory areas (Caudomedial Mesopallium, NCM)[38]. Furthermore, the difference between our and previous studies might be due to differences in the selection method for playback stimuli. We extracted distance calls from the normal communication flow, i.e. calls uttered when both partners remained in the same sound-proof box. In contrast, previous studies used provoked distance calls elicited by visually separating the individuals[14–17, 39, 40]. Because social context can influence the acoustic structure of a bird's call[40], it is possible that provoked calls emitted by birds in isolation exhibit enhanced call urgency in order to initiate contact with their partner. This call urgency might in turn increase the motivation of focal subjects to respond to these calls in a playback setup, possibly explaining why males in previous experiments showed a higher vocal response to their partner's calls than in our study.

We used natural rather than synthetic vocalizations for playback to ensure that stimuli contained all necessary acoustic structures, as altered call perception is possible when using artificially created vocalizations[41, 42]. Additionally, we employed long playback sessions to increase power for our analyses. Most playback studies use very few calls to avoid habituation. Instead, we attempted to mitigate habituation by continuously varying the interval between successive calls. Despite the long duration, calling rate did not decrease during the experiment or during the single series, indicating that the birds did not habituate to the playback. Additionally, differences between the results obtained for the first 30 playback calls compared to those for all 300 calls were negligible. When comparing the last 30 playback calls to the entire dataset the only noteworthy differences occurred during females' second trials which may indicate a certain degree of habituation. This is remarkable because we found that low numbers of answers (e.g. familiar kackle and hat in males, see Supplementary Figs S6a and S7a) actually increased improbable and false positive results. We repeated the entire experiment on two consecutive days to assess whether birds habituated to the playback design. Indeed we found differences between trials, but not concerning the stack call, which was always answered differentially according to familiarity level. The differences between trials are difficult to interpret and should be considered when planning experiments that contain multiple presentations of the same stimulus. Finally, we did not observe an effect of social context required for answer specificity[14, 15], the audience pairs, as their calls did not influence the results directly. Specifically, backpack microphones worn by the focal birds assured the individuality of the recordings[20], and the effect of the audience, quantified in our models, was limited.

Stress is also a possible confounding factor in our experimental setup. Notably, corticosterone levels in zebra finches increase 24 hrs after the separation of an established pair[43]. Moreover, these hormones are associated with a reduction in vocal discrimination ability[40]. Therefore, mate separation before and during the playback experiment might have increased stress levels in the test subjects, thereby impairing their discrimination abilities. Although we endeavoured to reduce the stress for focal birds by limiting the separation period to one night before the first experimental trial, the different personalities of the test subjects may have led to differences in stress response[44]. In addition, the quality of the pair bond itself might influence the level of stress birds experience when separated from their mate; couples sharing a stronger bond might be more strongly affected by separation than those having weaker pair bonds. This might partially explain why in males, which are highly repeatable in their response[22], we found a negative relationship between the proportion of answers during playback experiments and the baseline. Unfortunately, the small sample size of our study makes it difficult to generalize these findings; however, the correlation with measurements of pair strength is worth further investigation.

Vocal individual recognition of the so-called "soft calls" has not previously been tested in zebra finches; we provide the first evidence that at least one of these call types, the stack call, contains individual identity despite not being more individually distinct than other soft calls. This finding implies that soft unlearned acoustic signals are sufficient to determine a caller's identity and that visual cues are not required. We have identified the mechanisms underlying how birds vocally interact in a group. Namely, employing differential latency times when answering to different subjects allows a caller to address individuals specifically. Vocal recognition is a fascinating aspect of vocal communication because the ability to recognize individuals in a group of vocalizing conspecifics is a prerequisite for complex communication networks.

Material and Methods

Ethics statement. The use of audio transmitters, bird housing in sound-proof boxes, and conduct of all other experimental procedures were approved by the government of Upper Bavaria (record number 55.2-1-54-2532-21-2015). All further animal husbandry and handling was conducted according to the directives 2010/63/EU of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes.

Animals and housing conditions. A total of 12 adult zebra finches, six pairs, served as the focal birds for the experiment, plus 14 additional birds which served as the audience (i.e. as company for focal birds)[15]. All pairings were "forced", i.e. couples were formed by randomly selecting unrelated individuals from the breeding facilities of the Max-Planck-Institute for Ornithology, Seewiesen, Germany. Birds were kept in a 13/11 Light/Dark cycle, at 24 °C and 60–70% humidity. Food (mixed seeds, and "egg food"), fresh water and cuttlebone were provided *ad libitum*. We performed all experiments with birds from forced pairs in the non-breeding condition. All couples had been together for at least six months, raised at least one brood and had been housed without nesting material for three months prior to the experiment. Zebra finch couples were housed in single pair-cages (123.0 cm × 37.0 cm × 38.5 cm) in two separate rooms with three experimental couples per room. Within each room couples could see and hear each other, whereas there was no acoustic or visual contact between pairs housed in different rooms, making these two groups "unfamiliar" to each other. Experimental pairs were housed with other breeding pairs, seven of which served as "audience couples" during the playback experiments.

Experimental timeline, sound recording and playback. Zebra finch couples were moved to sound-proof boxes one week before the experiment to allow for acclimatization to the new conditions. Animals were equipped with custom-made light-weight (less than 5% of average body weight) wireless microphone transmitters fitted on their back via a leg-loop harness as previously described[20] which recorded continuously throughout the experiment. To determine the vocal relationship between males and females[10, 22], we audio recorded each pair for three consecutive days, using the recordings of the first and third day as the baseline for subsequent analysis of calling patterns. Audio was scored for four hours a day (12:00 to 16:00). Each sound-proof wooden box was equipped with a general microphone (TC20; Earthworks, USA) which was used to extract playback stimuli.

To create the familiarity level "familiar" (equivalent to a group member in the wild), we moved a couple from the same housing facility as the focal pair into their cage during the evening of day 3 (end of baseline recording). The two pairs shared a cage for approximately 24 h, separated by a wire mesh allowing acoustic and visual interaction. Their calls served as "familiar stimuli" during the playback experiments[37, 38]. During the night of day 4, the "familiar" couple was removed and the male and female of the focal couple were separated. Each focal bird was placed with an unfamiliar, established pair: the "audience couple". This audience provides a social context that increases the specificity of the answers[14, 15] and prevents social isolation[40]. Experiments were carried out during the morning and afternoon of days 5 and 6, resulting in one trial per bird per day. The time of testing (morning/afternoon) was randomized, and audience couples were changed between trials (i.e. in the evening of day 5).

The replay of calls was controlled via a computer connected to an amplifier (CS-PA1, SINTRON Vertriebs GmbH, Germany), and calls were broadcast via a loudspeaker (KFC-1761S, Kenwood Electronics, UK) placed at the back wall of the sound-proof box. The sound level of the experimental signals, measured at a distance of 1 m from the loud speaker (Sound Meter, Model HD600, Extech Instruments, U.S.A), was adjusted to a peak value between 50.03 dB ± 0.87 dB (mean ± SD; minimum for the lowest call type, the tet) and 74.05 dB ± 1.15 dB (corresponding to the loudest call type, the distance call reflecting a typical level of a natural distance call)[39], and was constant for all three familiarity levels of each call type.

The playback stimuli and the focal bird's calls were recorded (via external and backpack microphone, respectively) synchronously on separate audio channels for subsequent alignment. Each subject was presented with calls of three different individuals of the opposite sex, representing three familiarity levels: "mate" (m; partner of the focal individual), "familiar" (f; known individual), and "unfamiliar" (uf; unknown individual). Six different vocalization types were used for playback experiments: tet, stack, distance call, kackle, hat, and - in the case of females - song (for the original audio files see additional information). Playback calls were extracted from the general microphone recordings during the acclimatization phase and baseline period. In rare cases in which birds did not emit a specific call type during the sampling period, this type was omitted from the playback. Each vocalization was high-pass filtered (freq = 85 Hz), its amplitude normalized to 0.1 dB (maximal sample value) and the stimulus faded in and out to avoid rapid amplitude changes. Playback calls were presented in blocks, each block consisting of three series of playback calls of the same call type. Each series consisted of calls of an individual representing one of the three different familiarity categories (m, f, and uf). We used three randomly selected vocalizations of each call type per individual in order to mitigate pseudo-replication[45]. Within a series, each playback call was repeated 100 times for a total of 300 calls. To ensure that the playback was unpredictable, the inter-call intervals were changed randomly at each emission (within 2 ± 0.5 s, uniform distribution). Playback series within a block were interspaced by 70 ± 10 s of silence, and different call-type blocks by 130 ± 10 s of silence. The total duration of an experimental trial was approximately 2:45 h for males (15 series) and 3:30 h for females (18 series). The order of call-type blocks and of familiarity categories within blocks, as well as the order of single calls within series, were determined semi-randomly.

Sound analyses and sorting of vocalizations. Sorting of vocalizations from audio files proceeded as previously described[10, 20]. Briefly, sounds that exceeded a manually set amplitude threshold were extracted for further analysis. Using custom software written in Delphi Pascal for Windows (SoundExplorer; R. F. Jansen, MPIO, 2000; see ref. 11 for GitHub address), the following parameters of each sound were computed: average frequency, modal frequency, fundamental frequency (first peak), Wiener entropy, duration, and their standard deviations. The subsequent clustering process was based on a k-means clustering algorithm[46]. After noise detection and elimination, the results were refined manually: each cluster was checked and sorting errors were corrected, resulting in a separate cluster for each call type of a bird's repertoire[10, 11]. Information about the call type and the timestamp was saved for each vocalization. We used this information to determine the temporal relationship between all possible call type combinations of mates during the baseline[10, 11] and during the playback experiments. During the baseline period we considered all calls emitted within 0.5 seconds from the partner's call as answers. For each combination of call types, we calculated the number of answers and the proportion of answers from the total amount of emitted calls of that specific type (see ref. 22 for details on calculations).

How the variability of call type spectral features is related to the variation in conspecific response. We then explored whether the most individually distinct call types were the ones that were easier to discriminate. We extracted 14 acoustic features[12] of each call sonogram tested as playback stimuli (i.e. for each call type three calls for individual of each familiarity). which were then used to conduct principal component analysis (PCA)[47] for each call type. We extracted the mean frequency and its standard deviation, median frequency, skew, kurtosis, spectral flatness measure, entropy, mode frequency, frequency precision of the spectrum, peak frequency, fundamental frequency, dominant frequency, maximum dominant frequency, and duration (R package "seewave")[48]. Subsequently, we used the first two principal components as explanatory factors in a linear discriminant analysis (R function "lda")[49] with familiarity level as a predictor. We used the parameters obtained to predict (R function "predict")[49] the proportion of cases in which the calls were assigned to the correct familiarity. We ran this analysis for each focal bird and we present the average and SD of the incorrect assignments (see Supplementary Fig. S9). Call types in which the individuals are very distinct will have a lower proportion of incorrectly assigned calls. Lastly, we used the within call type variability as a predictor of the magnitude of the response to the conspecific (number of calls and latency, response to the mate – response to the familiar/unfamiliar – for the latency the average for each individual was taken) in a linear mixed model with random factors as described below.

Statistics. All statistical analyses were performed using R[50] in a Bayesian framework. We used linear mixed-effect models[51] to analyse the effect of different playback stimuli on the total number of response calls emitted during each playback series, and on the latency of the first answering call for each playback series (upper limit 1.5 s) separately by sex. Before interpreting the results, we checked whether model assumptions were met by inspecting the residuals for normality, homoscedasticity, and lack of remaining pattern. Both the total number of responses and latency were square root transformed to approximate normality. Three categorical variables served as fixed effects in both models: familiarity (3 levels: mate, familiar, and unfamiliar), playback call type (5 levels: distance, stack, kackle, hat, and tet; 6th level for females: song) and trial (2 levels: trial A and trial B), as well as all interactions. We included individual identity (12 levels), audience (i.e. the identity of the audience pair, 7 levels) and playback order (i.e. the order in which playback series were broadcast, 15–18 levels) as random factors. Model structure was based on the study design rather than model selection. Familiarity was expected to influence both outcome variables (total number of response calls and latency of the first answering call). Playback type and trial were also hypothesized to affect the outcome differentially by familiarity in the distinct playback call types. Therefore the interactions of all these variables were included. In order to obtain parameter estimates we used Maximum Likelihood (ML) because we were most interested in fixed effects[52]. We calculated credible intervals (CrI) using the function "sim" from the R package "arm"[53]. A total of 10000 values were simulated from the joint posterior distribution of the model parameters. If the CrI of different playback categories did not overlap, the results were considered significantly different from each other. In cases where CrIs overlapped, but the fitted values differed largely between playback series, a derived calculation from the aforementioned simulated values was performed. For this purpose, simulated values of the two groups of interest were compared (10000 comparisons), and we reported the number of cases in which the value of the first group was larger than the one of the second group. If this condition held true for less than 5% of the cases, the mean response of the first group was regarded as significantly smaller than that of the second.

To determine whether birds had habituated to the playback experiment, we investigated changes in calling rate throughout the experiment. First, we counted the events occurring in 500 s bins (roughly the length of a playback series) for each bird and trial. We then performed a linear mixed model with the number of calls as the dependent variable and bins (18 levels for males, 21 for females) as the explanatory variable with individual ID as a random factor. The parameters were simulated 10000 times to estimate fitted values and CrI from the outcome of the model. This allowed an assessment of patterns and enabled us to test whether there was a statistical difference in the number of calls between the first and the last bins.

We tested whether the familiarity of the playback affected the type of answer. For each trial we scored and counted the number of calls within 0.5 s of every playback call. We then calculated the proportion of each type of answer out of the total answers for each bird. Because very low counts might easily influence the proportions, we set a threshold of 5 answers to each playback series in order to be considered (dataset included as additional information, see Supplementary Table 3). For each playback series we compared the proportion of call types of the three familiarities using a non-parametric test (Kruskal-Wallis rank sum test). We ran the test only when there were at least 8 non-null values per series (i.e. the sample size was at least 8 times the number of explanatory variables)[54].

Additionally, we investigated whether we could explain individuals' answering rate by the pair's vocal relationship established before the experiment (during "baseline"). Among all call combinations used during "baseline", we selected only those in which each bird had used at least 3 calls in order to rule out inconsistent and rare combinations. For the resulting combinations we calculated a repeatability index[22] because only in case of a repeatable behaviour we can expect a consistent response during the experiment. We calculated repeatability according to the F ratio: the mean squares among groups divided by the mean squares within groups[55]. Finally, for repeatable call combinations, we quantified the correlation between the proportion of answers (i.e. of calls emitted within 0.5 s) during playback and baseline.

Data availability. Data available from the Dryad Digital Repository: D'Amelio B., P., Klumb., M., Adreani, N. M., Gahr, L. M. & ter Maat, A. Data from: Individual recognition of opposite sex vocalizations in the zebra finch. Dryad Digital Repository. doi:10.5061/dryad.4g8b7 (2017).

References

1. Bee, M. A. & Micheyl, C. The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology* **122**, 235–251 (2008).
2. Insley, S. J., Phillips, A. V. & Charrier, I. A review of social recognition in pinnipeds. *Aquatic Mammals* **29**, 181–201 (2003).
3. Deecke, V. B. Studying marine mammal cognition in the wild: a review of four decades of playback experiments. *Aquatic Mammals* **32**, 461–482 (2006).
4. Frommolt, K.-H., Goltsman, M. E. & Macdonald, D. W. Barking foxes, *Alopex lagopus*: field experiments in individual recognition in a territorial mammal. *Animal Behaviour* **65**, 509–518 (2003).
5. Lambrechts, M. M. & Dhondt, A. A. In *Current Ornithology* 115–139 (Springer US, 1995).
6. Sharp, S. P., McGowan, A., Wood, M. J. & Hatchwell, B. J. Learned kin recognition cues in a social bird. *Nature* **434**, 1127–1130 (2005).
7. Zann, R. A. *The zebra finch: a synthesis of field and laboratory studies* (Oxford Univ. Press, Oxford, 1996).
8. McCowan, L. S. C., Mariette, M. M. & Griffith, S. C. The size and composition of social groups in the wild zebra finch. *Emu* **115**, 191–198 (2015).
9. Elie, J. E. & Theunissen, F. E. Meaning in the avian auditory cortex: neural representation of communication calls. *The European journal of neuroscience* **41**, 546–67 (2015).
10. Ter Maat, A., Trost, L., Sagunsky, H., Seltmann, S. & Gahr, M. Zebra finch mates use their forebrain song system in unlearned call communication. *PloS one* **9**, e109334 (2014).
11. Gill, L. F., Goymann, W., Ter Maat, A. & Gahr, M. Patterns of call communication between group-housed zebra finches change during the breeding cycle. *eLife* **4**, e07770 (2015).
12. Elie, J. E. & Theunissen, F. E. The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition* **19**, 285–315 (2016).
13. Miller, D. B. The acoustic basis of mate recognition by female Zebra finches (*Taeniopygia guttata*). *Animal Behaviour* **27**, 376–380 (1979).
14. Vignal, C., Mathevon, N. & Mottin, S. Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behavioural processes* **77**, 191–8 (2008).
15. Vignal, C., Mathevon, N. & Mottin, S. Audience drives male songbird response to partner's voice. *Nature* **430**, 448–51 (2004).
16. Jacot, A., Reers, H. & Forstmeier, W. Individual recognition and potential recognition errors in parent–offspring communication. *Behavioral Ecology and Sociobiology* **64**, 1515–1525 (2010).
17. Reers, H., Forstmeier, W. & Jacot, A. Acoustic similarity to parental calls promotes response to unfamiliar calls in zebra finch fledglings. *Animal Behaviour* **86**, 159–167 (2013).
18. Mouterde, S. C., Theunissen, F. E., Elie, J. E., Vignal, C. & Mathevon, N. Acoustic communication and sound degradation: how do the individual signatures of male and female zebra finch calls transmit over distance? *PloS one* **9**, e102842 (2014).
19. Elie, J. E. *et al*. Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal Behaviour* **80**, 597–605 (2010).
20. Gill, L. F. *et al*. A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods in Ecology and Evolution* **7**, 1349–1358 (2016).
21. Ihle, M., Kempenaers, B. & Forstmeier, W. Fitness Benefits of Mate Choice for Compatibility in a Socially Monogamous Species. *PLoS biology* **13**, e1002248 (2015).
22. D'Amelio, P. B., Trost, L. & Maat, A. Ter. Vocal exchanges during pair formation and maintenance in the zebra finch (*Taeniopygia guttata*). *Frontiers in Zoology* **14**, 13 (2017).
23. Kondo, I. & Watanabe. Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. *Behaviour* **147**, 1051–1072 (2010).
24. Crane, J. M. S. *et al*. Chestnut-crowned babblers show affinity for calls of removed group members: A dual playback without expectancy violation. *Animal Behaviour* **104**, 51–57 (2015).
25. Robertson, B. C. Vocal mate recognition in a monogamous, flock-forming bird, the silvereye. *Zosterops lateralis. Animal Behaviour* **51**, 303–311 (1996).
26. Vicario, D. S., Naqvi, N. H. & Raksin, J. N. Sex differences in discrimination of vocal communication signals in a songbird. *Animal Behaviour* **61**, 805–817 (2001).
27. Butterfield, P. A. In Social behaviour in *birds and mammals* (ed. Crook, J. H.) 249–278 (Academic Press London, 1970).
28. Sewall, K. B., Young, A. M. & Wright, T. F. Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour* **120**, 163–172 (2016).
29. Sewall, K. B. Social complexity as a driver of communication and cognition. *Integrative and Comparative Biology* **55**, 384–395 (2015).
30. Boughman, J. W. & Moss, C. F. In *Acoustic Communication* 138–224 (Springer-Verlag, 2003).
31. Urbano, C. M., Aston, A. E. & Cooper, B. G. HVC contributes toward conspecific contact call responding in male Bengalese finches. *NeuroReport* **27**, 481–486 (2016).
32. Benichov, J. I. *et al*. The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology* **26**, 309–318 (2016).
33. Wiley, R. H. Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biological Reviews* **88**, 179–195 (2013).
34. Morris, D. The reproductive behaviour of the zebra finch (*Poephila Guttata*), with special reference to pseudofemale behaviour and displacement activities. *Behaviour* **6**, 271–322 (1954).
35. Miller, D. B. Long-term recognition of father's song by female zebra finches. *Nature* **280**, 389–391 (1979).
36. Clayton, N. S. Song discrimination learning in zebra finches. *Animal Behaviour* **36**, 1016–1024 (1988).
37. Menardy, F. *et al*. Social experience affects neuronal responses to male calls in adult female zebra finches. *The European journal of neuroscience* **35**, 1322–36 (2012).
38. Menardy, F., Giret, N. & Del Negro, C. The presence of an audience modulates responses to familiar call stimuli in the male zebra finch forebrain. *The European journal of neuroscience* **40**, 3338–50 (2014).
39. Mouterde, S. C., Elie, J. E., Theunissen, F. E. & Mathevon, N. Learning to cope with degraded sounds: female zebra finches can improve their expertise in discriminating between male voices at long distances. *The Journal of experimental biology* **217**, 3169–77 (2014).
40. Perez, E. C. *et al*. The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? *Hormones and behavior* **61**, 573–81 (2012).
41. Amador, A., Perl, Y. S., Mindlin, G. B. & Margoliash, D. Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature* **495**, 59–64 (2013).
42. Theunissen, F. E. & Elie, J. E. Neural processing of natural sounds. *Nature reviews. Neuroscience* **15**, 355–66 (2014).
43. Remage-Healey, L., Adkins-Regan, E. & Romero, L. M. Behavioral and adrenocortical responses to mate separation and reunion in the zebra finch. *Hormones and Behavior* **43**, 108–114 (2003).
44. Cockrem, J. F. Stress, corticosterone responses and avian personalities. *Journal of Ornithology* **148**, 169–178 (2007).
45. Kroodsma, D. E., Byers, B. E., Goodale, E., Johnson, S. & Liu, W.-C. Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour* **61**, 1029–1033 (2001).

46. Hartigan, J. A. & Wong, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society C* **28**, 100–108 (1979).
47. Vu, V. Q. Ggbiplot: A ggplot2 based biplot. *R package* (2011).
48. Sueur, J., Aubin, T. & Simonis, C. Equipment Review Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics* **18**, 213–226 (2008).
49. Ripley, R., Hornik, K., Gebhardt, A. & Firth, D. MASS: support functions and datasets for Venables and Ripley's MASS. *R package* (2011).
50. R Core Team. R: A Language and Environment for Statistical Computing (2015).
51. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software* **67**, 1–48 (2015).
52. Zuur, A. F., Ieno, E. N. & Elphick, C. S. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* **1**, 3–14 (2010).
53. Gelman, A. & Su, Y.-S. Data analysis using regression and multilevel/hierarchical models {arm}. *R package* (2015).
54. Forstmeier, W. & Schielzeth, H. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology* **65**, 47–55 (2011).
55. Lessells, C. M. & Boag, P. T. Unrepeatable repeatabilities: a common mistake. *The Auk* **104**, 116–121 (1987).

Acknowledgements

We are grateful to Lisa Trost, Hannes Sagunsky and Markus Abels for technical support and advice. We thank Robin Abbey-Lee and Erica Stuber for useful comments and suggestions. We thank the International Max Planck Research School (IMPRS) for Organismal Biology for training and support. This work was funded by the Max Planck Society.

Author Contributions

P.B.D. conceived the study. P.B.D., M.N.A. and M.K. designed the experiment. M.K. performed the experiment. M.K. processed the data. P.B.D. and M.K. analysed the data with input from all authors. P.B.D. and M.K. wrote the manuscript with input from all authors. A.T. and M.G. provided supervision throughout the project.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-05982-x

Competing Interests: The authors declare that they have no competing interests.

Supplementary Materials for

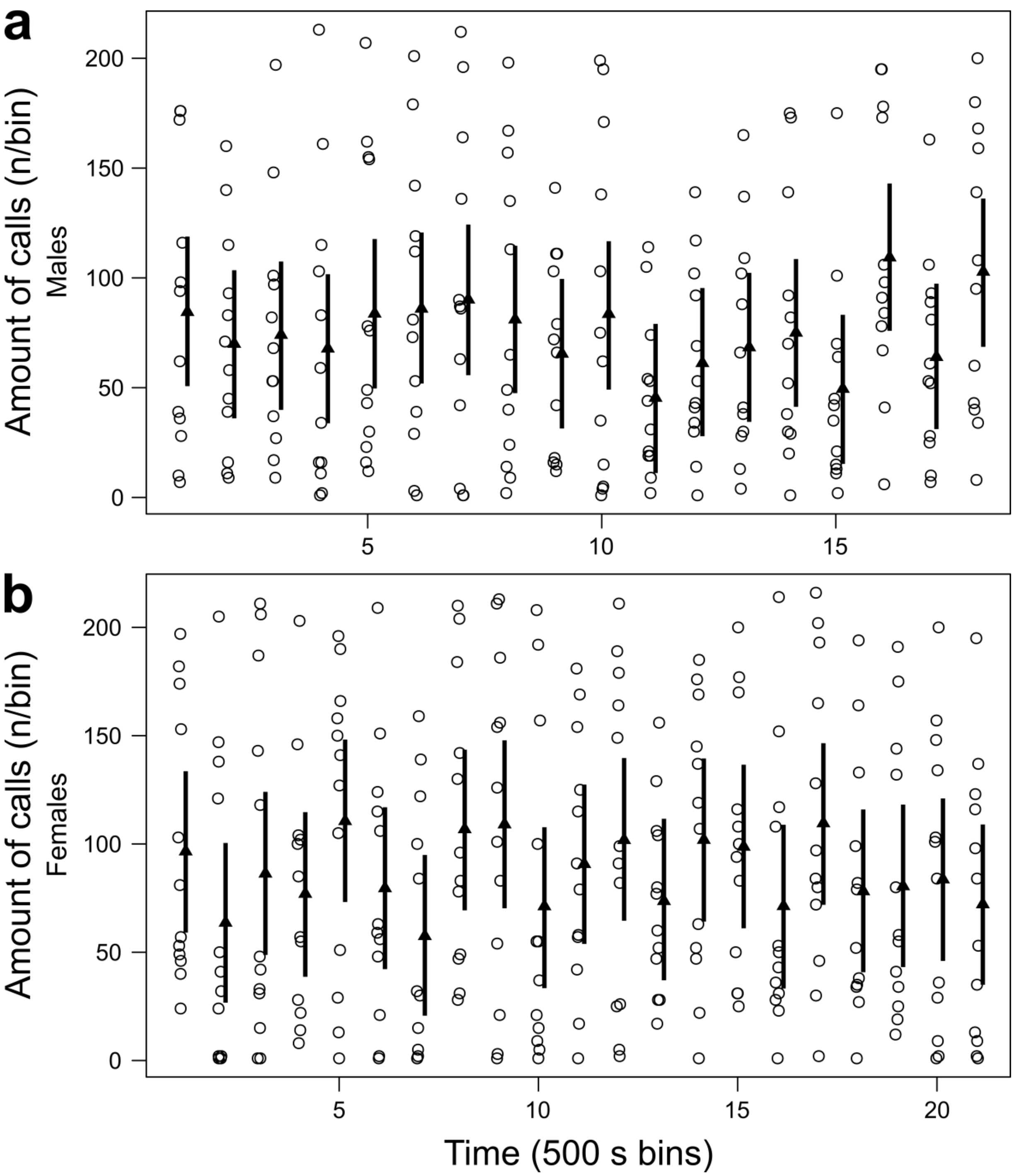
Individual recognition of opposite sex vocalizations in the zebra finch

Pietro B. D’Amelio*, Milena Klumb, Mauricio N. Adreani, Manfred L. Gahr, and Andries ter Maat

*Correspondence to: pdamelio@orn.mpg.de; pie.damelio@gmail.com

This file includes:

Supplementary Figures S1 to S9 and Table S1 to S9



Supplementary fig.1

Acclimatization of the number of call. Number of calls focal individuals emitted during the entire playback experiment, divided in 500 s bins. Shown are the raw data of both trials for all 12 individuals, divided in males (top, a) and females (bottom, b). Dots represent the sum of responses each individual emitted per 500 s bin by trials; error bars indicate estimated credible intervals and fitted values from the linear mixed model. We notice that in contrast to what we would expect in case of becoming accustomed to the playback stimuli; there is not a steady decline of the number of calls during the playback. There is no statistical difference between the first and the last bin.

Call type	**Familiarity**	**Latency**		**P. mate slower**	
		A	**B**	**A**	**B**
Tet	Mate	0.89 ± 0.45	0.63 ± 0.40		
	Familiar	0.80 ± 0.44	0.66 ± 0.42	0.9419	0.755
	Unfamiliar	0.66 ± 0.42	0.68 ± 0.42	0.743	0.8757
Stack	Mate	0.55 ± 0.38	0.52 ± 0.37		
	Familiar	0.55 ± 0.39	0.65 ± 0.42	0.4206	**0.0018**
	Unfamiliar	0.60 ± 0.41	0.64 ± 0.41	0.1343	**0.0018**
Distance	Mate	0.39 ± 0.30	0.43 ± 0.32		
	Familiar	0.44 ± 0.34	0.45 ± 0.33	0.1836	0.3585
	Unfamiliar	0.48 ± 0.38	0.53 ± 0.38	0.0549	0.0957
Kackle	Mate	0.76 ± 0.41	0.64 ± 0.41		
	Familiar	0.83 ± 0.38	0.72 ± 0.41	0.7985	0.1429
	Unfamiliar	0.76 ± 0.41	0.64 ± 0.42	0.6141	0.571
Hat	Mate	0.63 ± 0.42	0.72 ± 0.43		
	Familiar	0.71 ± 0.42	0.64 ± 0.40	0.1844	0.5638
	Unfamiliar	0.69 ± 0.48	0.62 ± 0.42	0.2669	0.889

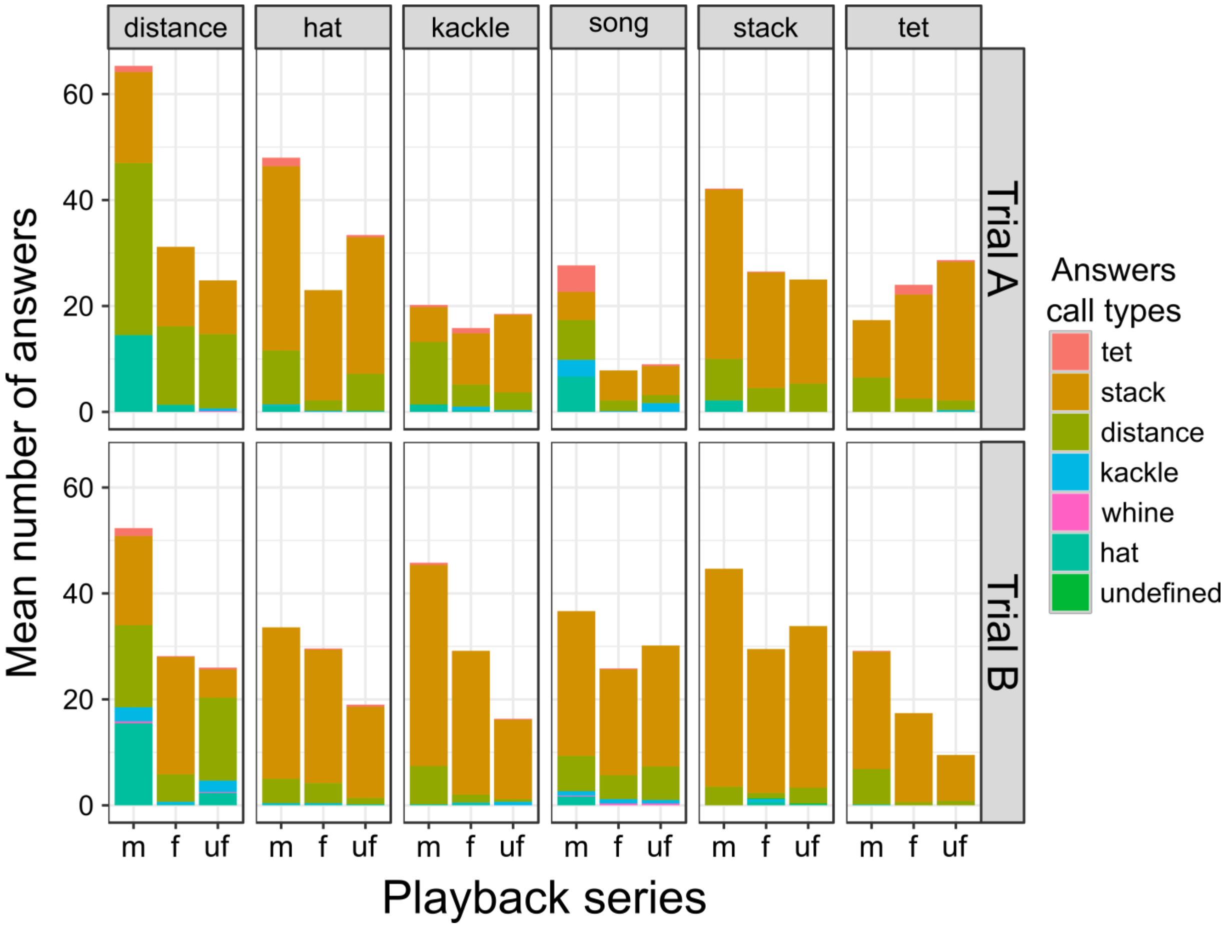
Supplementary table 1

Latency of males’ calls to females’ playback. The mean ± the standard deviation of each playback series is reported together with the probabilities (P) that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold.

Call type	**Familiarity**	**Latency**		**P mate slower**	
		A	**B**	**A**	**B**
Tet	Mate	0.76 ± 0.41	0.67 ± 0.40		
	Familiar	0.75 ± 0.41	0.77 ± 0.39	0.116	**0.0048**
	Unfamiliar	0.65 ± 0.39	0.75 ± 0.41	0.9474	**0.0196**
Stack	Mate	0.59 ± 0.40	0.65 ± 0.37		
	Familiar	0.70 ± 0.39	0.69 ± 0.39	**0.0004**	**0.0142**
	Unfamiliar	0.72 ± 0.39	0.67 ± 0.38	**0.0005**	**0.0036**
Distance	Mate	0.51 ± 0.35	0.58 ± 0.38		
	Familiar	0.64 ± 0.37	0.68 ± 0.37	**0.0004**	**0.0001**
	Unfamiliar	0.69 ± 0.35	0.64 ± 0.36	**0.016**	0.16
Kackle	Mate	0.75 ± 0.41	0.67 ± 0.40		
	Familiar	0.75 ± 0.39	0.70 ± 0.38	0.3322	**0.0256**
	Unfamiliar	0.74 ± 0.40	0.74 ± 0.40	0.5333	**0.0235**
Hat	Mate	0.62 ± 0.39	0.63 ± 0.39		
	Familiar	0.76 ± 0.40	0.65 ± 0.38	**0.0032**	0.3497
	Unfamiliar	0.71 ± 0.41	0.76 ± 0.41	**0.0187**	**0.0001**
Song	Mate	0.66 ± 0.41	0.68 ± 0.42		
	Familiar	0.79 ± 0.41	0.70 ± 0.42	**0.0001**	0.3855
	Unfamiliar	0.78 ± 0.38	0.71 ± 0.40	**0.0004**	0.2033

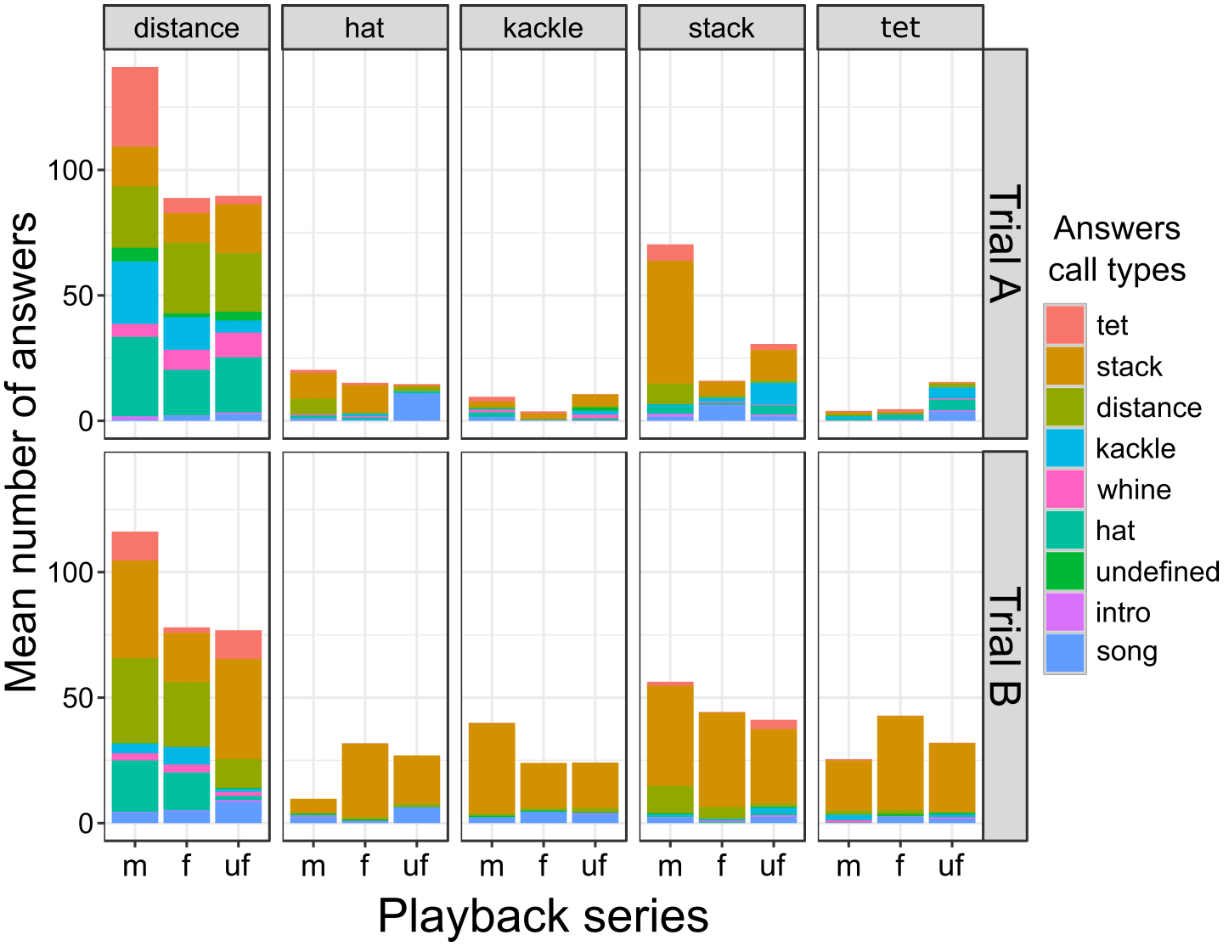
Supplementary table 2

Latency of females' calls to males' playback. The mean ± the standard deviation of each playback series is reported together with the probabilities (P) that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold.



Supplementary fig. 2

Call types that females used to answer male playback. Shown are the mean number of answers for each playback series in trials A and B. Colours represent the type of answering call (used). Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual.



Supplementary fig. 3

Call types that males used to answer female playback. Shown are the mean number of answers for each playback series in trials A and B. Colours represent the type of answering vocalization (used) ("intro" = song introductory syllable). Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual.

Trial	**Playback type**	**Answer type**	**Proportion of the specific call type used as answer (mean ± SD)**			**Test (N; K-W chi-squared; P-value)**
			Mate	Familiar	Unfamiliar	
A	Distance	Stack	0.31 ± 0.30	0.38 ± 0.30	0.42 ± 0.13	13; 0.7; 0.699
		Distance	0.54 ± 0.19	0.52 ± 0.32	0.56 ± 0.13	13; 0.03; 0.987
B	Distance	Stack	0.47 ± 0.28	0.75 ± 0.14	0.47 ± 0.38	13; 1.7; 0.418
		Distance	0.30 ± 0.22	0.16 ± 0.04	0.34 ± 0.37	13; 0.3; 0.875
A	Hat	Stack	0.67 ± 0.22	0.90 ± 0.08	0.69 ± 0.24	13; 5.4; 0.067
		Distance	0.28 ± 0.18	0.07 ± 0.09	0.25 ± 0.13	13; 5.4; 0.068
B	Hat	Stack	0.89 ± 0.09	0.89 ± 0.10	0.93 ± 0.06	9; 1.1; 0.576
A	Kackle	Stack	0.39 ± 0.37	0.52 ± 0.36	0.80 ± 0.18	15; 3.2; 0.200
		Distance	0.51 ± 0.38	0.26 ± 0.23	0.19 ± 0.18	15; 2; 0.365
B	Kackle	Stack	0.87 ± 0.10	0.94 ± 0.04	0.84 ± 0.23	15; 0.9; 0.630
		Distance	0.12 ± 0.10	0.05 ± 0.04	0.06 ± 0.07	15; 1.8; 0.397
A	Song	Distance	0.34 ± 0.32	0.37 ± 0.18	0.16 ± 0.11	8; 1.4; 0.488
B	Song	Stack	0.62 ± 0.21	0.87 ± 0.13	0.84 ± 0.16	10; 3; 0.223
A	Stack	Stack	0.86 ± 0.16	0.82 ± 0.12	0.77 ± 0.23	12; 0.7; 0.692
		Distance	0.11 ± 0.14	0.18 ± 0.12	0.23 ± 0.23	12; 1.4; 0.488
B	Stack	Stack	0.93 ± 0.08	0.92 ± 0.03	0.95 ± 0.07	15; 0.8; 0.673
		Distance	0.07 ± 0.08	0.04 ± 0.03	0.05 ± 0.07	15; 0.7; 0.692
A	Tet	Stack	0.68 ± 0.32	0.72 ± 0.41	0.82 ± 0.40	16; 2.9; 0.237
		Distance	0.32 ± 0.32	0.08 ± 0.09	0.05 ± 0.08	16; 5.5; 0.065
B	Tet	Stack	0.80 ± 0.20	0.98 ± 0.04	0.81 ± 0.32	12; 2.4; 0.307

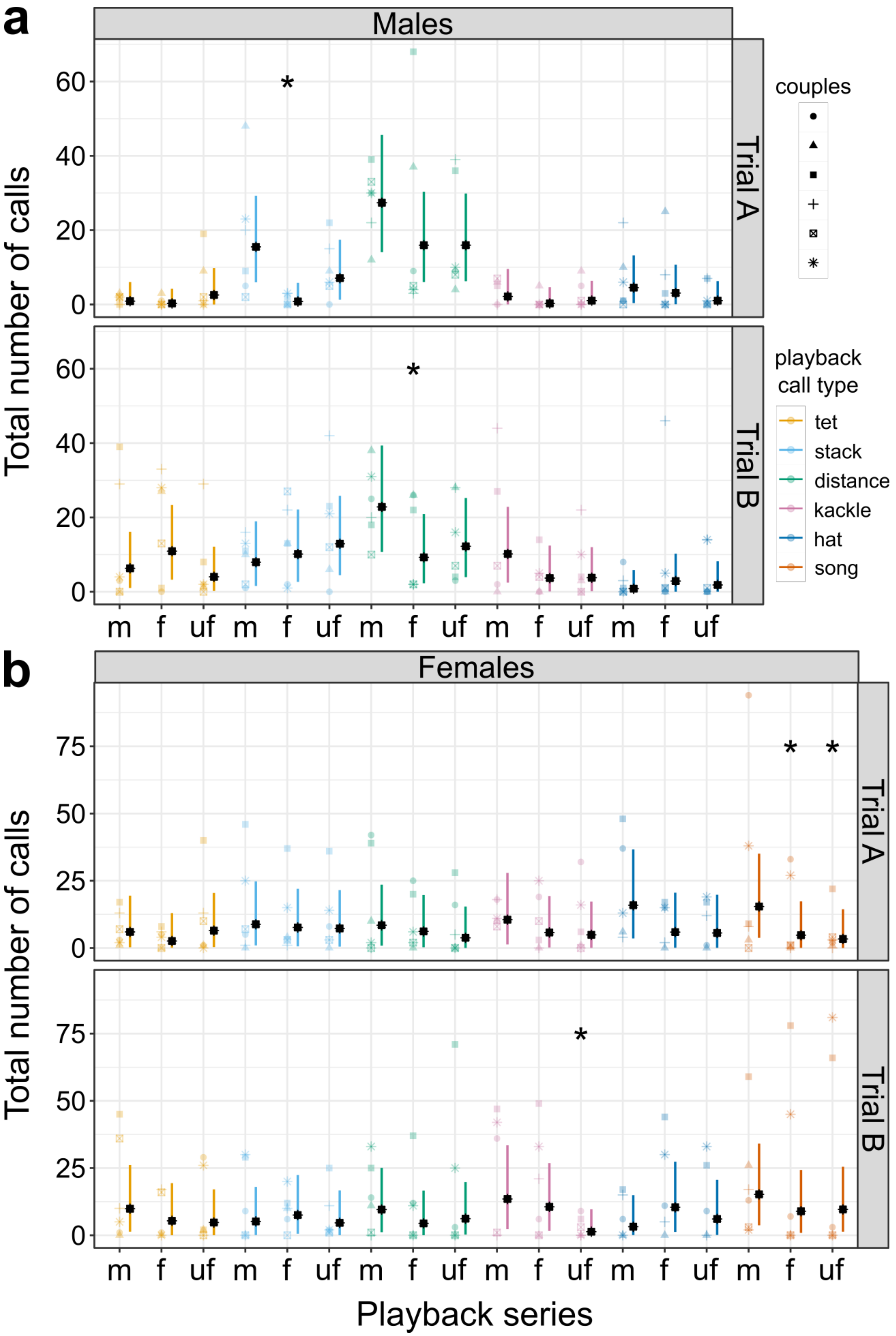
Supplementary table 3

Females' answers to males' playback of the three familiarities. Represented are the over threshold (at least 5 calls per bird per series, at least 8 non-null values per series) tested proportions of each answer type over all the answers for each series. A numerical example illustrates best our approach: to test whether birds used differential proportions of stacks out of all the answers to respond to play-backed stacks depending on the familiarity level we did the follow: If for example a bird used 30 calls to answer the stack mate calls and 15 of these answers were stacks, its proportion of answers with stack to mate's stack was 50%; if the same bird then used 20 calls to answer the familiar stacks and 15 of these answers were stacks, its proportion of answers with stacks to familiar stacks was 75%; finally, if the same bird used 20 calls to answer the unfamiliar stacks and 18 of these answers were stacks, its proportion of answers with stacks to unfamiliar stacks was 90%. We did the same for every tested bird and then compared the proportion of the three familiarities with a non-parametric test (Kruskal-Wallis rank sum test).

Trial	**Playback type**	**Answer type**	**Proportion of the specific call type used as answer (mean ± SD)**			**Test (N; K-W chi-squared; P-value)**
			Mate	Familiar	Unfamiliar	
A	Distance	Tet	0.22 ± 0.24	0.08 ± 0.10	0.03 ± 0.05	18; 5.3; 0.070
		Stack	0.11 ± 0.16	0.27 ± 0.33	0.25 ± 0.21	18; 0.9; 0.641
		Distance	0.20 ± 0.19	0.28 ± 0.35	0.27 ± 0.26	18; 0.04; 0.982
		Whine	0.03 ± 0.04	0.15 ± 0.27	0.17 ± 0.37	18; 0.04; 0.980
		Hat	0.23 ± 0.27	0.09 ± 0.17	0.13 ± 0.26	18; 2; 0.372
B	Distance	Tet	0.06 ± 0.10	0.03 ± 0.04	0.14 ± 0.34	18; 0.2; 0.904
		Stack	0.31 ± 0.26	0.49 ± 0.44	0.51 ± 0.39	18; 1.1; 0.567
		Distance	0.25 ± 0.21	0.21 ± 0.33	0.14 ± 0.13	18; 1; 0.612
		Whine	0.04 ± 0.08	0.03 ± 0.05	0.03 ± 0.03	18; 0.4; 0.829
		Hat	0.19 ± 0.37	0.15 ± 0.24	0.03 ± 0.05	18; 0.5; 0.781
A	Hat	Stack	0.55 ± 0.25	0.64 ± 0.44	0.07 ± 0.08	9; 2.2; 0.326
B	Hat	Stack	0.62 ± 0.31	0.92 ± 0.13	0.75 ± 0.20	11; 2.7; 0.255
A	Kackle	Stack	0.36 ± 0.34	0.67 ± 0.18	0.44 ± 0.30	9; 1.5; 0.472
B	Kackle	Stack	0.93 ± 0.06	0.82 ± 0.23	0.78 ± 0.26	15; 0.6; 0.754
		Distance	0.02 ± 0.02	0.03 ± 0.06	0.05 ± 0.05	15; 0.4; 0.820
A	Stack	Stack	0.64 ± 0.14	0.57 ± 0.41	0.57 ± 0.47	13; 0.1; 0.952
		Distance	0.13 ± 0.16	0.04 ± 0.07	0.04 ± 0.04	13; 1; 0.601
B	Stack	Stack	0.72 ± 0.25	0.85 ± 0.08	0.69 ± 0.31	17; 0.4; 0.816
		Distance	0.14 ± 0.23	0.10 ± 0.09	0.05 ± 0.05	17; 0.8; 0.666
B	Tet	Stack	0.83 ± 0.25	0.87 ± 0.07	0.81 ± 0.29	13; 1; 0.621

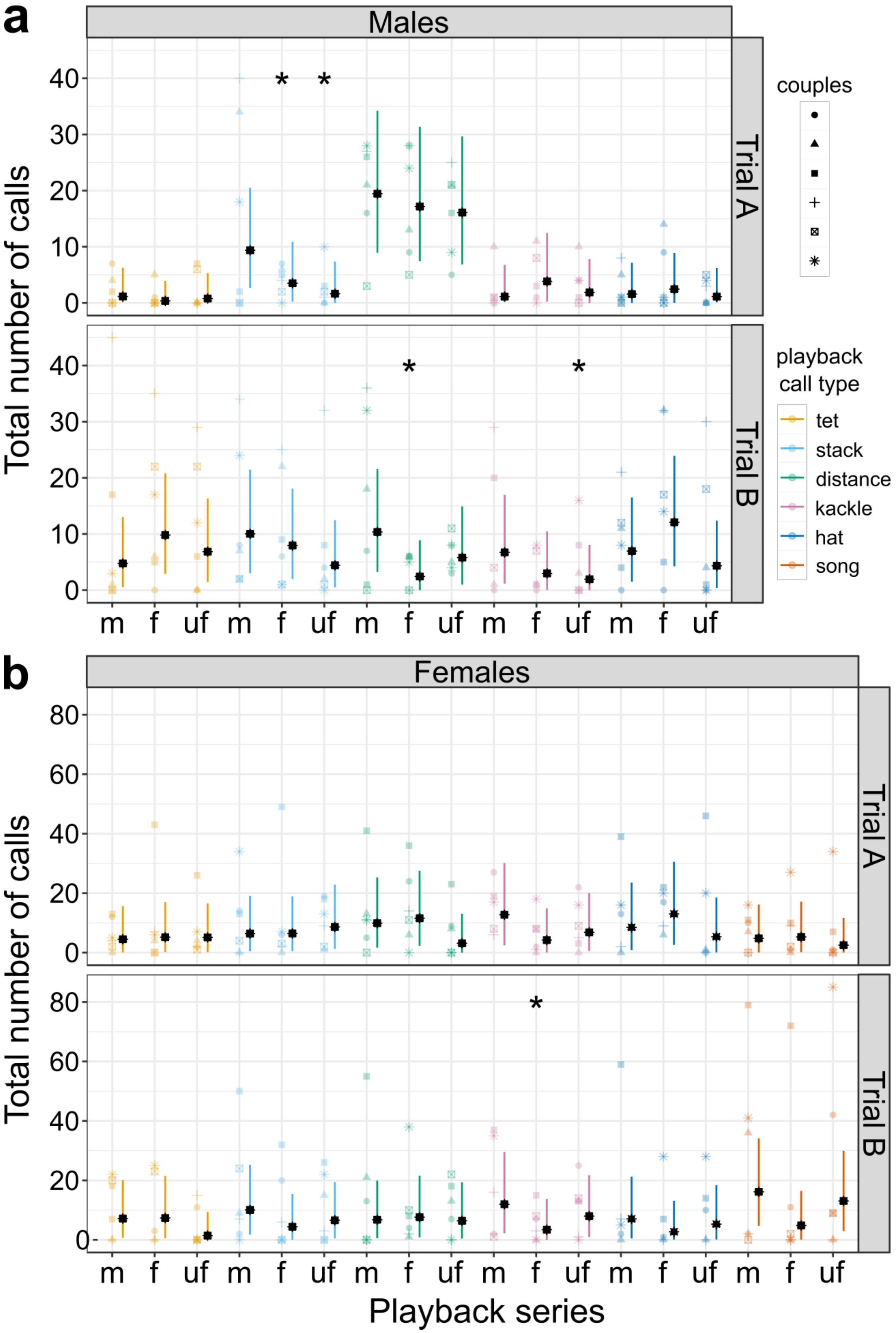
Supplementary table 4

Males’ answers to females’ playback of the three familiarities. Represented are the over threshold (at least 5 calls per bird per series, at least 8 non-null values per series) tested proportions of each answer type over all the answers for each series. For further explanation and a numerical example see caption of the table S3


a
Males
Trial A
Trial B
Total number of calls
0
20
40
60
*
*
m f uf m f uf m f uf m f uf m f uf
couples
●
▲
■
+
⊠
✳
playback
call type
tet
stack
distance
kackle
hat
song
b
Females
Trial A
Trial B
Total number of calls
0
25
50
75
*
*
*
m f uf m f uf m f uf m f uf m f uf m f uf
Playback series

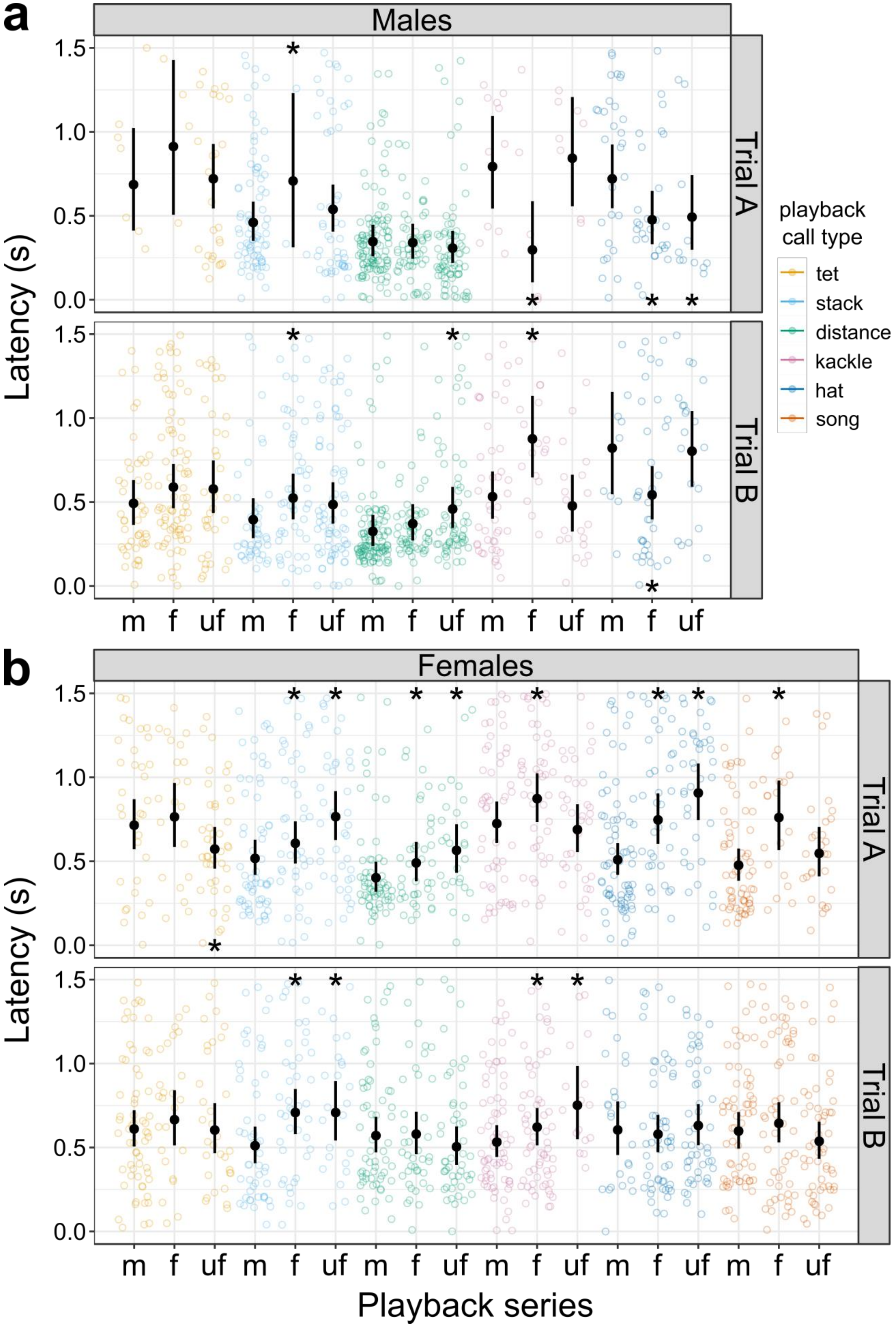
Supplementary fig. 4

Number of answering calls that focal individuals emitted during the broadcasting of the first 30 playback stimuli of the different playback series. For males (**A**) and females (**B**) and both trials raw data (symbols indicate responses of individual birds) and the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Colours represent the type of playback call broadcast; symbols indicate responses of individual birds. Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual; stars mark significant differences from the category "mate".


a
Males
Trial A
Trial B
Total number of calls
0
10
20
30
40
*
*
*
*
couples
●
▲
■
+
⊠
✳
playback call type
tet
stack
distance
kackle
hat
song
m
f
uf
m
f
uf
m
f
uf
m
f
uf
m
f
uf
b
Females
Trial A
Trial B
Total number of calls
0
20
40
60
80
*
m
f
uf
m
f
uf
m
f
uf
m
f
uf
m
f
uf
m
f
uf
Playback series

Supplementary fig. 5

Number of answering calls that focal individuals emitted during the broadcasting of the last 30 playback stimuli of the different playback series. For males (**A**) and females (**B**) and both trials raw data (symbols indicate responses of individual birds) and the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Colours represent the type of playback call broadcast; symbols indicate responses of individual birds. Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual; stars mark significant differences from the category "mate".


a
Males
Trial A
Trial B
Latency (s)
1.5
1.0
0.5
0.0
playback
call type
tet
stack
distance
kackle
hat
song
m
f
uf
b
Females
Playback series
*

Supplementary fig. 6

Latency to the first answering call, measured for the first 30 (playback) stimuli of the different playback series (analysed time interval: 0 - 1.5 s after the onset of the playback stimulus). Colours represent the type of playback call broadcast; dots indicate individual calls (raw data). Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual. For males (**A**) and females (**B**) and both trials the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Stars mark significant differences from the category "mate"; stars at the bottom indicate significance in the direction opposite to the expectation (i.e. slower answers to the mate than to non-mates).

Call type	Familiarity	Latency		P mate slower		Qualitatively consistent with full dataset?	
		A	B	A	B	A	B
	Mate	0.80 ± 0.43	0.49 ± 0.33				
Tet	Familiar	1.04 ± 0.42	0.65 ± 0.41	0.0836	0.0862	YES	YES
	Unfamiliar	0.67 ± 0.44	0.63 ± 0.44	0.4116	0.1147	YES	YES
	Mate	0.52 ± 0.34	0.41 ± 0.32				
Stack	Familiar	0.62 ± 0.55	0.61 ± 0.39	**0.0001**	**0.0247**	NO	YES
	Unfamiliar	0.56 ± 0.46	0.53 ± 0.39	0.0954	0.0799	YES	NO
	Mate	0.37 ± 0.27	0.33 ± 0.26				
Distance	Familiar	0.32 ± 0.20	0.36 ± 0.22	0.4982	0.4974	YES	YES
	Unfamiliar	0.32 ± 0.31	0.49 ± 0.36	0.8049	**0.0042**	YES	NO
	Mate	0.84 ± 0.40	0.57 ± 0.39				
Kackle	Familiar	0.53 ± 0.59	0.87 ± 0.32	**1**	**0**	NO	NO
	Unfamiliar	0.92 ± 0.30	0.51 ± 0.37	0.3716	0.7859	YES	YES
	Mate	0.83 ± 0.45	0.80 ± 0.39				
Hat	Familiar	0.55 ± 0.36	0.61 ± 0.46	**0.9968**	**0.9757**	NO	NO
	Unfamiliar	0.47 ± 0.43	0.84 ± 0.41	**0.995**	0.5419	NO	YES

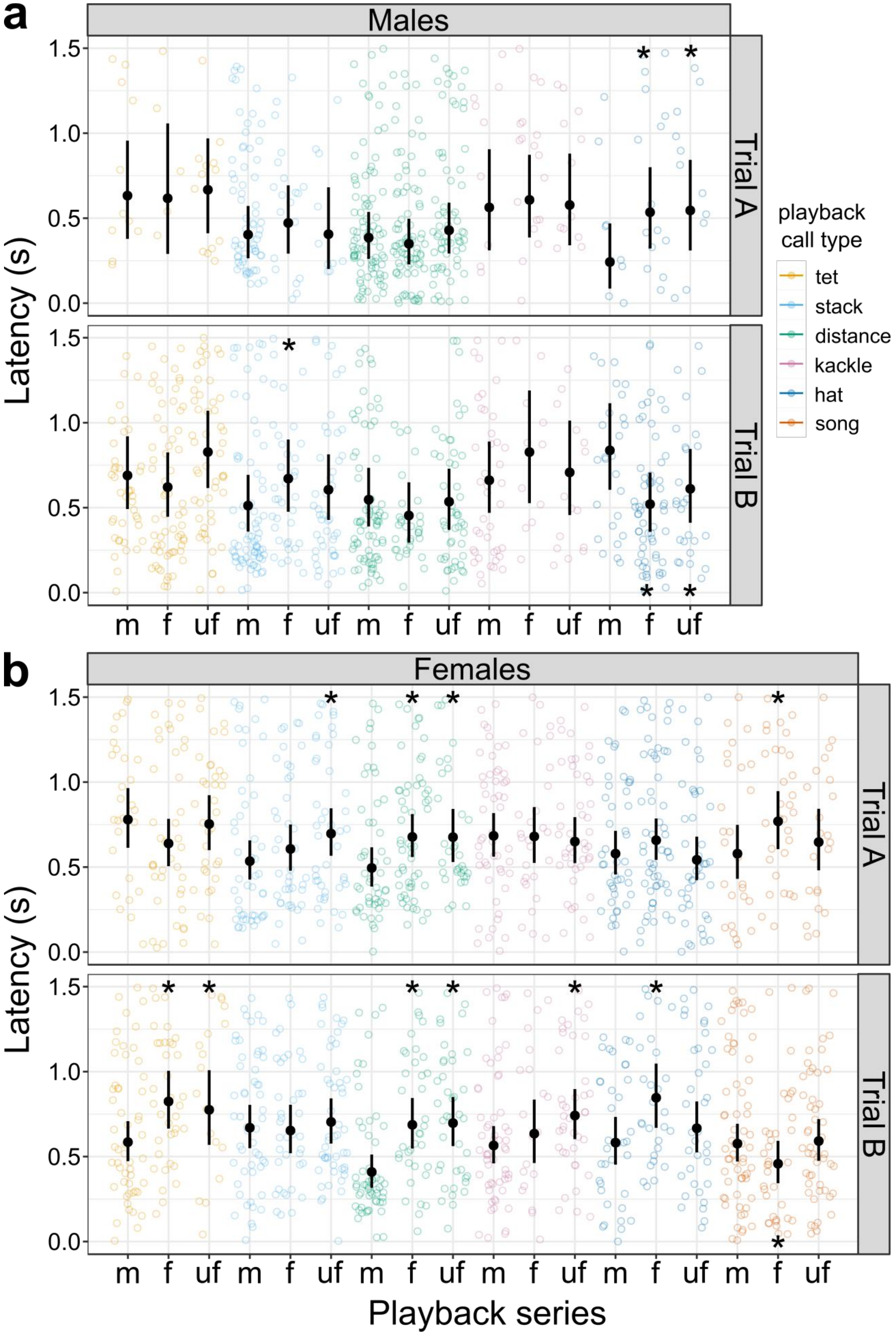
Supplementary table 5

Latency of males' calls to females' playback considering only the first 30 calls of each series. The mean ± the standard deviation of each playback series is reported together with the probability (P) value that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold. The two rightmost columns indicated whether the result is different from the one obtained considering the entire playback series.

Call type	Familiarity	Latency		P mate slower		Qualitatively consistent with full dataset?	
		A	B	A	B	A	B
Tet	Mate	0.81 ± 0.43	0.66 ± 0.40				
	Familiar	0.83 ± 0.40	0.74 ± 0.38	0.2611	0.1715	YES	NO
	Unfamiliar	0.62 ± 0.33	0.68 ± 0.41	**0.9758**	0.5432	NO	NO
Stack	Mate	0.57 ± 0.37	0.59 ± 0.41				
	Familiar	0.70 ± 0.42	0.77 ± 0.40	**0.0462**	**0.0002**	YES	YES
	Unfamiliar	0.76 ± 0.39	0.75 ± 0.38	**0**	**0.0001**	YES	YES
Distance	Mate	0.44 ± 0.31	0.62 ± 0.38				
	Familiar	0.52 ± 0.30	0.58 ± 0.37	**0.0271**	0.8984	YES	NO
	Unfamiliar	0.60 ± 0.33	0.55 ± 0.36	**0.0001**	0.431	YES	YES
Kackle	Mate	0.77 ± 0.42	0.58 ± 0.33				
	Familiar	0.91 ± 0.39	0.61 ± 0.36	**0.0107**	**0.0394**	NO	YES
	Unfamiliar	0.72 ± 0.37	0.79 ± 0.33	0.7197	**0**	YES	YES
Hat	Mate	0.56 ± 0.38	0.67 ± 0.40				
	Familiar	0.82 ± 0.42	0.62 ± 0.38	**0**	0.6271	YES	YES
	Unfamiliar	0.95 ± 0.39	0.65 ± 0.35	**0**	0.3757	YES	NO
Song	Mate	0.52 ± 0.33	0.65 ± 0.35				
	Familiar	0.74 ± 0.34	0.71 ± 0.44	**0**	0.2027	YES	YES
	Unfamiliar	0.59 ± 0.34	0.55 ± 0.36	0.0775	0.8752	NO	YES

Supplementary table 6

Latency of females' calls to males' playback considering only the first 30 calls of each series. The mean ± the standard deviation of each playback series is reported together with the probability (P) value that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold. The two rightmost columns indicated whether the result is different from the one obtained considering the entire playback series.


a
Males
Latency (s)
1.5
1.0
0.5
0.0
Trial A
Trial B
playback call type
tet
stack
distance
kackle
hat
song
m f uf m f uf m f uf m f uf m f uf
b
Females
Latency (s)
1.5
1.0
0.5
0.0
Trial A
Trial B
m f uf m f uf m f uf m f uf m f uf m f uf
Playback series

Supplementary fig. 7

Latency to the first answering call, measured for the last 30 (playback) stimuli of the different playback series (analysed time interval: 0 - 1.5 s after the onset of the playback stimulus). Colours represent the type of playback call broadcast; dots indicate individual calls (raw data). Familiarity categories: **m**= mate of the focal bird; **f**= familiar individual; **uf**= unfamiliar individual. For males (**A**) and females (**B**) and both trials the computed 95% credible intervals (error bars) as well as the fitted value (black symbols) are shown. Stars mark significant differences from the category "mate", stars at the bottom indicate significance in the direction opposite to the expectation (i.e. slower answers to the mate than to non-mates).

Call type	Familiarity	Latency		P mate slower		Qualitatively consistent with full dataset?	
		A	B	A	B	A	B
Tet	Mate	0.77 ± 0.47	0.61 ± 0.32				
	Familiar	0.87 ± 0.45	0.60 ± 0.42	0.5469	0.7434	YES	YES
	Unfamiliar	0.66 ± 0.34	0.80 ± 0.42	0.4131	0.1148	YES	YES
Stack	Mate	0.56± 0.37	0.50 ± 0.42				
	Familiar	0.58 ± 0.36	0.63 ± 0.37	0.2021	**0.0414**	YES	YES
	Unfamiliar	0.50 ± 0.37	0.59 ± 0.42	0.4894	0.1476	YES	NO
Distance	Mate	0.47 ± 0.36	0.52 ± 0.38				
	Familiar	0.44 ± 0.34	0.42 ± 0.24	0.5022	0.506	YES	YES
	Unfamiliar	0.56 ± 0.41	0.54 ± 0.38	0.2751	0.5683	YES	YES
Kackle	Mate	0.85 ± 0.49	0.63 ± 0.44				
	Familiar	0.80 ± 0.40	0.81 ± 0.42	0.3835	0.0706	YES	YES
	Unfamiliar	0.75 ± 0.38	0.62 ± 0.40	0.4619	0.3243	YES	YES
Hat	Mate	0.43 ± 0.41	0.77 ± 0.39				
	Familiar	0.75 ± 0.51	0.52 ± 0.37	**0.0076**	**0.9977**	NO	NO
	Unfamiliar	0.75 ± 0.45	0.60 ± 0.40	**0.0063**	**0.9721**	NO	NO

Supplementary table 7

Latency of males’ calls to females’ playback considering only the last 30 calls of each series. The mean ± the standard deviation of each playback series is reported together with the probability (P) value that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold. The two rightmost columns indicated whether the result is different from the one obtained considering the entire playback series.

Call type	Familiarity	Latency		P mate slower		Qualitatively consistent with full dataset?	
		A	B	A	B	A	B
Tet	Mate	0.85 ± 0.42	0.67 ± 0.40				
	Familiar	0.70 ± 0.46	0.83 ± 0.38	0.9476	**0.0001**	YES	YES
	Unfamiliar	0.80 ± 0.38	0.68 ± 0.37	0.6155	**0.0021**	YES	YES
Stack	Mate	0.60 ± 0.45	0.72 ± 0.39				
	Familiar	0.64 ± 0.40	0.69 ± 0.36	0.1143	0.6098	NO	NO
	Unfamiliar	0.74 ± 0.42	0.73 ± 0.37	**0.0054**	0.2941	YES	NO
Distance	Mate	0.58 ± 0.39	0.45 ± 0.34				
	Familiar	0.71 ± 0.37	0.73 ± 0.38	**0.0021**	**0**	YES	YES
	Unfamiliar	0.69 ± 0.36	0.74 ± 0.38	**0.0023**	**0**	YES	NO
Kackle	Mate	0.73 ± 0.40	0.62 ± 0.39				
	Familiar	0.72 ± 0.41	0.73 ± 0.45	0.5208	0.1153	YES	NO
	Unfamiliar	0.73 ± 0.35	0.79 ± 0.37	0.6942	**0.0021**	YES	YES
Hat	Mate	0.65 ± 0.37	0.61 ± 0.38				
	Familiar	0.72 ± 0.42	0.83 ± 0.36	0.1217	**0.0004**	NO	NO
	Unfamiliar	0.57 ± 0.41	0.75 ± 0.43	0.7051	0.1229	NO	NO
Song	Mate	0.66 ± 0.46	0.66 ± 0.43				
	Familiar	0.84 ± 0.39	0.49 ± 0.38	**0.0138**	**0.9869**	YES	NO
	Unfamiliar	0.68 ± 0.41	0.63 ± 0.37	0.2061	0.3962	NO	YES

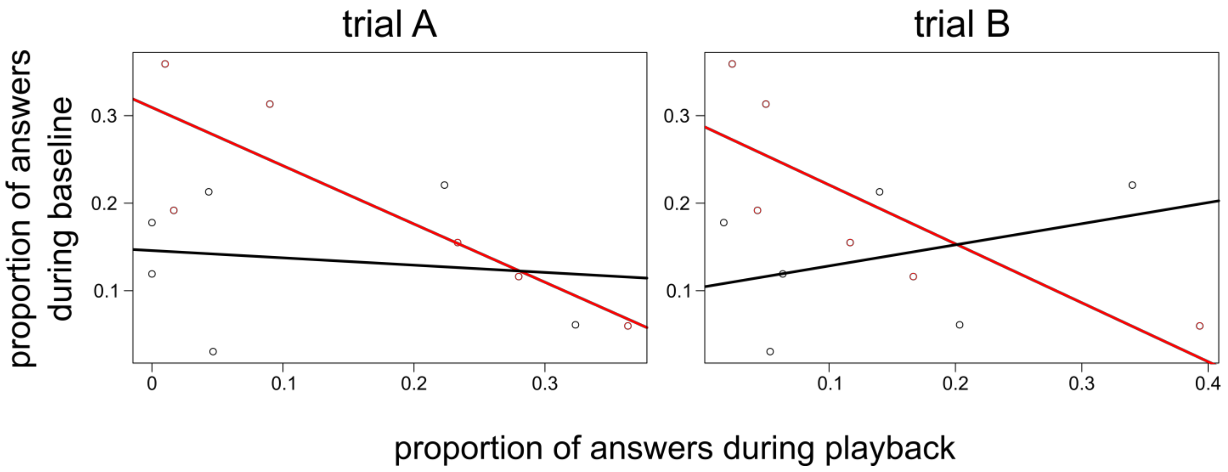
Supplementary table 8

Latency of females' calls to males' playback considering only the last 30 calls of each series. The mean ± the standard deviation of each playback series is reported together with the probability (P) value that calls emitted during mate series have a longer latency than calls emitted during the other familiarity categories. The statistical significant differences are reported in bold. The two rightmost columns indicated whether the result is different from the one obtained considering the entire playback series.

call_comb.	ans_m	awd_m	ans_f	awd_f	day	ans_m_p	awd_m_p	ans_f_p	awd_f_p	cp
stack_stack	148	125	124	152	d1	38.54	32.55	21.12	25.89	1
stack_stack	175	174	184	177	d3	33.27	33.08	21.47	20.65	1
stack_stack	208	243	243	203	d1	15.77	18.42	19.35	16.16	2
stack_stack	296	267	262	293	d3	15.23	13.74	16.21	18.13	2
stack_stack	1196	1346	1173	1045	d1	32.88	36.99	20.94	18.65	5
stack_stack	1041	1302	1274	1009	d3	29.77	37.23	23.20	18.38	5
stack_stack	70	84	86	75	d1	7.88	9.46	7.21	6.29	8
stack_stack	193	224	227	188	d3	4.06	4.71	16.62	13.76	8
stack_stack	79	87	91	81	d1	15.52	17.09	2.03	1.80	9
stack_stack	169	196	198	180	d3	22.84	26.49	4.06	3.69	9
stack_stack	64	127	120	60	d1	10.261	20.35	5.60	2.80	11
stack_stack	333	462	450	323	d3	12.97	17.99	6.60	4.75	11

Supplementary table 9

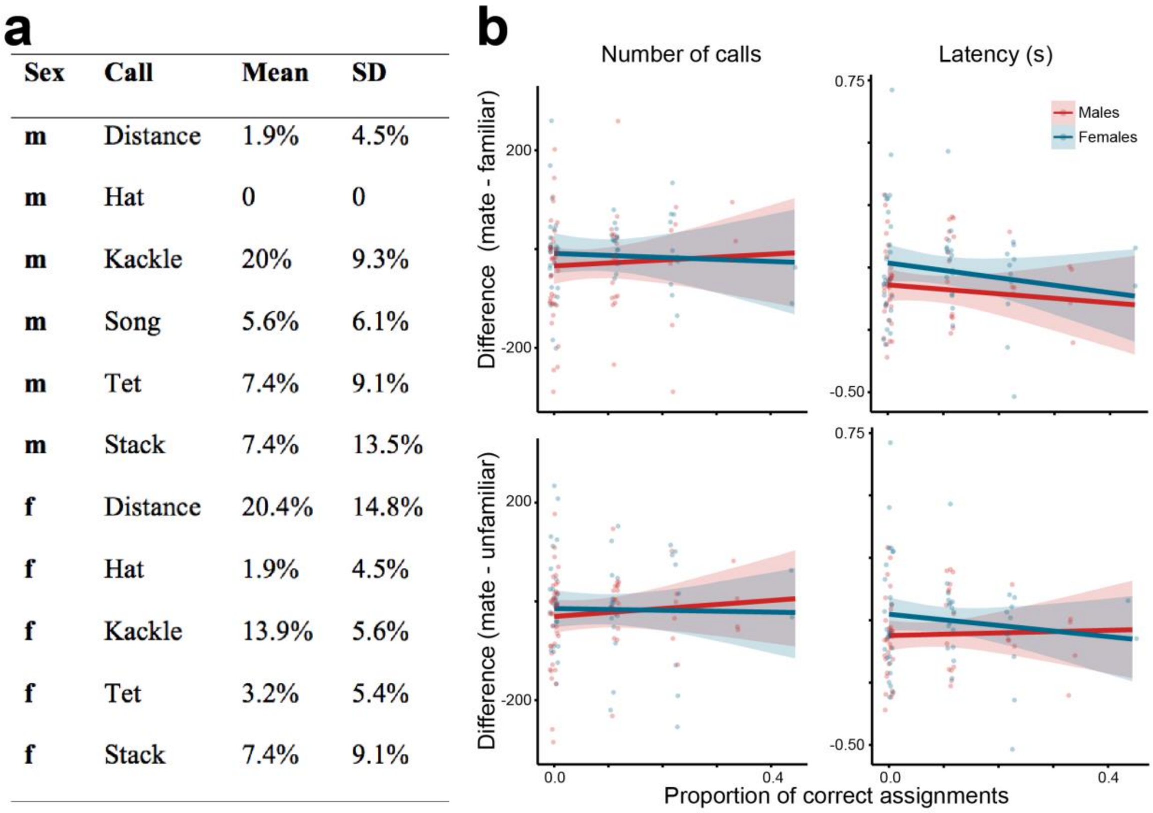
Stack-stack call interactions in all six focal pairs during the baseline (days 1 and 3). Absolute and relative numbers of answered stack calls and stack calls used as answers by males and females (ans_m, ans_f: number of stack calls that were used by male or female, respectively, to answer a partner's stack call; awd_m, awd_f: number of male or female stack calls, respectively, that were answered by a stack call of the partner; ans_m_p, ans_f_p: percentaged number of stack calls that were used by male or female, respectively, to answer a partner's stack call; awd_m_p, awd_f_p: percentage number of male or female stack calls, respectively, that were answered by a stack call of the partner; cp: couple ID).



Tria A, MALES, Baseline~playback experiment			
	Estimate	**Standard Error**	**P-value**
Intercept	3.09	0.44	
playback experiment	-6.66	2.07	**0.0324**
Tria B, MALES, Baseline~playback experiment			
Intercept	2.88	0.45	
playback experiment	-6.73	2.46	0.0523
Tria A, FEMALES, Baseline~playback experiment			
Intercept	1.45	0.47	
playback experiment	-0.83	2.93	0.7905
Tria B, FEMALES, Baseline~playback experiment			
Intercept	1.04	0.54	
playback experiment	2.41	3.09	0.4784

Supplementary fig. 8

Relationship between the proportion of stack-stack answers given during baseline (mean of both days) and those during trials A and B of the playback experiment. Dots represent individual birds, lines the regression line of the linear model between the 2 variables, whose results are reported in the table together with the correspondent P-value. The males are depicted in red and the females in black.



a

Sex	Call	Mean	SD
m	Distance	1.9%	4.5%
m	Hat	0	0
m	Kackle	20%	9.3%
m	Song	5.6%	6.1%
m	Tet	7.4%	9.1%
m	Stack	7.4%	13.5%
f	Distance	20.4%	14.8%
f	Hat	1.9%	4.5%
f	Kackle	13.9%	5.6%
f	Tet	3.2%	5.4%
f	Stack	7.4%	9.1%

Supplementary fig. 9

Relation between the variability in the call type (e.g., stack call) to the variation in conspecific response. (a) Values of incorrect assignment of the linear discriminant analysis. For each call type the percentage of incorrectly assigned calls are reported (mean value and the standard deviation, n=6 individuals per sex). Higher values correspond to more similar calls (i.e. call types more difficult to assign to the correct individual). (b) Results of linear mixed model (shaded areas) with row data (dots) correlating the within call type variability as a predictor of the conspecific response (number of calls left panel and latency in the right panels). Top panels contain the differences between mate and familiar and the bottom panels between mate and unfamiliar. Male calls are depicted in red and females in blues.

Chapter 3

Vocal exchanges during pair formation and maintenance in the zebra finch (*Taeniopygia guttata*)

Abstract

Background: Pair compatibility affects the success of a pair; however, its causes and mechanisms are not fully understood. Vocal exchange may be very important for pair formation, coordinating pair activities, maintaining the pair bond and mate guarding. To investigate the role of vocal exchange in pair formation and pair maintenance, we explored whether new and established pairs of zebra finches differed in their calling relationships. We used individualised backpack microphones to examine the entire daily vocal emission of pairs, with parallel video recording of behaviour. Results: We found that in non-breeding, isolated pairs, a specific type of call, the "stack call", was the most common. Furthermore, all pairs used the stack call for precisely timed antiphonal exchange. We confirmed a difference between new and established pairs in social behaviour, with the former spending less time in physical contact. Notably, we found that this was mirrored by a difference in calling behaviour: members of new pairs converged over time on a more symmetric calling relationship. Additionally, we observed different response rates to partners among individuals, but a repeatable relationship of answering within pairs, which may reflect different degrees of motivation to answer the partner.

Conclusions: Our findings show that there is plasticity in calling behaviour and that it changes during pair formation, resulting in a coordinated stack call exchange with a similar number of answers between partners once the pair is established. It is possible that some of the calling relationship measurements that we present reflect pair compatibility.

Published as: D'Amelio, P. B., Trost, L., & ter Maat, A. (2017). Vocal exchanges during pair formation and maintenance in the zebra finch (Taeniopygia guttata). *Frontiers in Zoology*, *14*(1), 13. https://doi.org/10.1186/s12983-017-0197-x

1

D'Amelio *et al. Frontiers in Zoology* (2017) 14:13
DOI 10.1186/s12983-017-0197-x

Frontiers in Zoology

RESEARCH Open Access

Vocal exchanges during pair formation and maintenance in the zebra finch (*Taeniopygia guttata*)

Pietro Bruno D'Amelio*, Lisa Trost and Andries ter Maat

Abstract

Background: Pair compatibility affects the success of a pair; however, its causes and mechanisms are not fully understood. Vocal exchange may be very important for pair formation, coordinating pair activities, maintaining the pair bond and mate guarding. To investigate the role of vocal exchange in pair formation and pair maintenance, we explored whether new and established pairs of zebra finches differed in their calling relationships. We used individualised backpack microphones to examine the entire daily vocal emission of pairs, with parallel video recording of behaviour.

Results: We found that in non-breeding, isolated pairs, a specific type of call, the "stack call", was the most common. Furthermore, all pairs used the stack call for precisely timed antiphonal exchange. We confirmed a difference between new and established pairs in social behaviour, with the former spending less time in physical contact. Notably, we found that this was mirrored by a difference in calling behaviour: members of new pairs converged over time on a more symmetric calling relationship. Additionally, we observed different response rates to partners among individuals, but a repeatable relationship of answering within pairs, which may reflect different degrees of motivation to answer the partner.

Conclusions: Our findings show that there is plasticity in calling behaviour and that it changes during pair formation, resulting in a coordinated stack call exchange with a similar number of answers between partners once the pair is established. It is possible that some of the calling relationship measurements that we present reflect pair compatibility.

Keywords: Monogamous songbirds, Vocal communication, Antiphonal calling, Individualized recording, Pair compatibility

Background

Individual quality does not necessarily predict the breeding success of a pair [1–3]. Instead, pair compatibility has been proposed to influence success because of synergistic effects between pair members [4, 5]. Furthermore, in several bird species, breeding success is positively related to pair-bond duration [6–9]. Although demonstrated mainly in long-lived non-Passeriformes, the benefits of "mate familiarity" and the "costs of mate change" may partially explain the effect of pair-bond duration [10]. Pair coordination is another factor that has been shown to have fitness benefits in various songbirds [11–15], and it is possible that vocal behaviour may be important for pair coordination. However, only a few studies have explicitly examined the influence of vocal exchanges between pair members on pair coordination [16]. Despite its possible interaction with pair compatibility, vocal coordination has been mainly examined in the context of duetting. Duets represent an extreme case of vocal synchronisation, where partners adjust timing and type of vocalisations to match their mate [17]. Duets have multiple, often independent, functions in different contexts [18]: they can be directed at outsiders [19–21], and can also be important in intra-pair communication for functions such as coordination of activities and pair-bond maintenance [22, 23]. Coordinated vocal exchanges between pair members may also play a critical role during pair formation [24]. Nevertheless, the role of intra-pair vocalisations during pair

---

* Correspondence: pie.damelio@gmail.com; pdamelio@orn.mpg.de
Department Gahr - Behavioural Neurobiology, Max Planck Institute for Ornithology, Eberhard-Gwinner-Straße, 82319 Seewiesen, Germany

formation has rarely been fully described or experimentally tested. Vocal exchange is a key factor in forming pair bonds, coordinating pair activity and maintaining pairs; it can thus provide us with an indication of pair compatibility.

Zebra finches are group-living songbirds that form lifelong, monogamous pairs in the wild [25, 26] and in captivity [27]. It has been hypothesised that they use two forms of communication, one with their partner and another with the rest of the group [28]. Zebra finches utter several thousand vocalisations each day and, with the exception of the song learned by males, the sexes have similar unlearned call types [15, 28, 29]. In zebra finches both sexes are involved in partner choice [30, 31]. Hence, behavioural coordination, potentially aided by vocal exchange using multiple call types, may be relevant for the choice of a mate and pair maintenance. The importance of song for pair formation in zebra finches has been extensively documented [32–34], and song after pair-bond establishment may be involved in stimulating the partner (i.e. females produced larger eggs with more orange yolks when paired to males with an high song output, [35]). However, song seems not to be critical for pair maintenance [36]. On the other hand, calling behaviour (e.g. the timing of calls and their interactions) and its importance in pair formation and maintenance has rarely been quantified. It is well documented that zebra finches initiate, and respond to, calls, taking turns in a vocal exchange [15, 37–39], a behaviour which is sometimes even termed duetting [40]. However, apart during environmental modification [41, 42], the importance and consistency of this alternating, antiphonal communication has not yet been assessed and high-resolution recording during pair formation is lacking.

Coordinated vocal exchanges within pairs could be achieved by assortative mating (i.e. choosing a partner because of a similar rate, or amount, of calling) or behavioural convergence (i.e. changing the calls' temporal patterns to answer the partner). However, evidence supporting both models is lacking. Therefore, a comprehensive description of calling behaviour, both during pair formation and after a pair bond has been established, may enhance our knowledge of the mechanisms of pair formation and maintenance of bonds independently. In the zebra finch, pair formation often takes less than a week and can start within minutes [43]. To measure calling behaviour during pair formation, we chose a time period that was sufficiently long to induce a relationship, but not long enough to be confounded with nest building, reproduction, or parenting [43–45]. To precisely quantify vocal exchange, minimally-invasive long-term recordings are necessary. Individual-based recordings enable unprecedented accuracy in quantifying calling-behaviour with minimal impact on the birds [37, 46]. Here we describe the vocal processes of pair formation and maintenance considering all vocalisations of both sexes, identifying the different call types and measuring their timing.

In this study we use week-long video and audio recordings to compare established pairs with new ones. We study new dyads from the very first encounters and hereafter we refer to this group as new pairs. We examine differences in social behaviour, to determine if differences in calling patterns (e.g. the presence or the pattern of antiphonal calling) are related to pair experience. If post-pairing behavioural convergence occurs, we expect the new and established pairs to be more similar in both social and vocal behaviour at the end of the recording period. Additionally, we hypothesise that motivation to call in response to the partner, measured as the proportion of answers out of the total number of calls, may differ from pair to pair depending on pair compatibility. The motivation to answer more frequently may be in turn correlated with time spent in physical contact, linking behavioural and vocal aspects of pair commitment. With backpack microphones, we recorded individual zebra finches and their partners without interfering with their daily activities, collecting nearly half a million vocalisation events. We mainly focussed our analysis on the stack call, one of several call types in the zebra finch repertoire [15, 28], as it was the most common call produced. Stack call was initially thought to just signal movement [28], but more and more evidence suggest that it is important in an affiliative context [29] and specifically during intra-pair communication [15, 37, 39, 47]. We identified antiphonal calling using stack calls in all pairs. New and established pairs differed in the symmetry of their calling relationship in term of number of stack calls used to answer their partner; this was paralleled by differences in social behaviour. We propose that antiphonal calling with this specific call type developed during pair formation may represent a private communication channel (i.e. the meaning of the interaction is only clear to the partners), which may enhance pair maintenance and pair synchronisation.

Methods

Study animals and recording scheme

A total of 24 mature adult zebra finches (over 120 days post-hatch) were housed in pairs and were video and audio recorded in sound-proof chambers for one week. Inside the sound-proof chambers pairs were kept in cages of 60 × 30 × 40 cm with grit, food (egg food and mixed seeds) and water *ab libitum*. The light cycle was 13 h light and 11 h dark with the day period spanning from 8:00-21:00 hours. Birds were audio-recorded for 12 hours (8:00-12:00, 12:00-16:00 and 17:00-21:00). Maintenance (cage cleaning, replacing food and water, etc.) was done between 16:00 and 17:00 so as not to interfere with the recordings. The birds were placed in the sound-proof chamber the morning of the first day of recording. Every second day we analysed 8 hrs of recording (8:00-16:00). Established pairs (N = 4), which had all bred successfully at least once prior to the experiment, were chosen randomly from breeding facilities at the Max Planck for Ornithology during a non-breeding period to capture normal daily vocal exchange. The members of new pairs (N = 8), unrelated and randomly chosen from our facilities, had never seen or heard the partners before the start of the experiment. Just prior to the experiment, the latter group was kept for at least 7 days in same-sex groups, acoustically and visually isolated from members of the opposite sex. Half of these birds had previous breeding experience and half were naïve. After the experiment we allowed the new pairs to breed in order to verify that they could raise offspring and were thus comparable to the established pairs; all 8 newly formed pairs bred successfully. While breeding, the new pairs were equipped with dummy backpacks of equal size and weight to the recording backpack, to ensure that the equipment had not impaired copulation or any other part of the breeding phase during the experiment.

Backpack microphone and recording selectivity

Transmitters backpacks, and their application and employment, are described in detail elsewhere ([15, 36, 43] the specific version used in our study is the one described for males in [15]). Individuals were fitted with backpacks custom-made for each bird (Fig. 1). Briefly, transmitters were equipped with microphones (Sparrow System, Fisher III, USA), attached to a backpack and mounted on the back of each animal. The full backpacks weighed approximately 1.3 g, which is equivalent to 8.3% of the weight of an average zebra finch (15.7 g) in our colony. The harness was made of a ring of ~19 cm silicone tube (1.7 mm outer diameter, Detakta, Germany); a second 7 mm long silicone tube (1.1 mm, Detakta, Germany) was stretched and pulled over the ring, thus separating it in two loops. The audio transmitter, protected with shrinkable tubing leaving a hole for the microphone, was fixed on the narrow part

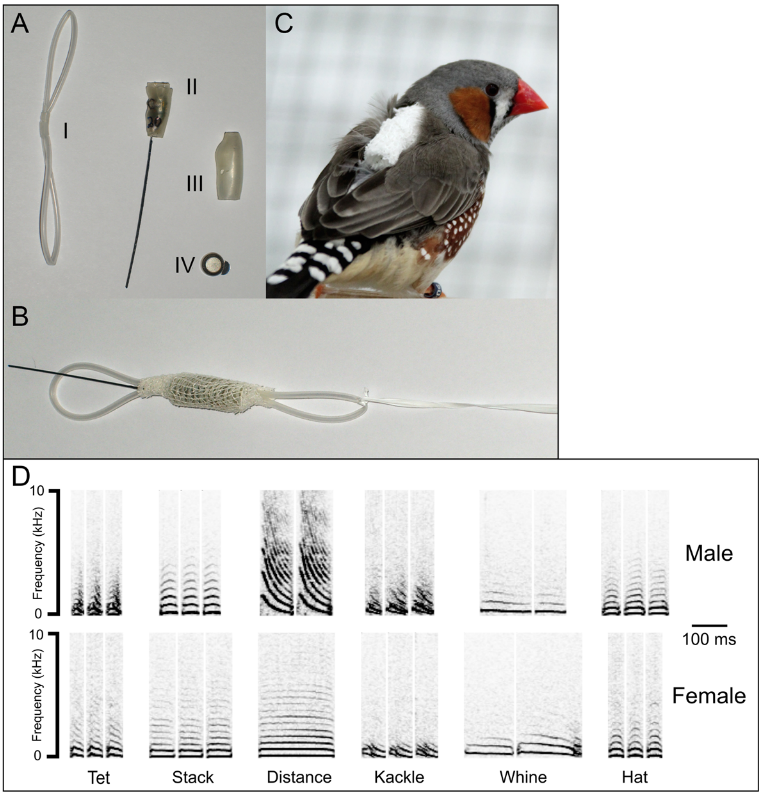


Fig. 1 Tools: telemetry transmitter and repertoire. Top picture: **a**) Different components of the telemetry transmitter. I) Elastic cord, the upper loop encircles the head and the bottom loop goes around the tail. II) Microphone transmitter body and antenna. III) Transmitter case. IV) 1.45 V battery. **b**) The assembled backpack. **c**) Zebra finch equipped with a transmitter. Bottom panel: **d**) Sonograms of the scored calls of the male and the female from a representative experimental pair (male only: song and introductory syllables are not shown): "Tet"; "Stack"; "Distance", "Kackle", "Whine", "Hat". Despite many generations of captivity, the repertoire is very similar to the one described by Zann (1996) in wild birds. We added one soft call type, called "Hat". The meaning and function of this call are yet to be determined, but it may be a modified version of the "Thuks" call used to indicate danger in wild populations [28]

between the loops with an adhesive elastic bandage (BSN medical Elastomull®haft). Finally, the transmitter was wrapped in gauze to protect it from damage and dust. One loop was placed around the neck, and one around the tail base, connected with 1.5 cm teflon tape. The posterior loop was placed rostral of the cloacal area, and the knot rested above the furcula. Backpack microphones were placed on the birds at least a week before the start of recording to allow the birds to acclimatise to the equipment [15, 46].

The AM-modulated radio signals sent by the microphones were detected using AOR8600 receivers (AOR, USA). The signal was processed in a 16-channel analogue-to-digital converter (Delta 1010, M-Audio, USA) operated at a sampling rate of 44100 Hz, and recorded using ASIO data streaming environment (Steinberg, Germany; interface adapted by Markus Kramer, MPIO Seewiesen). Each recording channel was stored as .wav file of 4 h duration.

The wireless microphone was mounted on the bird's back, facing the body, thus primarily recording the bird's own vocalisations [15, 37, 46]. On rare occasions, the recordings also included vocalisations emitted by other birds. However, during clustering processes these were clearly recognisable due to different basal frequency intensities, and removed [46].

Repertoire and vocalizations clustering

We classified calls into different categories using previously described criteria [15, 37, 46]. Briefly, we used the custom-written software "Sound Explorer" (see [15] for GitHub address) to analyse the sonograms. For each sonogram we calculated the following parameters: duration, mean frequency, mean frequency standard deviation (SD), mode frequency, mode frequency SD, first peak, first peak SD, zero crossing, maximum positive peak and minimum negative peak. These parameters were used for automatic sorting and the output clustering was subsequently manually refined. We refined clustering using visual features of the sonograms. During the screening the scorer was aware of the treatment (pair experience). However, he/she was blind to the time information used to extrapolate data for statistical analysis. These were automatically assigned and hidden therefore not a type of subjective behavioural recording [48]. Vocalisations were classified according to the criteria described in Zann (1996), with minor modifications (Fig. 1d, bottom panel). We divided vocalisations into 7 categories: 6 types of call (Fig. 1d, bottom panel, Additional file 1) and a separate category for the vocalisations which we were unable to assign to any call type (e.g. rare vocalisations or, since they are intergrading clusters, vocalisations with features of two call types). For males we included two additional categories: song and misplaced introductory syllables (those which were not followed by the song) (Additional file 1). Zebra finches are known to include some of their calls in their song [49], therefore as first step of clustering we ordered all vocalisations in their sequence of occurrence, and were thus able to distinguish which similar call types were used in songs from those present as single calls. Finally, the number of songs was calculated dividing the total number of syllables by the average number of song syllables of each male (see Fig. 2 for birds' daily emission of each call type).

Video recording and scoring

The video recordings were made with small cameras (Handykam Colour 420 line CCD high resolution camera, Handykam.com, Hayle, UK) positioned inside the sound boxes but outside the cage and started automatically with a set schedule. We analysed 30 min in the morning (8:00-8:30) and 30 min in the afternoon (12:00-12:30). The videos were scored with Observer® XT (Version 10, Noldus Information Technology, Wageningen, Netherlands) with the scorer blind to the treatment. The relative position between the two birds was coded as "Clumping" if the birds were in physical contact, "Close" if the distance between individuals was less than one bird, and "Distance" if the subjects were apart (Additional file 2). Additionally, the following behaviours were scored: "Perching – exploring", indicating that the bird was either moving or stationary in the cage; "Hopping" when the bird jumped between perches with less than 3 sec. intervals between hops, and "Preening" when the bird was cleaning its feathers (Additional file 3). Video and audio recordings were synchronised based on recognisable events (e.g. songs and/or the relative time between two vocalizations). Specifically, the audio channel of the video was extracted and aligned manually with the audio from one of the transmitters. Once synchronised and scored, the rate of each call type for each behaviour was calculated.

Cross correlation analysis

We used cross-correlation analysis to determine the synchronisation of pair vocalisations [37]. The onset times of the different vocalisations were used to shape cross-correlational density plots [50], where vocalisations of one individual were aligned with specific vocalisations of individualist partner. As a convention, we designated the female calls as always beginning at time 0. The length of the time window we evaluated was 2 seconds before and after each female call, divided into 100 bins. All the calls emitted half a second before or after the call of the focal individual (the females) were considered answered and answer calls respectively [15, 37]. Answer calls are vocalisations given in response to the focal individual within 0.5 seconds. Answered calls are those that receive a response by the focal in the same time period. We used the number of calls emitted during this interval to

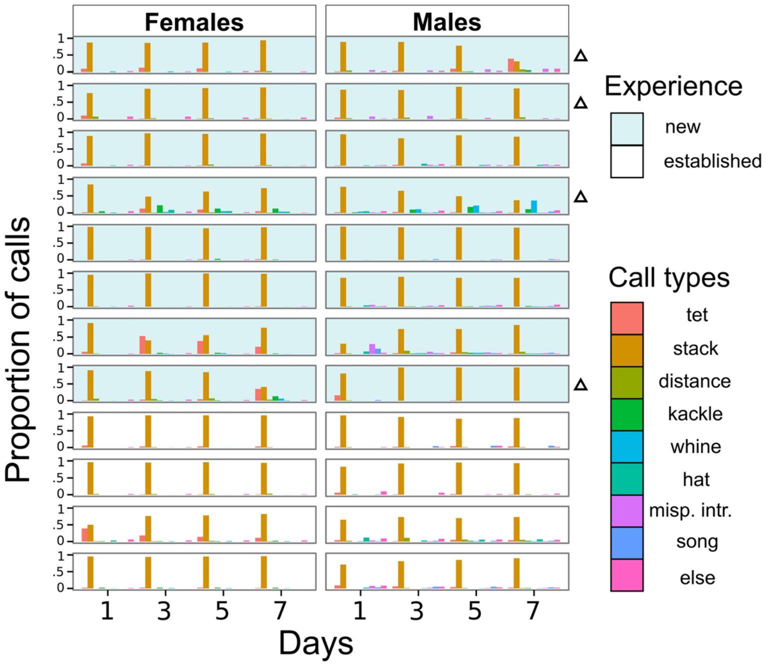


Fig. 2 Proportion of call types by day. The proportion of each call type is reported for all the birds of the study. Each row represents a pair and the two columns are for females (*left*) and males (*right*). Within each column the 4 days of the study are plotted next to each other. The different colours of the bars represent the different call types (Misp. Intr. is the abbreviation of misplaced introductory syllables: those which were not followed by the song). The height of the bars represents the proportion of call types relative to the bird and day (sum for each bird each day equal to 1). The top 8 pairs with the shaded background are pairs that never met before the experiment whereas the bottom 4 are already established pairs. -Triangles indicate new pairs formed by individuals which had never previously bred successfully with another mate

calculate the directionality index as follows: (Number of Answers - Number of Answered) / (Number of Answers + Number of Answered). Therefore the directionality index is 0 when the number of answers is the same for males and females. The female call is the focal stimulus, thus the directionality index is positive if the number of answers is greater for the male, and negative if greater for the female. Confidence limits were calculated using Poisson probabilities based on the baseline levels of the correlation which was defined as the period between 4 and 2 seconds before and 2 and 4 after the focal vocalisations. Hence, it was assumed that calls from these two 2 seconds periods had a random distribution [37].

Statistical analysis

Statistical analyses were conducted in R version 3.2.3 [51] using a Bayesian statistical approach with the packages "arm" [52] and "lme4" [53]. Linear Mixed Models (LMM) were calculated using the maximum-likelihood (ML) method to have a better estimation of the fixed effects [54]. Posterior means and their 95% credible intervals (CrI) were calculated (10000 simulations) using the function "sim". We used flat prior distributions (i.e. it does not influence the posterior distribution of the simulated data), therefore sensitivity analyses of prior distributions were not required. In all cases, the residuals were checked visually for the model fit with the following plots: residual vs. fitted; residual distribution; residual variance vs. fitted. In addition, we visually checked the assumption that the random effects were normally distributed. Tables with the full model results can be found in the Additional file 4. When it was necessary to compare subgroups within an analysis we performed a derived calculation: out of the 10 000 set of simulated parameters we report the number of cases for which the estimated value of the first group was larger than that of the second group, and report this value as "p". The threshold of 5% would be equivalent to significance level in a frequentist framework.

To explain the amount of time spent in physical contact, termed clumping, we included the experience of the pair (categorical, 2 levels) and the day (categorical, 4 levels) as explanatory variables in the LMM, and since we expected that the effect of familiarity changes with time (days), we included the interaction. Pair ID was added as random factor (categorical, 12 levels) (Fig. 3a). Clumping ~ Experience * Day + (1|PairID)

To study the directionality index over time of new and established pairs we ran a LMM with experience and day as explanatory variables. We used the absolute value of the index because we focused in its difference from 0 and not its direction. Because we were interested to know whether the two treatments changed over time we included the interaction between pair experience and day (Fig. 3b). |Directionality| ~ Experience * Day + (1|PairID)

To determine the relative distance at which birds vocalised most frequently, we used a LMM with relative position as the explanatory factor (categorical, 3 levels) of the calling rate. For this model we only considered the calling rate of stack calls. The square root of the calling rate was taken, to achieve normal distribution of the residuals. Pair ID and day were used as random factors to account for repeated measures (Additional file 5). Calling rate ~ Relative position + (1|Day) + (1|PairID)

To study the correlation of the numbers of stack calls between males and females we used 2 LMMs considering either the males' total number of calls or the males' number of answers as outcome variables and the corresponding females' variables as explanatories. Both models had day of recording and pair experience nested into pair ID as random factors (Fig. 4). To represent the two models in the same plot we normalized the data dividing, for each relationship, by the highest number of calls. Total male calls ~ Total female calls + (1|Day) + (1|Experience/PairID); Male answer calls ~ Female answer calls + (1|Day) + (1|Experience/PairID)

To determine if vocalisations were related to behavioural aspects, we modelled the proportion of calls used as answers (out of the total number of calls of the focal individual) as a function of the time spent in physical contact (Fig. 5). We ran 2 separate models for the 2 sexes. For these LMMs, we used the same random effect structure as the models of stack calls described above. Additionally, since measurement units were different and measurement values were several orders of magnitude apart, we standardized both variables using z-scores to simplify the interpretation. We also ran the same model excluding the null clumping values, days in which birds did not clump, to confirm the result. Percentage of answers ~ Clumping + (1|Day) + (1| Experience/PairID). The repeatability of the directionality index was calculated according to Lessells & Boag (1993) [55], where the among-groups variance component describes variance among pairs and the within-group variance component describes the variance within a pair across different days.

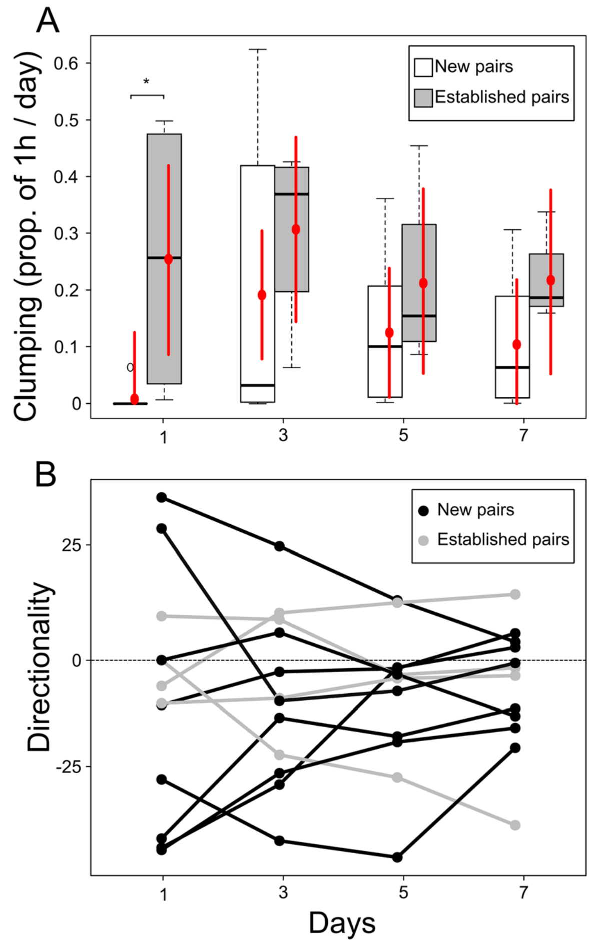


Fig. 3 Social and calling behaviour over time. Proportion of time spent in physical contact (**a**) and directionality index (**b**) of new and established pairs over time. **a**) Amount of physical contact ("clumping"), given as a proportion of the overall time, for each of the four days. The boxplots represent the row value, the red dots the estimated Bayesian values and the red segments the Credible Intervals (CrI) estimated from the LMM. An asterisk indicates a lack of overlap between CrI and the fitted values (equivalent to frequentist significance set at <5%). **b**) Directionality index over time. Each line is a pair; established pairs are shown in grey whereas new pairs are in black. The variance in new pairs is bigger at the beginning than at the end, indicating that the relationship becomes quantitatively more symmetrical

Results

Proportion of different calls by day

We first looked at the proportion of different call types emitted by individual birds exposed to our experimental conditions. We recorded and categorised 475 903 vocalisations. Only a small portion of the vocalisations (mean ± SD per recording; 2.62 ± 2.43%, N = 96) were not assigned to one of the depicted call types (Fig. 1d). In 94 cases out of 96 (12 pairs recorded for four days in a week) the stack call was the most frequently emitted call type (0.84 ± 0.16%, N = 96) (Fig. 2, Additional file 6). This was the same in both new (8-hour recordings where stack calls were the most common call type / total number of recordings; 62/64) and established (32/32) pairs, and it did not change during the observed period (day1: 24/24; day3: 23/24; day5: 24/24; day7: 23/24).

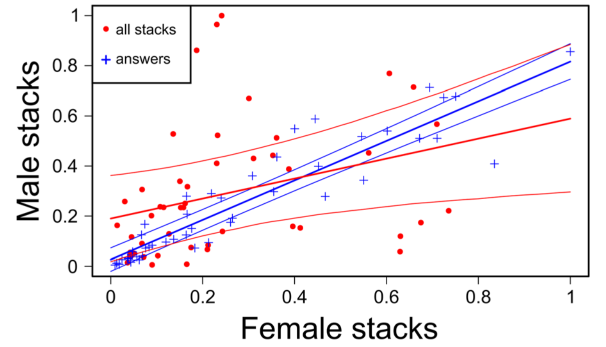


Fig. 4 Correlation of the numbers of stack calls and answers between males and females. Total number of stack calls (*red dots*) and number of stack call replies to stacks (*blue crosses*) for males and females for each day for each pair (values are normalised dividing by the highest number of calls). Bold lines represent the relationships extracted from the estimated posterior of the linear mixed models; thinner lines are the respective 95% credible intervals. Number of answers is tightly correlated between males and females, while call number is only loosely correlated, indicating that different individuals use differing proportions of calls to answer the partner

Stack calls were almost always the most frequently-emitted call in both new and established pairs.

Social behaviour of new and established pairs

We asked whether the proportion of time spent in physical contact (clumping) differed depending on pair experience. We found that new pairs spent very limited time clumping during the first day (time of clumping expressed in seconds and as percentage of the total time scored, mean ± SD; 29 ± 81 sec., 0.8 ± 2.2%, N = 8) (Fig. 3a), whereas established pairs clumped for much longer (919 ± 923 sec, 25.5 ± 25.7%, N = 4). Credible intervals do not overlap zero, indicating a marked difference (899 sec., CrI: 181 – 1620, p = 0.0032). In the following days, the trend that new pairs spent less time clumping continued, but the credible intervals of new and established pairs overlapped (Fig. 3a) indicating that there was no longer a clear difference (day3: p = 0.0832; day5: p = 0.1409; day7: p = 0.0817). Only the established pairs spent a considerable amount of time in physical contact during the first day, whereas the newly introduced pairs spent less time in contact.

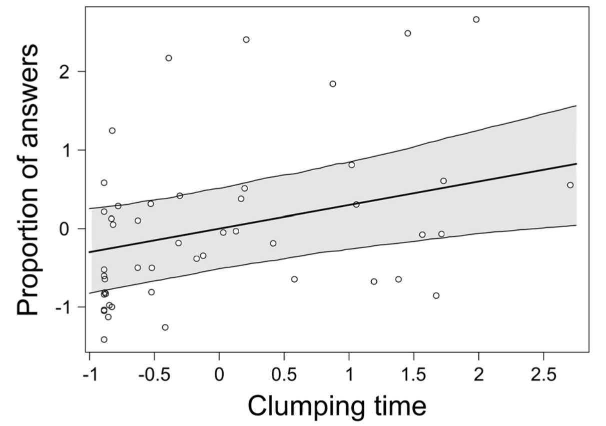


Fig. 5 Correlation between proportion of male answers and time spent in physical contact. Amount of time in physical contact and proportion of answers (standardised by z-scores). The dots are the raw values, one point for each male each day, the bold line is the fitted regression line drawn from the posterior distribution of the value estimated from the LMM. The thinner lines and the shaded area represent the credible intervals

Calling behaviour of new and established pairs

To see if new and established pairs differed in vocal coordination, we used cross-correlations to show whether there was a specific answer to our focal stimulus, the partner calls. Pair members used different combinations of call types to respond to their partner, and these combinations of call types showed a coordinated pattern of replies (i.e. over-threshold number of calls within the analysed time window, see method "cross-correlation") (Additional file 7). However, only one combination, stack-stack, was present among all the studied pairs on each day (Additional file 7, Additional file 8). We confirmed that mates answer to each other with very precise latency and low rate of overlapping calls [37] (Additional file 8). Observing the stack-stack calling we found that both new and established pairs tended to respond to their partners, with pairs varying in the number of replies (antiphonal calls) and total calls (Additional file 6, Additional file 8). The shape of the cross-correlation histogram, which shows the amount of replies of the partners compared to baseline calling, can therefore be used to describe the calling relationship (Additional file 8). The shape can be summarised by the directionality index, which changed from pair to pair (some pairs were asymmetrical, others were symmetrical), and also over time within the pairs. We found very high repeatability (r ± SE: 0.94 ± 0.03, *N* = 12) in the directionality index, indicating that each pair develops a specific calling relationship. The directionality index values (Fig. 3b) of new pairs were very wide in range on the first day (mean ± SD, day 1: -12.27 ± 32.64, *N* = 8) and tended to converge to a more symmetrical relationship over time (day 7: -5.90 ± 10.21, *N* = 8). The absolute value of the directionality index statistically differed between the first and the last day ($p <$ 0.0001, *N* = 8). In contrast, the index of established pairs did not change significantly (day1: -1.43 ± 8.74, day 7: -7.05 ± 22.32; *p* = 0.1011, *N* = 4). Furthermore we observed a more symmetrical relationship of established pairs compared with new ones during the first day (probability that new pairs had higher directionality index than established ones, $p = 0.0006$, $N = 12$); the difference was not significant during the last day ($p = 0.8084$, $N = 12$).

Correlations of total number of stack calls and reply stack calls

We defined motivation to answer the proportion of calls used as answers out of the total number of call emitted. Consequently, to understand whether the motivation to reply differed within and between pairs, we compared the total number of stacks between partners with the proportion of the number of stacks used as replies (to other stacks) (Fig. 4). If the two distributions were similar it would mean that each individual used the same proportion of calls to answer the partner (i.e. the motivation to answer was similar among individuals). In contrast, we found that the two relationships differed greatly in shape and dispersion, indicating that each bird answers to the partner with a different proportion of calls. If the slope of the relationships were 1 and the intercept 0, it would mean that the number of calls, either the total calls or only the replies, was equal between males and females. On the contrary we found a difference between males and females in the total number of calls (estimated regression line; $y = 0.190 + 0.398x$). However, when considering the number of replies, the number of calls used was more similar ($y = 0.026 + 0.790x$). Most interestingly, if the credible intervals (CrI) were narrow it would indicate the use of a similar strategy across pairs. We found that this was the case for the number of answers (0.711 - 0.870), which was very similar between and across pairs, whereas the total number of calls had a wide CrI (0.066 - 0.739) and was only loosely correlated between and across pairs. To further explore the difference between the correlation of number of answers and total number of calls we measured the goodness-of-fit of the models, marginal and conditional r^2-values (i.e. how much of the variance is explained by fixed effects alone and total respectively; [56, 57]). We found that for total amount of calls, marginal (r^2m) and conditional (r^2c) r^2-values were 0.102 and 0.777, whereas for the number of replies, $r^2m = 0.860$ and $r^2c = 0.943$. We found that the experience of the pair did not explain any variance and most of the variance explained by the random factors was due to differences between pairs. Furthermore, for the model including the total number of calls, the residuals against the random factor "day" showed a specific pattern. This probably was because they called much less during the first day than predicted from the model.

Relationship between clumping time and proportion of answers

As we found that different individuals answer with different percentages of calls (Fig. 4) we tested the correlation of vocal with the social behaviour. We combined information from the video and audio recordings to calculate the relationship between the time spent in physical contact (mean ± SD expressed in seconds; 556 ± 627 sec., N = 12) and the proportion of replies of the males' stack calls (expressed in %, 15 ± 10.1%, N = 12) (Fig. 5) and of the females' stack calls (17.7 ± 10.8%, N = 12) on their total number of stacks. A higher proportion of calls used as answers might reflect a higher motivation in answering, and also a longer time spent in clumping might reflect a stronger motivation to stay in contact. We found a positive relationship between standardised time spent clumping and the proportion of replies (after standardisation, see "statistical analysis" in methods, slope: 0.300, CrI: 0.097 - 0.499; Fig. 5). This means that each increasing unit of clumping time (expressed as change in standard deviation) yields an increased expected proportion of replies by about a third of a unit. For instance, an increase of 627 sec. of clumping time would lead to an increase of 3% in the proportion of calls that a male uses as replies. The number of 0 s in the clumping values may bias the model. Interestingly, also excluding occurrences in which the birds did not clump, yielded to a very similar result (slope: 0.282, CrI: 0.073 - 0.492). For the proportion of calls used by females as answers, this relationship was not as strong (slope: 0.068, CrI: -0.182 - 0.315). Interestingly, the experience of the pair did not explain any variance in the proportion of answers of male or female stack call. We conclude that clumping time can predict, to a certain extent, the proportion of replies of the male.

Discussion

Antiphonal calling with stacks was a common feature for all zebra finch pairs in this study; this alternating calling behaviour was temporally precise, characterised by a very low rate of overlap between calls and a high level of alternation between mates. As both new and established pairs displayed this behaviour, we propose that this pattern of alternating stack calls could define a private channel of communication between mates in non-breeding situations, possibly a display of monogamous pairs [58]. Physical contact, termed clumping, has been used by many studies as a behavioural indicator of bonded pairs [59–61]. As expected, when comparing new and established pairs, we confirmed that only the latter spend time in physical contact during the first day [43]. Interestingly, the difference in social behaviour between new and established pairs is mirrored by a difference in calling patterns. We found that both new and established pairs exchanged stack calls; however, new pairs were more variable with regard to the directionality of the calling relationship during the first day (i.e. often the relationships are asymmetrical, meaning that one member answered more than the other). A week later, these new pairs had more symmetrical calling relationships. We can describe the observed pattern as behavioural convergence, labelling it as post-pairing adjustment [62]. The number of stack calls of males and females within a pair was loosely correlated, whereas the number of calls used to answer the partner was similar between pair members. This suggests that quantitatively pairs had a balanced vocal exchange, but each bird used a specific and different percentage of calls to answer the partner. The percentage of replies ('answer calls') by the males positively correlates with the amount of clumping exhibited by the pair. We tentatively interpret this as stronger motivation towards the partner expressed by both vocalisation and affiliative behaviour. The vocal exchange of stack calls did not occur when the birds were clumping, but rather when they were distant from each other. This suggests a function of vocal exchange during locomotion using this call type, perhaps when birds are relatively close [28, 29].

Previous studies have already described antiphonal calling involving stacks [15, 37]. Here, we added a detailed and quantitative description of the pattern of stack call usage during establishment of new pairs and the consistency of stack usage over time in already established pairs. High behavioural similarity between partners may make cooperation more effective, and may have fitness consequences in species with bi-parental care (part of the "mate familiarity effect", reviewed [10]). Coordination of other behaviours, such as provisioning and foraging, has been found to be beneficial in zebra finches in the wild [13]. The antiphonal exchange might aid the coordination between partners, and possibly improve decision making processes (e.g. during foraging behaviour). Alternatively, or additionally, as a display that continues after the formation of the pair bond, it could be important for pair maintenance [58, 63], or potentially support mate guarding. However, whether the symmetrical communication has a functional value is still an open question. Experiments measuring fitness parameters are necessary to answer it.

We observed a large difference between marginal and conditional r^2-values in the model correlating the total number of calls of males and females. The pattern of the residuals vs. random effect "day" did not follow a normal distribution and partially explains this result: the daily number of calls changed during the experiment and changed differentially for males and females. On the contrary, the random factor "day" did not explain any variance in the model correlating the number of answers. Therefore, showing that the answering relationship was acquired early in the pair development and the proportion of answers remained consistent over time despite the change in the overall amount of calls. The pair forms quickly within the first days [45], and we showed that in conjunction, the vocal relationship stabilized early. This partially explains why the experience of the pair did not influence either the relationship between number of calls and number of answers or the relationship between the latter parameter and the amount of clumping. However, the very high variability between pairs and the small sample size might mask the differences between groups. From the analysis comparing number of answers and number of calls it is possible to draw further conclusions. We observed a high behavioural similarity between paired males and females in the number of calls used to answer the partner [37]. However, when we considered the total amount of calls produced, we observed asymmetry between partners. This might reflect a different motivational state and interest of the birds towards their partner [64], since in our experimental design, individuals could not choose their partners. The quality of the match might therefore differ substantially among pairs, producing different patterns of calling and replying. In addition, we found that clumping time predicts the percentage of answers of the male. Males that spent more time in physical contact also used a higher proportion of their calls to answer their partner. Hence antiphonal calling could be tested as an indicator of pair compatibility. To find methods to quantify pair quality, compatibility, is very important since fitness can depend on it [2, 5]. Further, specifically designed experiments involving mate choice are needed to better clarify the relationship between answering rate and clumping. Likewise latency to the first occurrence of affiliative behaviours, such as clumping, allopreening, copulation, and their quantity, can be studied in correlation to fitness and vocal behaviour to find what factors better predict pair compatibility.

The vocal repertoire was similar among different birds and the most common call type for isolated pairs in a non-breeding situation was almost invariably the stack call. Different authors have reported other calls to be the most common; Zann (1996, ch. 10, p. 197) described the tet, others the distance call [65, 66]. This could be due to the context in which the recordings were made, or the tools used to record vocalisations. For example, the cited studies used an external microphone that might have failed to detect vocalisation with low amplitude [65, 66]. Also, the birds used by Zann were in groups and allowed to breed. That breeding status affects the type of calls that are emitted is supported by Gill et al. (2015), which shows a change in abundance according to the context; this could alter the relevance of some calls according to the deviance (i.e. the relative abundance of a particular call type [67]). Furthermore, Gill et al. (2015) showed that in a group situation, in contrast with our isolated pairs, the stack call is not always the most common type and other call type combinations other than stack-stack were always present between pair members. These differences suggest a social meaning for other call types and call combinations (i.e. some calls and call combinations might have a message for members of the group other than the mate).

Given the diversity of calls and their different uses, studying the temporal relationships of vocalisations could improve our understanding of complex communication [68] and the 'linguistics' of calling. Most language usage is interactive, involving rapid turn-taking characterised by short turns and very rapid responses [69]; zebra finch vocal exchange mirrors this pattern. The zebra finch calling system is clearly lacking flexibility in its messages compared to that of humans; however, turn-taking patterns and tempos of the different systems can be compared. Hence investigating the dynamic pattern of calling could help to understand the role of turn-taking in vocal communication [70]. The first step in this direction would be to verify that these calls are enough to identify the caller, making possible to select the interlocutor. Moreover, this fast exchange model of vocal communication can be investigated from the point of view of behavioural neurobiology [37]. During antiphonal calling, a bird must provide the specific appropriate response within a few milliseconds of an auditory stimulus. When the bird hears a call, it needs to process it: that is extract the type of call and the calling individual's identity, recall the memory of that individual, and choose and utter an extremely rapid response, which makes our system ideal for investigating processing recognition and answer choice.

Our approach, with the use of backpack microphones and continuous recording, allowed an extremely high level of precision and accuracy in our measurements. However, despite the high repeatability of the turn-taking behaviour, due mainly to the time consuming procedures, the sample size is a limitation of our study, and it is therefore difficult to confidently generalise all of our results. Nevertheless, we are convinced that the results and the approach presented here may spur further research on calling patterns because of its relevance for different fields.

Conclusion

Here, we document differences between new and established zebra finch pairs, shedding light on the role of alternating (antiphonal) calling. Members of both new and established pairs use stack calls to answer their partner. While birds in established pairs respond to their partner with equal number of calls, the newly formed pairs begin with one bird calling more, but then develop a calling relationship that becomes more symmetrical over time. We therefore found post-pairing behavioural convergence between pair members, whereby they adjusted the number of calls used to answer their partner. In addition, within both groups, pairs differed from each other, but were internally consistent. Furthermore, in males, reply frequency was positively correlated with the time spent in physical contact with their partner. The high repeatability together with the possible reflection of a motivational state leads us to postulate that the study of such calling relationships might add information on pair compatibility. We propose that the patterned exchange of vocalisations may represent a fundamental part of the pair bond, and may serve as a private channel of communication within the pair.

Additional files

Additional file 1: Audio examples of the call repertoire of the zebra finch. (same pair as Fig. 1). Five calls for each sex and call type are spaced by one second silences. We randomly selected calls to be presented from the ones not containing noise. Sounds were recorded with backpack microphones and their amplitude normalised to -0.1 dB (maximal sample value). (7Z 240 kb)

Additional file 2: Video example for each behaviour scored for the relative position. 30 seconds for each behaviour are shown. (MOV 8114 kb)

Additional file 3: Video example for each behaviour scored for individual behaviour. 15 seconds for each behaviour are shown. (MOV 4853 kb)

Additional file 4: Tables with estimates, Standard Errors (SE), Credible Intervals (CrI) and random factor and residual variance of each LMM used. First is presented model structure, together with mean and SD of the raw data. The graphical representation, if present, is referred after the model. (DOCX 24 kb)

Additional file 5: Figure S1. Rate of stack calls during different relative position of pair members. Rate of stack calls (n/sec) for the 3 different relative positions scored. Clumping: the pair is in physical contact. Close: the space between the birds is less than one bird. Distance: the birds are apart. Boxplots are drawn using raw data, the red dots are the estimated Bayesian values and the red segments the Credible Intervals (CrI) estimated from the LMM.. Both males and females used different rates of stack calls depending on their relative position (data not shown for each sex separately). While in physical contact, (clumping), the birds called the least (measured in calls/sec., mean ± SD, 0.090 ± 0.115 calls/sec., N = 12), followed by close proximity, (close), (0.160 ± 0.126 calls/sec, N = 12); whereas when they were spatially separated, (distance), the pairs had a higher rate of calling (0.243 ± 0.183 calls/sec., N = 12). Using the output of the LMM we calculated the probability that estimated values of one of the relative positions would be higher than the ones of another; asterisks indicate $p < 0.05$. We found that the probability that the calling rate during "Clumping" was higher than "Close" was $p = 0.0074$, and "Close" higher than "Distance" was $p = 0.0556$, indicating strong differences between these categories. Thus, relative position influences the amount of elicited calls. (PNG 333 kb)

Additional file 6: Table S1. Total and proportion of different call type by bird. Total number and proportion (in brackets) of each call type for each individual. The total number of calls is the mean of the four days of recording and the proportion is calculated from this mean. Misp. Intr. is the abbreviation of misplaced introductory syllables (those which were not followed by the song). (XLSX 11 kb)

Additional file 7: Table S2. Response strength for each pair each day each combination. Within a time window of interest of 4 s before and 4 s after call onset for our cross-correlation histograms we counted the number of calls. We divided the time window with a binwidth of 50 ms (i.e. 160 bins in total). The number of calls in the bins in the first 0.5-s (Nbase, i.e. the calls between 4 and 3.5 seconds before the focal calls) was used as baseline and those in the 0.5-s bins after call onset (Nresponse) as the response. We calculated the response strength index for each call combination for each pair for each day as follows: Rresponse = (Nresponse −Nbase) / (Nresponse + Nbase). The index range between -1 and +1, positive values correspond to an increase of calling after the stimuli (partner calls) compared to the baseline, negative value to an inhibition of calling, values close to 0 to maintenance of baseline calling (all the values are multiplied by 100 to aid readability). We set a threshold to avoid weak correlations to bias the results: if the number of calls within the time window considered did not reach 160 (i.e. one call for each bin) the index was not computed (reported as 0). (XLSX 31 kb)

Additional file 8: Figure S2. Stack-stack cross-correlation for each pair. Each row represents a pair, and each column shows a different day of the experiment. For each scored day per pair, a cross-correlation graph [50] is presented of the stack-stack call relationship between male and female. Cross-correlation histograms show the temporal correlation between one male and one female call type within a given time window. Histograms were aligned on female vocalisations. The y-axis represents the number of calls, normalised by the bin with the highest number of occurrences (between 0 and 1). The interval considered on the x-axis is 0 ± 2 sec. The 0.99 Poisson confidence limits are shown with horizontal red lines [37]. Typically, the shape of the histogram is characterized by a sharp inhibition in the bins next to the 0, because of the little overlap between calls, and a spike of events within 0.5 sec, often over the set confidence interval. Therefore, the calls in the window within ± 0.5 sec. from the focal calls are considered as replies and coloured according to the sex. The replies of the males are depicted in orange and females in grey. Within each cross-correlation the numbers on the top represent the total amount of stack calls over the 8 hours of recording, and the number used to reply to the stack calls of the partner. (PNG 3557 kb)

Acknowledgments

We want to thank Hannes Sagunsky and Markus Abels for technical support. Further, we thank Wolfgang Goymann, Manfred L. Gahr, Maude Baldwin and 2 anonymous referees for their helpful comments on previous versions of the manuscript. We thank the International Max Planck Research School (IMPRS) for Organismal Biology for training and support.

Funding

This work was funded by the Max-Planck-Gesellschaft.

Availability of data and material

Raw data have not been uploaded and made publicly available because they are extremely large files but are available from the corresponding author on reasonable request (examples are available as Additional files 1, 2, 3, 4). The generated and/or analysed data during the current study are made available as Additional file 6 and 7 and the rest are available from the corresponding author on reasonable request.

Authors' contributions

PBD, LT & AT conceived the study and designed the experiment. PBD performed the experiment. PBD processed the data. PBD & AT analysed the data with input from all authors. PBD wrote the manuscript with input from all authors. AT & LT provided supervision during the entire project. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval

The use of audio transmitters and bird maintenance in sound-proof boxes was approved by the government of Upper Bavaria, "Sachgebiet 54 – Verbraucherschutz, Veterinärwesen, 80538 München" with the record number: Az. 55.2-1-54-231-25-09. All animal husbandry and handling was conducted according to the directives 2010/63/EU on the protection of animals used for scientific purposes.

The application of the backpack was fast (time comparable to [46]), and its long battery life meant that birds did not have to be handled again for the duration of the experiment [46]. The habituation phase of one week was chosen based on previous experiments to ensure that the birds had fully recovered from any behavioural effects of backpack application and carrying [15]. Although birds would at times peck at the backpack, no aberrant behaviours were noted during the daily check (see methods, "study animal and recording scheme"), nor while scoring videos. No signs of injury were noted when removing the backpacks. After completion of the experiment, the birds returned to our breeding colony.

Received: 18 October 2016 Accepted: 14 February 2017

References

1. Nisbet ICT, Dann P. Reproductive performance of little penguins *Eudyptula minor* in relation to year, age, pair-bond duration, breeding date and individual quality. J Avian Biol. 2009;40:296–308.
2. Ihle M, Kempenaers B, Forstmeier W. Fitness benefits of mate choice for compatibility in a socially monogamous species. PLoS Biol. 2015;13:e1002248.
3. Leu ST, Burzacott D, Whiting MJ, Bull CM. Mate familiarity affects pairing behaviour in a long-term monogamous lizard: evidence from detailed bio-logging and a 31-year field study. Ethology. 2015;121:760–8.
4. Sánchez-Macouzet O, Rodríguez C, Drummond H. Better stay together: pair bond duration increases individual fitness independent of age-related variation. Proc R Soc B Biol Sci. 2014;281.
5. Spoon TR, Millam JR, Owings DH. The importance of mate behavioural compatibility in parenting and reproductive success by cockatiels, Nymphicus hollandicus. Anim Behav. 2006;71:315–26.
6. Bradley JS, Wooller RD, Skira IJ, Serventy DL. The influence of mate retention and divorce upon reproductive success in short-tailed shearwaters *Puffinus tenuirostris*. J Anim Ecol. 1990;59:487–96.
7. Ens BJ, Safriel UN, Harris MP. Divorce in the long-lived and monogamous oystercatcher, *Haematopus ostralegus*: incompatibility or choosing the better option? Anim Behav. 1993;45:1199–217.
8. van de Pol M, Heg D, Bruinzeel LW, Kuijper B, Verhulst S. Experimental evidence for a causal effect of pair-bond duration on reproductive performance in oystercatchers (*Haematopus ostralegus*). Behav Ecol. 2006;17:982–91.
9. Griggio M, Hoi H. An experiment on the function of the long-term pair bond period in the socially monogamous bearded reedling. Anim Behav. 2011;82:1329–35.
10. Black JM, Hulme M. Partnerships in birds: the study of monogamy. UK: Oxford University Press; 1996.
11. Mariette MM, Griffith SC. Nest visit synchrony is high and correlates with reproductive success in the wild Zebra finch *Taeniopygia guttata*. J Avian Biol. 2012;43:131–40.
12. Bebbington K, Hatchwell BJ. Coordinated parental provisioning is related to feeding rate and reproductive success in a songbird. Behav Ecol. 2016;27:652–9.
13. Mariette MM, Griffith SC. The adaptive significance of provisioning and foraging coordination between breeding partners. Am Nat. 2015;185:270–80.
14. Raihani NJ, Nelson-Flower MJ, Moyes K, Browning LE, Ridley AR. Synchronous provisioning increases brood survival in cooperatively breeding pied babblers. J Anim Ecol. 2010;79:44–52.
15. Gill LF, Goymann W, Ter Maat A, Gahr M. Patterns of call communication between group-housed zebra finches change during the breeding cycle. Elife. 2015;4:e07770.
16. Boucaud ICA, Mariette MM, Villain AS, Vignal C. Vocal negotiation over parental care? Acoustic communication at the nest predicts partners' incubation share. Biol J Linn Soc. 2015;117:322–36.
17. Hall ML. Chapter 3 a review of vocal duetting in birds. Adv Study Behav. 2009;40:67–121.
18. Mennill DJ, Vehrencamp SL. Context-dependent functions of avian duets revealed by microphone-array recordings and multispeaker playback. Curr Biol. 2008;18:1314–9.
19. Wickler W. Duetting songs in birds: Biological significance of stationary and non-stationary processes. J Theor Biol. 1976;60:493–7.
20. Hall ML, Peters A. Coordination between the sexes for territorial defence in a duetting fairy-wren. Anim Behav. 2008;76:65–73.
21. Koloff J, Mennill DJ. The responses of duetting antbirds to stereo duet playback provide support for the joint territory defence hypothesis. Ethology. 2013;119:462–71.
22. Templeton CN, Ríos-Chelén AA, Quirós-Guerrero E, Mann NI, Slater PJB. Female happy wrens select songs to cooperate with their mates rather than confront intruders. Biol Lett. 2013;9:20120863.
23. Benedict L. California towhee vocal duets are multi-functional signals for multiple receivers. Behaviour. 2010;147:953–78.
24. Balsby TJS, Scarl JC. Sex-specific responses to vocal convergence and divergence of contact calls in orange-fronted conures (*Aratinga canicularis*). Proc R Soc B Biol Sci. 2008;275:2147–54.

25. Immelmann K. Drought adaptations in Australian desert birds. Proc. 18th International Ornithological Congress. 1963. p. 649–57.
26. Griffith SC, Holleley CE, Mariette MM, Pryke SR, Svedin N. Low level of extrapair parentage in wild zebra finches. Anim Behav. 2010;79:261–4.
27. Butterfield PA. The pair bond in the zebra finch. In: Crook JH, editor. Soc. Behav. birds Mamm. Academic Press London; 1970. 249–78.
28. Zann R. The zebra finch : a synthesis of field and laboratory studies. Oxford: Oxford University Press; 1996.
29. Elie JE, Theunissen FE. The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. Anim Cogn. 2016;19:285–315.
30. Forstmeier W, Birkhead TR. Repeatability of mate choice in the zebra finch: consistency within and between females. Anim Behav. 2004;68:1017–28.
31. Jones KM, Monaghan P, Nager RG. Male mate choice and female fecundity in zebra finches. Anim Behav. 2001;62:1021–6.
32. Tomaszycki ML, Adkins-Regan E. Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. Anim Behav. 2005;70:785–94.
33. Collins SA, Hubbard C, Houtman AM. Female mate choice in the zebra finch ? The effect of male beak colour and male song. Behav Ecol Sociobiol. 1994;35:21–5.
34. Riebel K. Chapter 6 song and female mate choice in zebra finches: a review. Adv Study Behav. 2009;40:197–238.
35. Bolund E, Schielzeth H, Forstmeier W. Singing activity stimulates partner reproductive investment rather than increasing paternity success in zebra finches. Behav Ecol Sociobiol. 2012;66:975–84.
36. Adkins-Regan E, Tomaszycki M. Is male song quality important in maintaining pair bonds? Behaviour. 2006;143:549–67.
37. Ter Maat A, Trost L, Sagunsky H, Seltmann S, Gahr M. Zebra finch mates use their forebrain song system in unlearned call communication. PLoS One. 2014;9:e109334.
38. Blaich CF, Norman M, Syud FA, Benitez G, Frost J, Ravenscroft J, et al. The use of distance calls to maintain pair contact in zebra finches (*Taeniopygia guttata*). Bird Behav. 1996;11:25–30.
39. Blaich CF, Steury KR, Pettengill P, Mahoney KT, Guha A. Temporal patterns of contact call interactions in pair-bonded domestic zebra finches (Taeniopygia guttata). Bird Behav. 1996;11:59–69(11).
40. Elie JE, Mariette MM, Soula HA, Griffith SC, Mathevon N, Vignal C. Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? Anim Behav. 2010;80:597–605.
41. Perez EC, Fernandez MSA, Griffith SC, Vignal C, Soula HA. Impact of visual contact on vocal interaction dynamics of pair-bonded birds. Anim Behav. 2015;107:125–37.
42. Villain AS, Fernandez MSA, Bouchut C, Soula HA, Vignal C. Songbird mates change their call structure and intrapair communication at the nest in response to environmental noise. Anim Behav. 2016;116:113–29.
43. Silcox AP, Evans SM. Factors affecting the formation and maintenance of pair bonds in the zebra finch, Taeniopygia guttata. Anim Behav. 1982;30:1237–43.
44. Caryl PG. Sexual behaviour in the zebra finch *Taeniopygia guttata*: response to familiar and novel partners. Anim Behav. 1976;24:93–107.
45. Pedersen A, Tomaszycki ML. Oxytocin antagonist treatments alter the formation of pair relationships in zebra finches of both sexes. Horm Behav. 2012;62:113–9.
46. Gill LF, D'Amelio PB, Adreani NM, Sagunsky H, Gahr MC, ter Maat A. A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. Methods Ecol Evol. 2016;7:1349–58.
47. Swaddle JP, Page LC. High levels of environmental noise erode pair preferences in Zebra finches: implications for noise pollution. Anim Behav. 2007;74:363–8.
48. Tuyttens FAM, de Graaf S, Heerkens JLT, Jacobs L, Nalon E, Ott S, et al. Observer bias in animal behaviour research: Can we believe what we score, if we score what we believe? Anim Behav. 2014;90:273–80.
49. Zann R. Structure, sequence and evolution of song elements in wild Australian Zebra finches. Auk. 1993;110:702–15.
50. Abeles M. Quantification, smoothing, and confidence limits for single-units' histograms. J Neurosci Methods. 1982;5:317–25.
51. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for statistical computing; 2016. http://www.R-project.org/.
52. Gelman A, Su YS. Arm: Data analysis using regression and multilevel/hierarchical models. R package version 1.9-3. 2016. https://CRAN.R-project.org/package=arm.
53. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using {lme4}. J Stat Softw. 2015;67:1–48.
54. Zuur AF, Ieno EN, Elphick CS. A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol. 2010;1:3–14.
55. Lessells CM, Boag PT. Unrepeatable Repeatabilities: a common mistake. Auk American Ornithologists' Union. 1987;104:116–21.
56. Nakagawa S, Schielzeth H. A general and simple method for obtaining R 2 from generalized linear mixed-effects models. Methods Ecol Evol. 2013;4:133–42.
57. Barton K. MuMIn: Multi-model inference. R package version 1.15.6. 2016. https://CRAN.R-project.org/package=MuMIn.
58. Wachtmeister C-A. Display in monogamous pairs: a review of empirical data and evolutionary explanations. Anim Behav. 2001;61:861–8.
59. Witte K. Time spent with a male is a good indicator of mate preference in female zebra finches. Ethol Ecol Evol. 2006;18:195–204.
60. Cate CT. Directed song of male zebra finches as a predictor of subsequent intra- and interspecific social behaviour and pair formation. Behav Processes. 1985;10:369–74.
61. Tomaszycki ML, Banerjee SB, Adkins-Regan E. The role of sex steroids in courtship, pairing and pairing behaviors in the socially monogamous zebra finch. Horm Behav. 2006;50:141–7.
62. Laubu C, Dechaume-Moncharmont F-X, Motreuil S, Schweitzer C. Mismatched partners that achieve postpairing behavioral similarity improve their reproductive success. Sci Adv. 2016;2:e1501013.
63. Servedio MR, Price TD, Lande R. Evolution of displays within the pair bond. Proc R Soc B Biol Sci. 2013;280:20123020.
64. Iwasaki M, Poulsen TM, Oka K, Hessler NA. Sexually dimorphic activation of dopaminergic areas depends on affiliation during courtship and pair formation. Front Behav Neurosci. 2014;8:210.
65. Vignal C, Mathevon N, Mottin S. Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. Behav Processes. 2008;77:191–8.
66. Vignal C, Mathevon N, Mottin S. Audience drives male songbird response to partner's voice. Nature. 2004;430:448–51.
67. Beckers GJL, Gahr M. Large-scale synchronized activity during vocal deviance detection in the zebra finch auditory forebrain. J Neurosci. 2012; 32:10594–608.
68. Arnold K, Zuberbühler K. Language evolution: semantic combinations in primate calls. Nature. 2006;441:303.
69. Levinson SC. Turn-taking in Human Communication – Origins and Implications for Language Processing. Trends Cogn Sci. 2015;20:6–14.
70. Henry L, Craig AJFK, Lemasson A, Hausberger M. Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. Front Psychol. 2015;6:1416.

Supplementary Materials for

Vocal exchanges during pair formation and maintenance in the zebra finch (*Taeniopygia guttata*)

Pietro B. D’Amelio*, Lisa Trost and Andries ter Maat

*Correspondence to: pdamelio@orn.mpg.de; pie.damelio@gmail.com

The following content can be retrieved at the following DOI: 10.1186/s12983-017-0197-x

Additional file 1. Audio examples of the call repertoire of the zebra finch. (same pair as Figure 1). Five calls for each sex and call type are spaced by one second silences. We randomly selected calls to be presented from the ones not containing noise. Sounds were recorded with backpack microphones and their amplitude normalised to -0.1 dB (maximal sample value).

Additional file 2. Video example for each behaviour scored for the relative position. 30 seconds for each behaviour are shown.

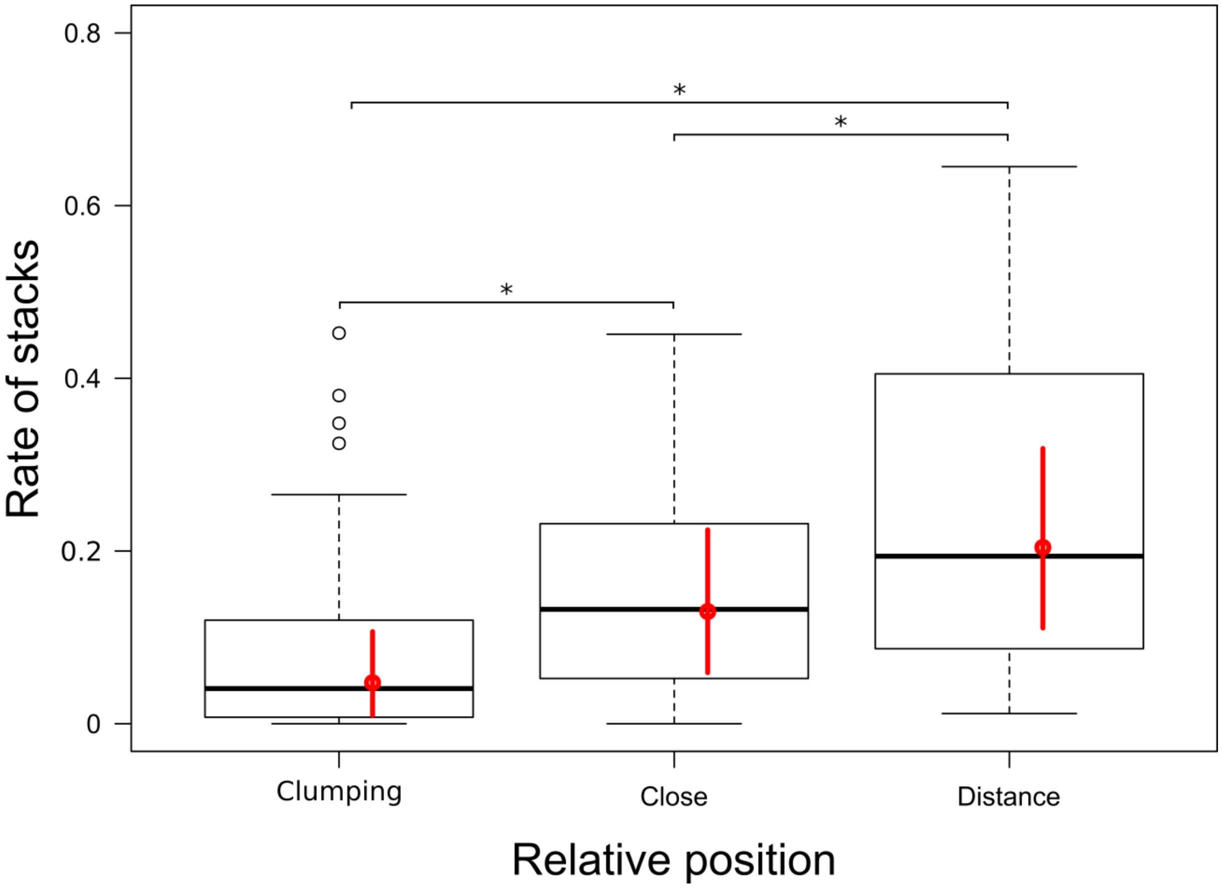
Additional file 3. Video example for each behaviour scored for individual behaviour. 15 seconds for each behaviour are shown.

Additional file 4: Estimates, Standard Errors (SE), Credible Intervals (CrI) and random factor variance of each LMM used. First is presented model structure, together with mean and SD of the raw data. The graphical representation, if present, is referred after the model.

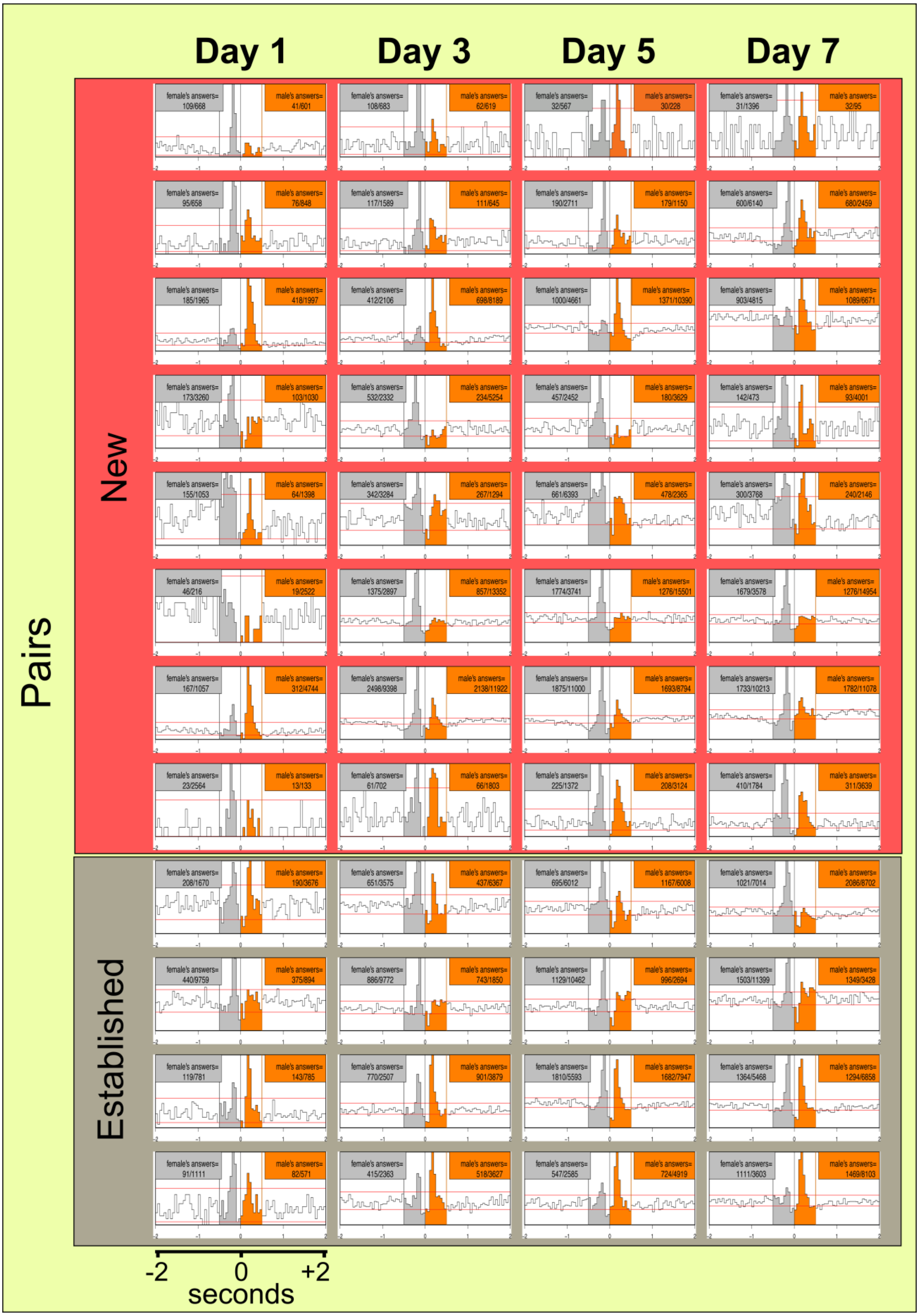
Additional file 6: Table S1. Total and proportion of different call type by bird. Total number and proportion (in brackets) of each call type for each individual. The total number of calls is the mean of the four days of recording and the proportion is calculated from this mean. Misp. Intr. Is the abbreviation of misplaced introductory syllables (those which were not followed by the song).

Additional file 7: Table S2. Response strength for each pair each day each combination. Within a time window of interest of 4 s before and 4 s after call onset for our cross-correlation histograms we counted the number of calls. We divided the time window with a binwidth of 50 ms (i.e. 160 bins in total). The number of calls in the bins in the first 0.5-s (Nbase, i.e. the calls between 4 and 3.5 seconds before the focal calls) was used as baseline and those in the 0.5-s bins after call onset (Nresponse) as the response. We calculated the response strength index for each call combination for each pair for each day as follows: Rresponse =(Nresponse –Nbase) / (Nresponse +Nbase). The index range between -1 and +1, positive values correspond to an increase of calling after the stimuli (partner calls) compared to the baseline, negative value to an inhibition of calling, values close to 0 to maintenance of baseline calling (all the values are multiplied by 100 to aid readability). We set a threshold to avoid weak correlations to bias the results: if the number of calls within the time window considered did not reach 160 (i.e. one call for each bin) the index was not computed (reported as 0).

Additional File 5



Additional file 5, Figure S1: Rate of stack calls during different relative position of pair members. Rate of stack calls (n/sec) for the 3 different relative positions scored. Clumping: the pair is in physical contact. Close: the space between the birds is less than one bird. Distance: the birds are apart. Boxplots are drawn using raw data, the red dots are the estimated Bayesian values and the red segments the Credible Intervals (CrI) estimated from the LMM.. Both males and females used different rates of stack calls depending on their relative position (data not shown for each sex separately). While in physical contact, (clumping), the birds called the least (measured in calls/ sec., mean ± SD, 0.090 ± 0.115 calls/sec., N = 12), followed by close proximity, (close), (0.160 ± 0.126 calls/sec, N = 12); whereas when they were spatially separated, (distance), the pairs had a higher rate of calling (0.243 ± 0.183 calls/sec., N = 12). Using the output of the LMM we calculated the probability that estimated values of one of the relative positions would be higher than the ones of another; asterisks indicate $p < 0.05$. We found that the probability that the calling rate during “Clumping” was higher than “Close” was p = 0.0074, and “Close” higher than “Distance” was p = 0.0556, indicating strong differences between these categories. Thus, relative position influences the amount of elicited calls.


Day 1
Day 3
Day 5
Day 7
Pairs
New
Established
female's answers= 109/668
male's answers= 41/601
female's answers= 108/683
male's answers= 62/619
female's answers= 32/567
male's answers= 30/228
female's answers= 31/1396
male's answers= 32/95
female's answers= 95/658
male's answers= 76/848
female's answers= 117/1589
male's answers= 111/645
female's answers= 190/2711
male's answers= 179/1150
female's answers= 600/6140
male's answers= 680/2459
female's answers= 185/1965
male's answers= 418/1997
female's answers= 412/2106
male's answers= 698/8189
female's answers= 1000/4661
male's answers= 1371/10390
female's answers= 903/4815
male's answers= 1089/6671
female's answers= 173/3260
male's answers= 103/1030
female's answers= 532/2332
male's answers= 234/5254
female's answers= 457/2452
male's answers= 180/3629
female's answers= 142/473
male's answers= 93/4001
female's answers= 155/1053
male's answers= 64/1398
female's answers= 342/3284
male's answers= 267/1294
female's answers= 661/6393
male's answers= 478/2365
female's answers= 300/3768
male's answers= 240/2146
female's answers= 46/216
male's answers= 19/2522
female's answers= 1375/2897
male's answers= 857/13352
female's answers= 1774/3741
male's answers= 1276/15501
female's answers= 1679/3578
male's answers= 1276/14954
female's answers= 167/1057
male's answers= 312/4744
female's answers= 2498/9398
male's answers= 2138/11922
female's answers= 1875/11000
male's answers= 1693/8794
female's answers= 1733/10213
male's answers= 1782/11078
female's answers= 23/2564
male's answers= 13/133
female's answers= 61/702
male's answers= 66/1803
female's answers= 225/1372
male's answers= 208/3124
female's answers= 410/1784
male's answers= 311/3639
female's answers= 208/1670
male's answers= 190/3676
female's answers= 651/3575
male's answers= 437/6367
female's answers= 695/6012
male's answers= 1167/6008
female's answers= 1021/7014
male's answers= 2086/8702
female's answers= 440/9759
male's answers= 375/894
female's answers= 886/9772
male's answers= 743/1850
female's answers= 1129/10462
male's answers= 996/2694
female's answers= 1503/11399
male's answers= 1349/3428
female's answers= 119/781
male's answers= 143/785
female's answers= 770/2507
male's answers= 901/3879
female's answers= 1810/5593
male's answers= 1682/7947
female's answers= 1364/5468
male's answers= 1294/6858
female's answers= 91/1111
male's answers= 82/571
female's answers= 415/2363
male's answers= 518/3627
female's answers= 547/2585
male's answers= 724/4919
female's answers= 1111/3603
male's answers= 1469/8103
-2
-1
0
1
2
-2
0
+2
seconds

Additional file 5, Figure S2: Stack-stack cross-correlation for each pair. Each row represents a pair, and each column shows a different day of the experiment. For each scored day per pair, a cross-correlation graph [50] is presented of the stack-stack call relationship between male and female. Cross-correlation histograms show the temporal correlation between one male and one female call type within a given time window. Histograms were aligned on female vocalisations. The y-axis represents the number of calls, normalised by the bin with the highest number of occurrences (between 0 and 1). The interval considered on the x-axis is 0 ± 2 sec. The 0.99 Poisson confidence limits are shown with horizontal red lines [37]. Typically, the shape of the histogram is characterized by a sharp inhibition in the bins next to the 0, because of the little overlap between calls, and a spike of events within 0.5 sec, often over the set confidence interval. Therefore, the calls in the window within ± 0.5 sec. from the focal calls are considered as replies and coloured according to the sex. The replies of the males are depicted in orange and females in grey. Within each cross-correlation the numbers on the top represent the total amount of stack calls over the 8 hours of recording, and the number used to reply to the stack calls of the partner.

Chapter 4

Synchronized recording of position and vocalization helps to understand the function of bird vocalizations

Abstract

The function of a communication signal can be accurately understood only if the global context is interpreted correctly. Vocalizations are often emitted in the presence of conspecifics, whose behaviour can subsequently influence the emission of further vocalizations. In a small songbird, the zebra finch (*Taeniopygia guttata*), we asked whether and how the position of a conspecific (receiver) relative to the focal bird (sender) predicted the types and rate of acoustic signals emitted. We recorded the vocalizations of individual birds using on-board devices, and automatically extracted individual position from video recordings. Both sources of information were synchronized to integrate them with fine temporal resolution. We demonstrate that only one type of vocalization is strictly related to movement, and discuss its potential functions. Our results indicate that the relative position of senders and receivers is important to correctly interpret the vocal repertoire of a species, and opens up the possibility of generating ethograms to clarify the context-dependent usage of vocalizations.

Prepared as: D'Amelio, P. B., Gahr, M., & ter Maat, A.: Synchronized recording of position and vocalization helps to understand the function of bird vocalizations

Introduction

Birds use a wide variety of vocalizations. Precise descriptions of bird species' vocal repertoires are essential for vocal communication studies. The more we investigate vocal networks (Benetos, Stowell, & Plumbley, 2018), neural correlates of vocal communication (Elie & Theunissen, 2015a) and evolution of vocal repertoires (Leighton, 2017), the more important a precise characterization of each vocalization type becomes. The repertoire characterization should not only include the description of the spectral features of vocalizations but also of their usage. To define the entire array of vocalizations and their functions the integration of different kinds of information (e.g. age, sex of the vocalizing individuals, behavioural context) is needed, since communication relies extensively on contextual information (Smith, 1969). Tools to explore vocal repertoires have been recently expanding to include continuous remote recordings (Stepanian et al., 2016), on-board devices (Gill et al., 2016) and unsupervised clustering algorithms (Murray, Mercado, & Roitblat, 1998; Parsons & Jones, 2000). Although an increasing effort is put in the automation of ethograms (Anderson & Perona, 2014; Sakamoto et al., 2009), methods that integrate different kinds of information, e.g. context and spectral features (Elie & Theunissen, 2015b), are still underdeveloped. Despite the increasing interest and technological advances, research on bird vocal repertoires is still based on broad categorizations of behaviours, and the position of the individuals is never considered. A better, yet unexplored approach would be to record positional information in real-time.

Recently, a variety of studies have focused on zebra finch vocal repertoire and its functions (Boucaud, Mariette, Villain, & Vignal, 2016; Elie & Theunissen, 2015b; Gill, Goymann, Ter Maat, & Gahr, 2015). This increased attention has been fuelled by the application of new methods: on-board recording devices and the possibility of analysing spectral features of very large datasets (Elie & Theunissen, 2015b; Gill et al., 2016). These studies have confirmed that calls distinguishable by different spectral shapes are typically elicited in different contexts. However, many questions remain unanswered. For example, the definition of certain calls remains ambiguous. Specifically, two calls with different spectral features, the "stack" and the "tet", were found to be emitted during the same behavioural context (Elie & Theunissen, 2015b). The tet and the stack calls are unlearned affiliative contact calls (D'Amelio, Trost, & ter Maat, 2017; Zann, 1996). The tet call is described as the most common call type in zebra finch groups, where it does not elicit a specific behavioural response from the receiver (Elie & Theunissen, 2015b; Zann, 1996; but see Gill et al., 2015). Tet calls are also important for communication at the nest (Elie et al., 2010). The stack call has been anecdotally described as an infrequent vocalization mainly produced just before taking off and in flight, especially by males while leading females in search for nest sites (Zann, 1996). Zann (1996) also noticed that the stack call is the most commonly emitted call of isolated zebra finches (also in Gill et al., 2016). Subsequent studies found additional important characteristics of this call type: 1) stack calls are used in antiphonal communication (D'Amelio, Trost, et al., 2017; Gill et al., 2015; Ter Maat, Trost, Sagunsky, Seltmann, & Gahr, 2014); and, 2) are individually vocally recognized by both males and females (D'Amelio, Klumb, Adreani, Gahr, & ter Maat, 2017). These two points demonstrate that stack calls can be directed to specific individuals. 3) Stack calls are mostly life-stage-specific (i.e. their frequency of emission changes during different life stages. Specifically it decreases in the breeding context) (Gill et al., 2015). 4) They are sexually dimorphic (Elie & Theunissen, 2015b), 5) their rate and timing of production can be modulated both at short- (Benichov et al., 2016) and long-time scales (D'Amelio et al., 2017). 6) this modulation is probably controlled by the telencephalic areas of the song system (Benichov et al., 2016). Besides the tet and the stack calls, the adult zebra finch repertoire includes a "distance" call used to communicate with mates in absence of visual contact (Blaich et al., 1996); an alarm/alert call called "tuck" (likely named as "hat" too) (D'Amelio et al., 2017; Elie & Theunissen, 2015b); and different types of breeding calls emitted during different life-stages, from nest building to copulation (Elie & Theunissen, 2015b).

In this study, we equipped captive zebra finch couples with on-board microphones in a non-breeding setting, and employed a system for continuous position detection to explore call usage with high temporal resolution. We monitored birds continuously, recording synchronously their position in the aviary and their vocal activity. Specifically, we asked whether the type and/or frequency of calls changed depending on the distance between the birds. Whilst considering all the vocalizations produced, our setup allowed us to focus on one call type emitted predominantly, the stack call. By adding information about the birds' relative distance, we are now in a position to further investigate the use of this call type. We propose that the real-time estimation of relative distance between partners and/or group members is essential to clarify the function of different types of vocalizations.

Material and Methods

Ethics statement

The use of audio transmitters and all other experimental procedures were approved by the government of Upper Bavaria with the record number 55.2-1-54-2532-21-2015 and conducted accordingly. All further animal husbandry and handling were conducted according to the directives 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes.

Animal housing

We recorded the behaviour of 3 established pairs of adult zebra finches (3 males, 3 females) that had successfully reproduced at least once and had not bred for at least a month at the time of the experiment. Each pair was housed separately in the experimental facilities of the Max Planck Institute for Ornithology in Seewiesen on a 13:11 light:dark cycle with 22–26 °C temperature, 40–50% humidity, and *ad libitum* food, grit and water. During the experimental recordings the birds, one pair at a time, were moved to a separate room, without visual or acoustical contact with other birds. The experimental aviary was 1x1x1 m with a transparent ceiling for video recording from the top. The camera (Handykam HK100441, Hayle, UK) was placed at the centre 65 cm over the cage ceiling. The aviary had two perches at the same height. Food, grit and water were provided, but no breeding items were present.

Experimental timeline, video and sound analysis

Vocalizations were recorded with light-weight wireless microphone transmitters fitted on the bird's back via a leg-loop harness (Gill et al., 2016). Birds were fitted with a backpack for a week before being moved to the experimental room where they stayed four days prior the recording, to allow habituation and acclimatization which normally takes 3 days (Gill et al., 2016).

We recorded for two hours in the morning (8:30-10:30), two at midday (12:00-14:00), and two in the afternoon (15:30-17:30) and analysed one hour for each recording of the first and third day after acclimatization. Video and audio were recorded simultaneously by the same device and software and were thus synchronized. After separating video and audio tracks the latter were processed in order to classify call types as previously described (D'Amelio et al., 2017).

From the video track, we extracted one frame per second (Quick time 7, different time resolution can be chosen according to the question) and saved it as an image. In ImageJ (version 1.51p) we used two java macros to remove the background of each image leaving the two birds visible. The second macro determined the coordinates of each bird at all time points. Once the position of the two birds was determined, we computed their distance to each other (i.e. relative position). When the coordinates of the birds corresponded to the beforehand identified coordinates of the perches we identified the bird as sitting on a perch and corrected the distance accordingly. We had to consider that the camera view is conical; therefore we measured one object of known length on the perch and on the ground. The ratio between these measurements was calculated to correct for the distance of birds sitting on a perch or on the ground. To distinguish between "on a perch" and "under a perch" we used a size threshold because objects on perches look bigger being closer to the camera. Finally, when a bird was flying no distance was calculated because we could not estimate the height.

Using the original timestamps we then integrated spatial and vocal information in R environment (R Core Team, 2017).

Time series and statistical analysis

Vocal and positional information were converted to time series (vocal event per second and distance between birds per second). The time of emission of each vocalization was approximated to the nearest integral second. The time series for the positional information were defined using 1 second intervals starting at 0 for the first image. Movement was calculated from distance as the absolute difference in relative position between two consecutive seconds; all the values were normalized to the maximum distance of each recording before being processed. The movement values were assigned to the time bin that precedes the change of position (e.g. at time 1, relative distance was y_1 and movement $z_{1,}$ is the difference of position, $y_2 - y_1$). To calculate the total amount of movement for each session we summed all the absolute movement values.

We extracted the cross-correlation values between distance/movement and call type for each pair, day and part of the day. The reference time series was always the relative position, the second was the vocalization. To avoid the detection of false positives we controlled the autocorrelation of the time series used. Since only one of the two time series, the vocalization, had autocorrelation our results will not be spurious (Dean & Dunsmuir, 2016). To test whether the cross-correlation values of a specific time lag of the different pairs deviated from random we devised a permutation test. For each time series the values were permutated, using blocks of four values to correct for autocorrelation, and a new cross-correlation value calculated. The process was repeated 1000 times and the 95% confidence interval of the time lags ± 5 seconds were extracted and averaged. We present the probability that the real value at time lag 0 of the cross-correlation between distance/movement and vocal production differs from the calculated fitted value.

To test whether the amount of movement and the vocal emission were correlated we fitted linear mixed models using the package “arm”(Gelman & Su, 2015). The couple identity was included as random factor. Residuals were checked visually for normality and absence of patterns. We used the package "MuMI" (Bartoń, 2016) to estimate r^2. The package estimates marginal and conditional r^2-values which correspond to the variance explained by fixed effects alone and total respectively. We report only the marginal r^2-values since we were interested in the fixed part of the model.

Results

Isolated pairs kept in a non-breeding context emitted at least 6 call types besides the male song. However, the majority of the classified calls were stacks (proportions: stack 0.87 ±0.02, tet 0.04 ±0.02, distance 0.04±0.02, hat 0.03± 0.03, breeding calls 0.01±0.01).

The birds spent the most of the time on perches (84.1 % on perches vs. 15.9 % on ground). Periods of intense activity, hopping, were alternated with periods where very little physical activity occurred both with birds separated and in physical contact (figure 1 A, purple line).

For the most common vocalization type, stack calls, a clear pattern emerges: their emission is correlated with hopping whereas when birds did not move, very few vocalizations were uttered, and this was not influenced by whether they were apart or in physical contact (figure 1).

Movement predicts stack emission

We tested whether movement or distance between birds was the best predictor of stack emission rate. We used 4 different time bins (1s, 2s, 5s, and 10s) to determine the temporal window that best describes the use of the stack call (figure 2). The time lags with the highest cross-correlation value between movement and stack emission were almost always within the 1 second interval (1s: 17/18, 2s: 17/18, 5s: 16/18, and 10s: 17/18). There was no tight correlation between Stack emission and the distance between birds.

Instead, the highest cross-correlation values were higher for movement than for distance regardless of bin width (Mean ± SD (1s), Movement: 0.38 ± 0.11; Distance 0.24 ± 0.11). Independently from the time resolution used the time lag with highest cross-correlation value was 0, meaning that movement and vocal emission occurred at the same time.

Usage of stacks in males and females

The use of stack calls was not different between males and females (figure 3). In both males and females, there was no strong relationship between calling and distance for any time lag (mean ccf values ± SD and p-values for time lag 0, males: 0.12±0.12, p=0.17, females: 0.12±0.12, p=0.27). In contrast, for movement there was a significant difference from the baseline in the cross-correlation value when there was no lag between the time series (males: 0.32±0.13, p<0.0001; females: 0.29±0.09, p<0.0001).

Comparison of different call types

Furthermore, we compared whether different call types were used differently (figure 4). Again at time lag 0 stack calls emission had a significant relationship with movement (mean ccf values±SD and p-values, 0.37±0.12, p<0.0001) but not distance (0.14±0.14, p=0.17). In contrast, distance calls at time lag 0 did not show any change in cross-correlation values for either distance or movement. Interestingly, as the names suggest, cross-correlation values were higher for distance than movement (movement: 0.046±0.044, p=0.11; distance: 0.056±0.033, p=0.055). Also, tet calls were not correlated with distance (0.032±0.075, p=0.5) and movement (0.014 ± 0.035, p= 0.444); but, notably, for movement the time lag with the highest cross-correlation value was -1 second (values for time lag -1: 0.021±0.042, p=0.333).

Total amount of movement and call types

Finally, we tested whether higher overall movement correlated with a global higher propensity to vocalize, independently of call type. We found that the amount of movement strongly correlated only with stack call production (r^2-values: stack calls= 0.88, distance calls= 0.40, breeding calls= 0.18, song= 0.12, hat calls= 0.08, tet calls= 0.06; figure 5).

Discussion

Movement, rather than distance between individuals, predicted the stack call emission in male and female zebra finches. No other call type was strictly connected to motion, suggesting a specific function of the stack during movements. Despite having distinct acoustic features, previous studies have lumped two affiliative contact calls, tet and stack, into one category arguing that they were emitted during the same behavioural context (Elie & Theunissen, 2015b). Our results show that context classification is not enough to define the vocal repertoire and that fine, real time, behavioural monitoring is necessary to differentiate between certain call types. We found clear differences in the usage of tet and stack calls suggesting that they indeed have different functions. We did not find differences in the use of stack calls between males and females. Therefore, despite being sexually dimorphic (Elie & Theunissen, 2015b), they probably serve the same function(s). Nomenclature of zebra finch calls has not always been consistent (e.g. see Elie & Theunissen, 2015b; Zann, 1996), but the increased interest about them necessitates a universally accepted labelling of calls. A precise and shared repertoire has implications for multiple fields from the understanding of the evolution of complicated vocal communication (Seyfarth & Cheney, 2014) to mechanisms of perception (Elie & Theunissen, 2015a) and production (Elemans, 2014). Integration between spectral feature, context and real-time evaluation of distance of group components is likely to be necessary to achieve the final repertoire description.

Interestingly, at the smallest time resolution that we employed, 1 second, the highest cross-correlation values between movement and stack emission were the ones with no time lag. We aligned distance with "movement that is going to come"; consequently, no time lag between movement and vocalizations indicates that stack calls were emitted mostly during or following the movement. However, finer time resolution would be necessary to study the causality between movement and vocalisations in detail, because in 1 second birds can perform more than one jump (pers. obs.). We also observed that calls can be emitted before/after take-off and during flight; hence rather than the probability of locomotion they might signal the action itself.

Zebra finches produce basically nothing but stack calls when in total isolation (Gill et al., 2016). Consequently, one hypothesis is that stack calls merely co-occur with any action, hop or flight, independently of the social context. However, we also know that stack calls contain individual identity and can be directed to specific individuals (D'Amelio et al., 2017; Stowell, Gill, & Clayton, 2016; Ter Maat et al., 2014). Combining the notions about antiphonal calling and individual recognition with the presented findings we can infer that the stack call might signal locomotion to a specific individual. Zebra finches are life-long monogamous and a pair moves together throughout the year (Zann, 1994). The maintenance of such a stable pair bond is linked to close spatial contact that can be facilitated by vocal communication. Interestingly, besides signalling locomotion, the continuous chattering of zebra finch pairs might have developed relevance for other behaviours such as pair maintenance and commitment to the partner (D'Amelio et al., 2017; Hernandez, Perez, Mulard, Mathevon, & Vignal, 2016). During movement the birds in our study hopped from one perch to another or to the cage's mesh, often with exaggerated jumps, resembling a behavioural display (pers. obs.). However, the meaning of the observed hopping in the wild is yet to be discovered. From qualitative and anecdotal observations in the wild, it was observed that stack calls are emitted during nest search and take off (Zann, 1996). Therefore, stack call emission during hopping could be merely a by-product of the caged environment, a stereotypical behaviour (Mason, 1991), yet other hypotheses should be considered. For example, stack calls might also act as “moving calls”, aiding the quorum decision in moving groups (Bousquet, Sumpter, & Manser, 2011). Alternatively, they could be part of a sort of ritual dance since zebra finch (Ullrich, Norton, & Scharff, 2016) and other species in the taxonomic group, Estrildidae, perform ritualized movements during courtship (Langmore & Bennett, 1999; Ota, Gahr, & Soma, 2015). Interestingly both stack and song number were correlated with the total amount of movement; hence possibly also the stack call exchange is part of pair maintenance and reciprocal stimulation behaviour (Bolund, Schielzeth, & Forstmeier, 2012; D’Amelio et al., 2017).

Albeit relatively rarely, the birds emitted other call types beside stack calls in our experiment; however, more active pairs did not simply produce more calls of any kind. Nevertheless, none of the call types recorded were correlated to specific behaviours, possibly because of our impoverished environment (Elie & Theunissen, 2015b). Distance and tet calls were not strictly correlated to either distance or movement (although tets had a weak tendency to be emitted before movements). To understand the precise rules of emission of specific calls, we need to integrate more information, and/or to design specific experimental arenas. For example, including information about movement related to feeding would help identifying types of calls or call rates referred to food (Clay, Smith, & Blumstein, 2012); placing nest material and nest boxes can elicit breeding calls (Zann, 1975), and so on. To explore the function of calls used at the group level it would be necessary to add and follow multiple individuals at once. The approach described here can help to produce ethograms to accurately describe the proximate cause of vocal emission.

Detailed classification of behaviour, aided by learning algorithms and real-time tracking brought important breakthroughs in the study of behaviour (Anderson & Perona, 2014; Berman, Bialek, & Shaevitz, 2016; Sakamoto et al., 2009). However studies tracking birds' relative position are still limited (Dell et al., 2014) and none, to our knowledge, integrated vocalizations with positional information. Besides aiding a finer description of call functions a continuous unsupervised monitor of position can have applications in different fields. For example, the synchronicity of behaviours such as offspring provisioning at the nest has been found to have fitness consequences (Mariette & Griffith, 2015). Thus, automating behavioural scoring using our method could enhance quantification and ease data collection. Or, in conjunction with neural recordings (Schregardus et al., 2006; Ter Maat et al., 2014), the here described system can be used to study the neural correlates of positional information and vocalizations.

Using the approach outlined here, recording bigger groups and adding different contexts will certainly lead to a precise understanding of the call usage of zebra finches and the same methods can be applied to other species. Overall, the methods used in this study can lead to a novel strategy for drawing ethograms, a more objective and quantitative approach to measure behaviour and its global context.

Acknowledgements

We are grateful to Lisa Trost for help conducting the experiments, advice and to her and Hannes Sagunsky for technical support. We also thank Michelle Land and Gabrielle Zuniga for help during the method development. PBD thanks the IMPRS research school for training and support. This work was funded by the Max-Planck-Gesellschaft.

References

Anderson, D. J., & Perona, P. (2014). Toward a Science of Computational Ethology. *Neuron*, *84*(1), 18–31. https://doi.org/10.1016/J.NEURON.2014.09.005

Bartoń, K. (2016). MuMIn: multi-model inference. *R Package*, *version 1*.

Benetos, E., Stowell, D., & Plumbley, M. D. (2018). Approaches to Complex Sound Scene Analysis. In *Computational Analysis of Sound Scenes and Events* (pp. 215–242). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-63450-0_8

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, *26*(3), 309–318. https://doi.org/10.1016/j.cub.2015.12.037

Berman, G. J., Bialek, W., & Shaevitz, J. W. (2016). Predictability and hierarchy in Drosophila behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(42), 11943–11948. https://doi.org/10.1073/pnas.1607601113

Blaich, C. F., Norman, M., Syud, F. A., Benitez, G., Frost, J., Ravenscroft, J., … Ware, P. (1996). The use of distance calls to maintain pair contact in zebra finches (*Taeniopygia guttata*). *Bird Behavior*, *11*(1), 25–30. https://doi.org/10.3727/015613896791748889

Bolund, E., Schielzeth, H., & Forstmeier, W. (2012). Singing activity stimulates partner reproductive investment rather than increasing paternity success in zebra finches. *Behavioral Ecology and Sociobiology*, *66*(6), 975–984. https://doi.org/10.1007/s00265-012-1346-z

Boucaud, I. C. A., Mariette, M. M., Villain, A. S., & Vignal, C. (2016). Vocal negotiation over parental care? Acoustic communication at the nest predicts partners' incubation share. *Biological Journal of the Linnean Society*, *117*(2), 322–336. https://doi.org/10.1111/bij.12705

Bousquet, C. A. H., Sumpter, D. J. T., & Manser, M. B. (2011). Moving calls: a vocal mechanism underlying quorum decisions in cohesive groups. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1711), 1482–1488. https://doi.org/10.1098/rspb.2010.1739

Clay, Z., Smith, C. L., & Blumstein, D. T. (2012). Food-associated vocalizations in mammals and birds: what do these calls really mean? *Animal Behaviour*, *83*(2), 323–330. https://doi.org/10.1016/J.ANBEHAV.2011.12.008

D'Amelio, P. B., Klumb, M., Adreani, M. N., Gahr, M. L., & ter Maat, A. (2017). Individual recognition of opposite sex vocalizations in the zebra finch. *Scientific Reports*, *7*(1), 5579.

https://doi.org/10.1038/s41598-017-05982-x

D'Amelio, P. B., Trost, L., & ter Maat, A. (2017). Vocal exchanges during pair formation and maintenance in the zebra finch (Taeniopygia guttata). *Frontiers in Zoology*, *14*(1), 13. https://doi.org/10.1186/s12983-017-0197-x

Dean, R. T., & Dunsmuir, W. T. M. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, *48*(2), 783–802. https://doi.org/10.3758/s13428-015-0611-2

Dell, A. I., Bender, J. A., Branson, K., Couzin, I. D., de Polavieja, G. G., Noldus, L. P. J. J., … Brose, U. (2014). Automated image-based tracking and its application in ecology. *Trends in Ecology and Evolution*, *29*(7), 417–428. https://doi.org/10.1016/j.tree.2014.05.004

Elemans, C. P. (2014). The singer and the song: The neuromechanics of avian sound production. *Current Opinion in Neurobiology*, *28*, 172–178. https://doi.org/10.1016/J.CONB.2014.07.022

Elie, J. E., Mariette, M. M., Soula, H. A., Griffith, S. C., Mathevon, N., & Vignal, C. (2010). Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal Behaviour*, *80*(4), 597–605. https://doi.org/10.1016/j.anbehav.2010.06.003

Elie, J. E., & Theunissen, F. E. (2015a). Meaning in the avian auditory cortex: neural representation of communication calls. *The European Journal of Neuroscience*, *41*(5), 546–67. https://doi.org/10.1111/ejn.12812

Elie, J. E., & Theunissen, F. E. (2015b). The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition*, *19*(2), 285–315. https://doi.org/10.1007/s10071-015-0933-6

Gelman, A., & Su, Y.-S. (2015). Data analysis using regression and multilevel/hierarchical models {arm}.

R Package. Retrieved from https://cran.r-project.org/package=arm

Gill, L. F., D'Amelio, P. B., Adreani, N. M., Sagunsky, H., Gahr, M. C., & ter Maat, A. (2016). A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods in Ecology and Evolution*, *7*(11), 1349–1358. https://doi.org/10.1111/2041-210X.12610

Gill, L. F., Goymann, W., Ter Maat, A., & Gahr, M. (2015). Patterns of call communication between group-housed zebra finches change during the breeding cycle. *eLife*, *4*, e07770. https://doi.org/10.7554/eLife.07770

Hernandez, A. M., Perez, E. C., Mulard, H., Mathevon, N., & Vignal, C. (2016). Mate call as reward: Acoustic communication signals can acquire positive reinforcing values during adulthood in female zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology*, *130*(1), 36–43. https://doi.org/10.1037/a0040027

Langmore, N. E., & Bennett, A. T. D. (1999). Strategic concealment of sexual identity in an estrilid finch. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1419), 543–550. https://doi.org/10.1098/rspb.1999.0670

Leighton, G. M. (2017). Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings*. *Biological Sciences*, *284*(1868), 20171508. https://doi.org/10.1098/RSPB.2017.1508

Mariette, M. M., & Griffith, S. C. (2015). The adaptive significance of provisioning and foraging coordination between breeding partners. *The American Naturalist*, *185*(2), 270–80. https://doi.org/10.1086/679441

Mason, G. J. (1991, June 1). Stereotypies: a critical review. *Animal Behaviour*. Academic Press. https://doi.org/10.1016/S0003-3472(05)80640-2

Murray, S. O., Mercado, E., & Roitblat, H. L. (1998). The neural network classification of false killer whale (Pseudorca crassidens) vocalizations. *The Journal of the Acoustical Society of America*, *104*(6), 3626–3633. https://doi.org/10.1121/1.423945

Ota, N., Gahr, M., & Soma, M. (2015). Tap dancing birds: The multimodal mutual courtship display of males and females in a socially monogamous songbird. *Scientific Reports*, *5*, 16614. https://doi.org/10.1038/srep16614

Parsons, S., & Jones, G. (2000). Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *Journal of Experimental Biology*, *203*(17), 2641–2656.

R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F., & Wanless, S. (2009). Can Ethograms Be Automatically Generated Using Body Acceleration Data from Free-Ranging Birds? *PLoS ONE*, *4*(4), e5379. https://doi.org/10.1371/journal.pone.0005379

Schregardus, D. S., Pieneman, A. W., Ter Maat, A., Jansen, R. F., Brouwer, T. J. F., & Gahr, M. L. (2006). A lightweight telemetry system for recording neuronal activity in freely behaving small animals. *Journal of Neuroscience Methods*, *155*(1), 62–71. https://doi.org/10.1016/J.JNEUMETH.2005.12.028

Seyfarth, R. M., & Cheney, D. L. (2014). The evolution of language from social cognition. *Current Opinion in Neurobiology*, *28*, 5–9. https://doi.org/10.1016/J.CONB.2014.04.003

Smith, W. J. (1969). Messages of Vertebrate Communication. *Science*, *165*(3889), 145–150. Retrieved from http://www.jstor.org/stable/1726424

Stepanian, P. M., Horton, K. G., Hille, D. C., Wainwright, C. E., Chilson, P. B., & Kelly, J. F. (2016).

Extending bioacoustic monitoring of birds aloft through flight call localization with a three-dimensional microphone array. *Ecology and Evolution*, *6*(19), 7039–7046. https://doi.org/10.1002/ece3.2447

Stowell, D., Gill, L., & Clayton, D. (2016). Detailed temporal structure of communication networks in groups of songbirds. *Journal of the Royal Society, Interface*, *13*(119), 20160296. https://doi.org/10.1098/rsif.2016.0296

Ter Maat, A., Trost, L., Sagunsky, H., Seltmann, S., & Gahr, M. (2014). Zebra finch mates use their forebrain song system in unlearned call communication. *PloS One*, *9*(10), e109334. https://doi.org/10.1371/journal.pone.0109334

Ullrich, R., Norton, P., & Scharff, C. (2016). Waltzing Taeniopygia: integration of courtship song and dance in the domesticated Australian zebra finch. *Animal Behaviour*, *112*, 285–300. https://doi.org/10.1016/J.ANBEHAV.2015.11.012

Zann, R. A. (1975). Inter- and Intraspecific Variation in the Calls of Three Species of Grassfinches of the Subgenus Poephila (Gould) (Estrildidae). *Zeitschrift Für Tierpsychologie*, *39*(1–5), 85–125. https://doi.org/10.1111/j.1439-0310.1975.tb00902.x

Zann, R. A. (1994). Reproduction in a Zebra Finch Colony in South-eastern Australia: the Significance of Monogamy, Precocial Breeding and Multiple Broods in a Highly Mobile Species. *Emu*, *94*(4), 285–299. Retrieved from https://doi.org/10.1071/MU9940285

Zann, R. A. (1996). *The zebra finch: a synthesis of field and laboratory studies*. *Oxford Univ, Press*. Retrieved from http://books.google.com/books?id=5KO6cZH0WbEC&pgis=1

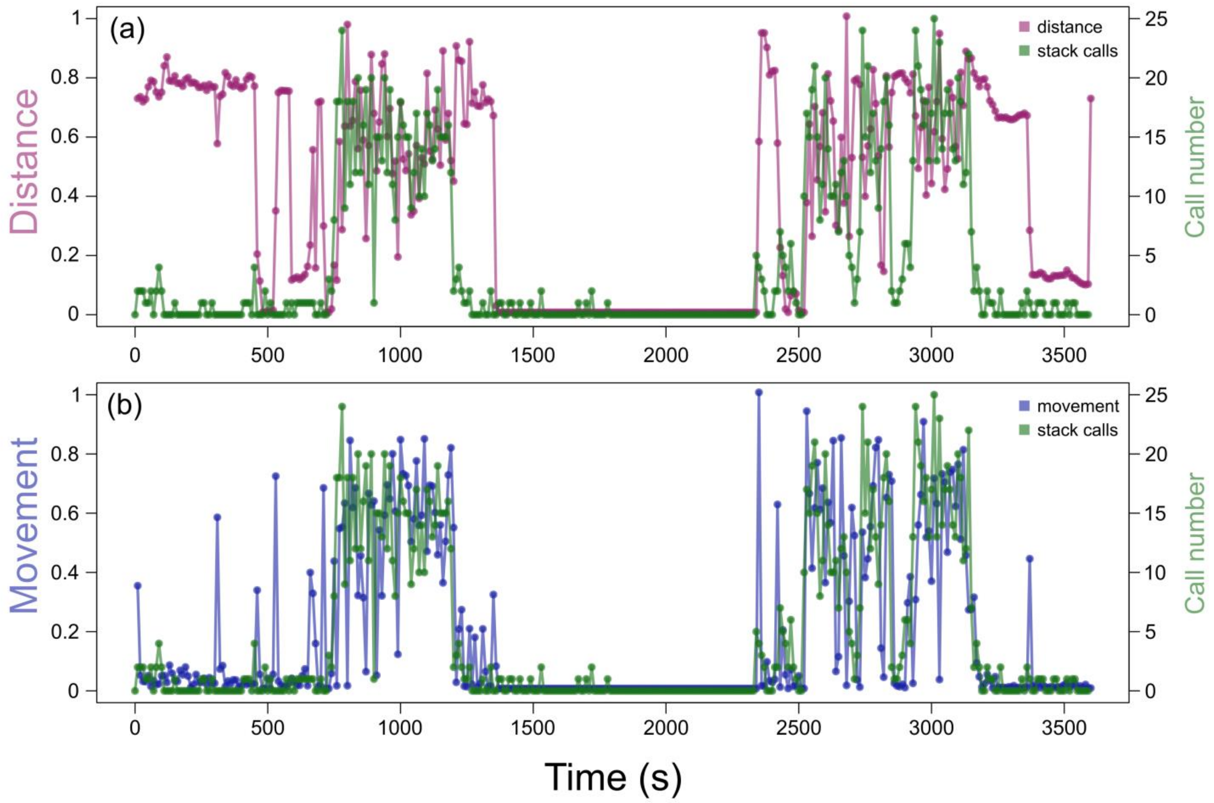


Figure 1 Relative position and calling over time

Example of the raw data binned every 10 seconds of the birds' relative position over time (1 hour). On the y-axis are plotted distance (A-purple) or movement (B-blue) and number of stacks (green) for each bin normalized by the maximum value reached in the recording. The three variables (call number, movement, distance) were used as time series data for the cross-correlation analysis.

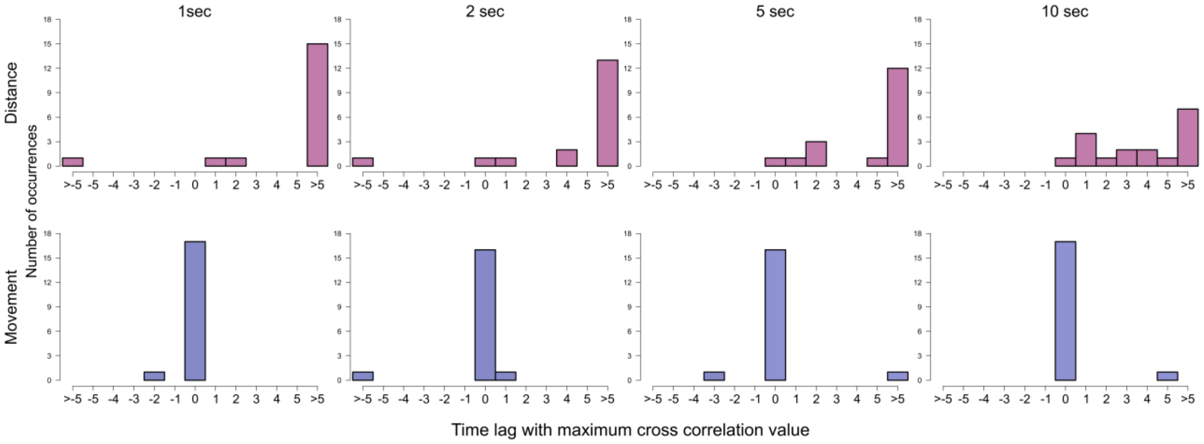


Figure 2 Maximum value of cross-correlation for different time lags

Distribution over time lags of the maximum cross-correlation values between stack calls and distance (purple)/movement (blue) at different bin widths. The cross-correlation values of vocalizations with distance are on top and with movement at the bottom. For any bin widths considered high cross-correlation values are concentrated around time lag 0 when correlating vocalizations and movement, whereas the high values are disperse when correlating vocalizations and distance.

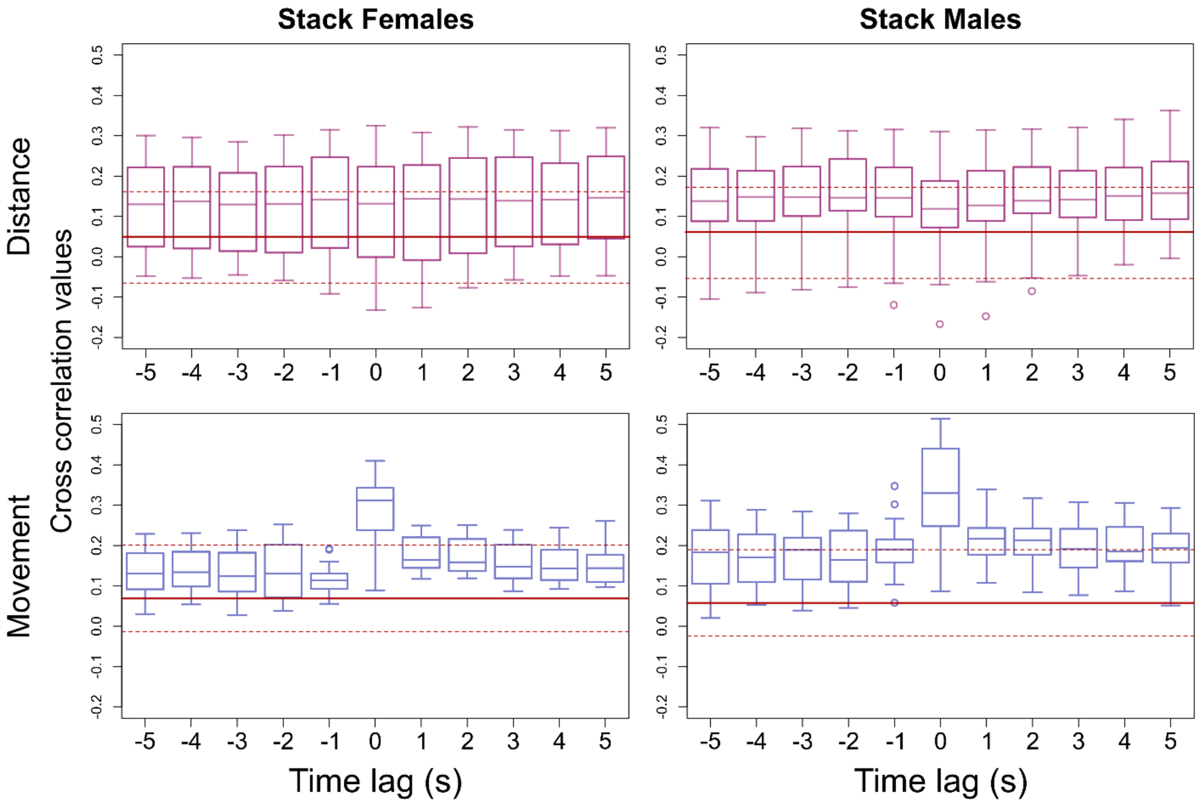


Figure 3 Differences between male and females calls

Cross-correlations between distance (top, purple) / movement (bottom, blue) and male (right) and female (left) stack calls for all the considered pairs for ±5 s time lags. Dotted red lines represent the estimated 95% confidence intervals and the solid line the fitted values. At time lag 0 the cross-correlation between movement and vocal emission has very high values. The Stack calls emission is associated with movement rather than distance and males and females have very similar patterns.

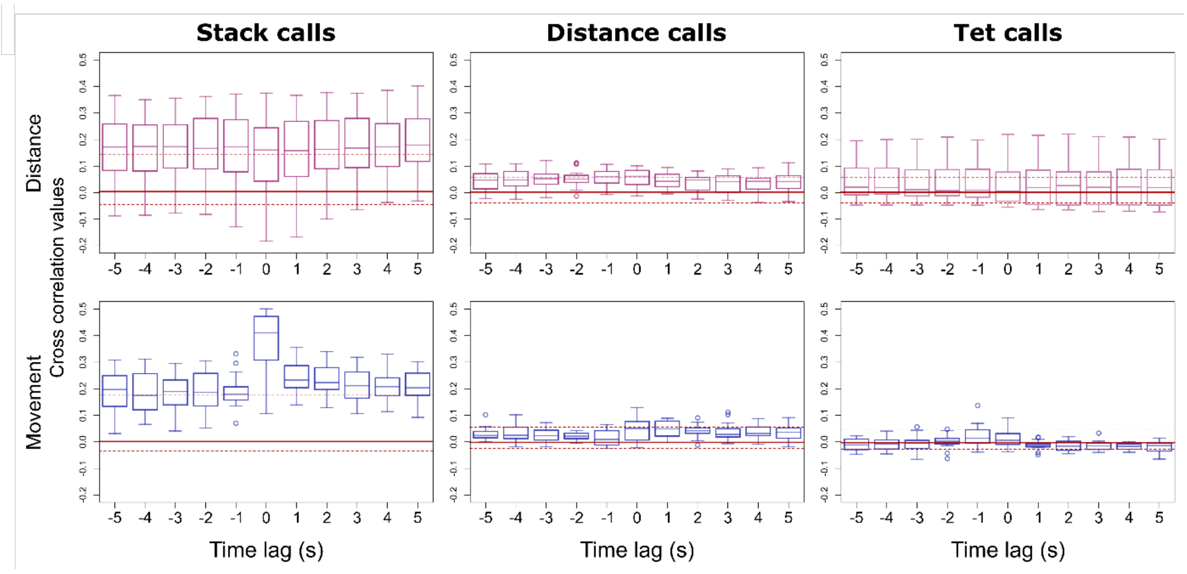


Figure 4 Cross-correlation between distance (top) / movement (bottom) and different call types

Three different affiliative contact calls are shown Stack, Distance and Tet call. On the top (purple) is shown the correlation between distance and the vocal emission. On the bottom (blue) is shown the correlation between movement and the vocal emission. For all the considered pairs the ±5 s time lags are shown. Dotted red lines represent the estimated 95% confidence intervals and the solid line the fitted values. The only strong correlation is between movement and stack call when there is no time lag between the two.

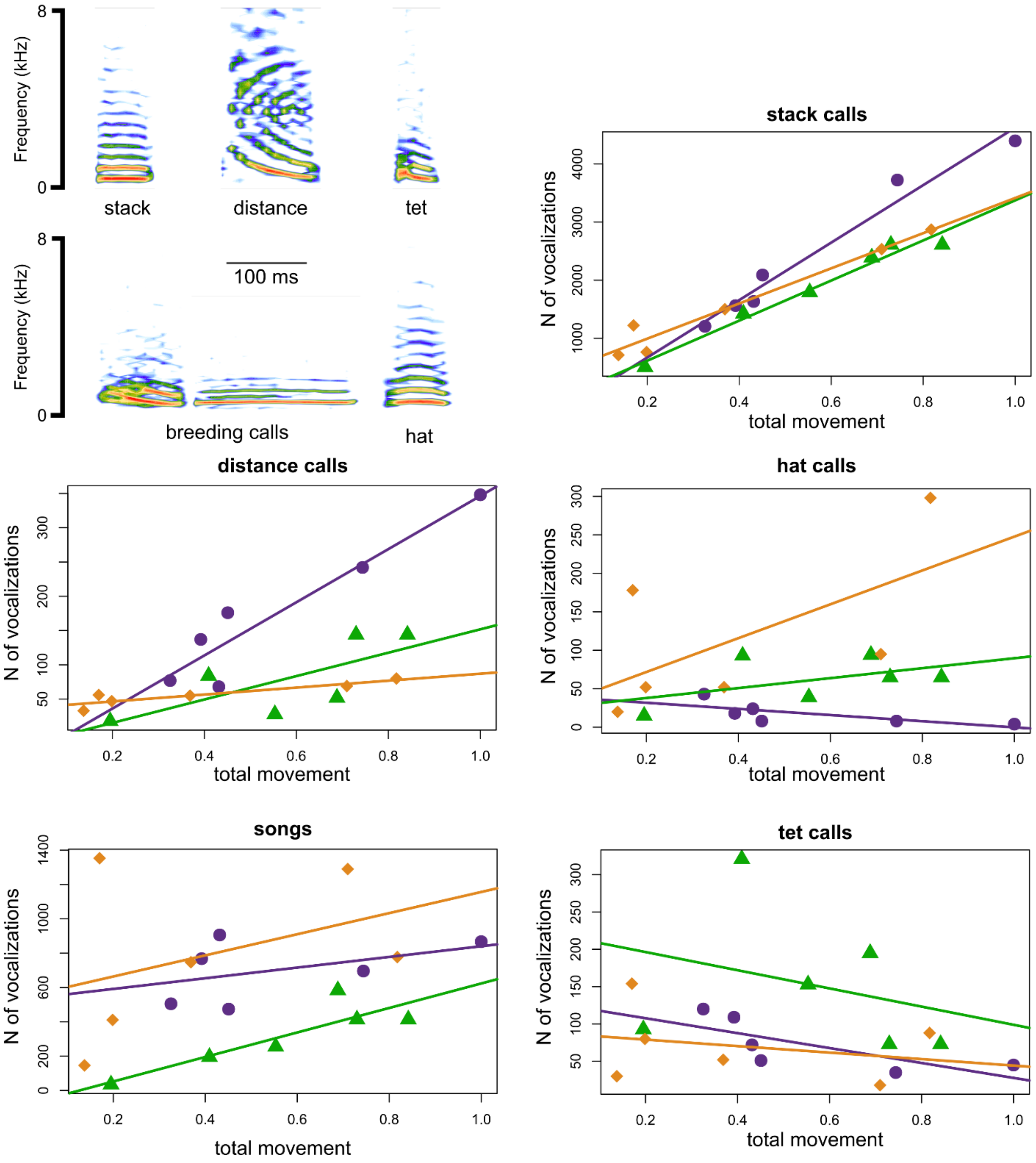


Figure 5 Correlation between total amount of movement and vocal emission

Top left, representative spectrograms of the vocalization types considered. Rest of the panels, each box contains the relationship for different vocalization types between vocal emission (y-axes) and total amount of movement (x-axes). Each colour/symbol represents a couple. Each data point corresponds to one hour of behavioural recording. The amount of movement is not simply correlated with the vocal emission of any call type but only with stack calls.

General discussion

In my thesis, I studied the vocal exchanges of zebra finches as a model for vocal communication of social living species. Instead of focusing on a single vocalization type I decided to consider the entire species' repertoire. For example, I regarded songs, the most studied auditory communication signal in birds, just as one of the several types of vocalization. To quantify the vocal production of each individual with high precision I first contributed to the development and testing of a tool, the backpack microphone (**chapter 1**). Using this device I was able to describe the vocal interactions between individuals, focusing on the most important social unit for this species: the pair (**chapter 2, 3, 4**). I first studied whether and how zebra finches recognize and reply to specific individuals within a group; I found that they are capable of individual vocal recognition of short innate calls (**chapter 2**). Then, I described the vocal relationship within bonded pairs and its development during pair formation; I found that turn-taking vocal behavior is stable in established pairs and needs to develop in newly formed ones (**chapter 3**). Finally, I integrated the relative distance between pair members with their vocal activity and discovered that one of the zebra finch affiliative contact calls is connected to movement, which helped us to characterize the vocal repertoire of the species (**chapter 4**). Overall the present thesis takes an important step towards the comprehension of the usage of learned and unlearned vocalization in songbirds. I am convinced that the outlined findings will represent: i) the basis for studying the fine mechanisms of the described behaviors at the neural level; ii) the foundation for a multispecies comparative analysis of call based communication and, iii) once the relevance of calling interactions is established in the wild, the starting point to unravel the impact of calling dynamics on fitness.

New tools for old questions: the importance and challenge of quantifying individual vocal behavior

Knowing who says what to whom is fundamental for many questions in vocal communication (Gill, Goymann, Ter Maat, & Gahr, 2015; McGregor, 2005). Recording the vocal output of individuals without influence on the studied behavior is of substantial importance (Caccamise & Hedin, 1985). Doing it in small flying animals with short and rather similar inter-individual vocalizations is, however, not an easy task. For these animals, individualized recording has been achieved by either separating individuals (Fernandez, Vignal, & Soula, 2017), or triangulating the auditory signals after recording with microphone arrays (Chiu, Xian, & Moss, 2009; Suzuki et al., 2018), or using on-board devices (**Chapter 1;** Anisimov et al., 2014; Hiryu, Shiori, Hosokawa, Riquimaroux, & Watanabe, 2008). Miniaturization of on-board microphones has been difficult for different reasons. Compared to most data collected via remote sensing (e.g. position), acoustic recordings are very demanding in terms of sampling rate (44100 Hz in all recordings of the present thesis) resulting in high energy consumption and storage requirements. In **chapter 1** we describe how we overcame these challenges.

We demonstrated that our microphone-transmitters enabled us to obtain individualized and synchronized recordings with a recording duration (up to 16 days) that exceeded the habituation period (up to 3 days), resulting in unbiased behavioral quantification. The microphone-transmitters carried as backpacks, required minimal handling of the birds and in addition, the battery exchange after ~14 days biased only the day of manipulation with vocal and locomotor activity recovering to baseline within 24 hours. The quality of the recordings remained very high regardless of the environmental and social noisiness. The device also allowed successful copulation and egg laying. Since it affects bird behavior only temporarily and enables normal physical and vocal interactions, we achieved high flexibility in the study design:

multiple individuals, noisy conditions, developmental studies and complex aviary design, they all do not hamper the recording quality.

Depending on the specific question and model species the tool can be customized to suit different circumstances. Larger animals, for example, could carry heavier devices equipped with bigger batteries resulting in longer recording periods and/or larger recording ranges. Larger animals could also carry the usually heavier loggers (see introduction **chapter 1** for a review comparison loggers vs. transmitter); however, loggers cannot be synchronized precisely hence they are not suitable for vocal interaction related questions. The current transmitting range of the microphone-backpacks (up to 26m) allows recordings even in the wild under certain conditions. For example, so far, to record habituated groups of animals researchers had to follow them with a microphone (Bousquet, Sumpter, & Manser, 2011; Golabek & Radford, 2013); now each individual of the group could carry a transmitter and the researcher can follow them with an antenna. This would allow recording synchronously each animal in the group. Moreover, it would be possible to place the antenna in strategic positions where sender carrying birds assemble, such as roosts or feeding places. With individual-level resolution and synchronized recordings, we can now study duets, identify dialogues in vocalizing groups or eavesdropping on the private whispers at a night roost in the wild.

However, the data recorded with our microphone transmitters also pose new challenges. The amount of data is simply massive: many thousands if not millions of vocalizations had to be handled for each study presented here (**chapter 2, 3, 4**). The data analysis necessitates new techniques, due to the fact that cross-correlation analysis is not sufficient anymore when more than two birds are involved (Stowell, Gill, & Clayton, 2016). Moreover, the analysis became complicated and data hungry when we compared/correlated multiple types of vocalizations. Specific analytical methods will need to be developed to understand the interplay between different call combinations and individuals in vocal social network.

In summary, the backpack-microphones transmitter we described in **chapter 1** is currently the best tool for recording individualized vocal activity in small animals and was, therefore, the basic tool used in each thesis chapter. I was able to test the exact influence that it has on the studied behavior. Therefore, it allowed me to precisely quantify vocal behavior in relevant social settings and it was particularly useful for unlearned calls of songbirds, which are often soft, short and spectrally very similar between different individuals (Marler, 2004; Reichard & Anderson, 2015; Zann, 1996). On-board microphones will be the future of vocal communication studies. I could demonstrate that this system is suitable for captive conditions and that is ready to be tested in the wild.

Individuality in zebra finch unlearned calls and the importance of timing

Auditory learning leading to individual vocal recognition is ubiquitous among vertebrates (Jarvis, 2006). It has been proven in zebra finches when tested with the male song (Miller, 1979) and both sexes' distance calls (Vignal, Mathevon, & Mottin, 2004, 2008). Besides the song and the distance calls the zebra finch repertoire contains short vocalizations, which, as mentioned above are: unlearned, very similar between individuals and often very variable within a type (**chapter 2,** Elie & Theunissen, 2015b; Zann, 1996). No study so far has investigated whether these vocalizations can be assigned to individuals by the fellow birds. Recording vocalizing groups with synchronized microphone–transmitters provided evidence that zebra finches respond to specific individuals (Gill et al., 2015), which requires individual recognition. Individual recognition can be achieved via different channels (e.g. visual, olfactory, acoustic). We now provide evidence that the acoustic channel can be used efficiently to identify individuals (**chapter 2**). Employing miniaturized microphone-transmitters (**chapter 1**), we demonstrated for the first time that unlearned soft vocalizations contain markers of individual identity. Moreover, our result also provides a mechanism by which animals can address specific individuals in a group: the latency of the answer. Compared to the past notion, we mark a change of paradigm: we now show that what was thought to be an undirected and unspecific vocalization (Zann, 1996) can convey messages to specific individuals. We also found that timing and not the number of calls identifies who the bird is directing the response to, stressing the importance of timing in zebra finch vocal networks. The same call can have two opposite meanings ("I ignore you" vs. "I respond to you") just by changing the latency to the sender call by ~150 milliseconds. As a consequence of this finding we now suggest that call latency, rather than the number of calls or type, should be strictly controlled when establishing communication networks or when using call based response as a proxy for preference (Chen, Clark, & Woolley, 2017).

In parallel with the main finding of **chapter 2** we also tested hypotheses never considered in acoustic learning studies so far:

i) Do vocal relationships prior to the experiment influence the performance during the trial?

ii) Does the type of call used as answer differ according to the familiarity of the playback?

We know that the calling relationship in established pairs is stable over time but differs from pair to pair (**chapter 3**). Hence we can expect that a stronger calling association prior to the experiment will lead to stronger answering during the playback trial. Interestingly, the relationship prior to the experiment seemed to influence the answering during playback in a direction opposite to the expectations: males that answered with a higher proportion of calls during the baseline responded with the lowest proportion of calls during the playback. It is possible that this finding is just a false positive due to the small sample size but there are alternative explanations. For example, pairs with a higher answering rate might have stronger bonds (**chapter 3**) and produce more corticosterone, experience more stress, when divided for the playback experiment which in turn impairs hearing ability (Perez et al., 2012).

In addition, we found that the type of call used to answer did not change depending on the familiarity of the playback. This was possibly a consequence of the testing context, in that it was appropriate only for certain call types (**chapter 4**, Elie & Theunissen, 2015b), and caused the bird to use mainly *stack* and *distance* calls.

Behavioural habituation in playback experiments can influence the resulting outcome (Kroodsma, 1986). Despite the long trials, we found only limited signs of habituation. Our strategy was to use several different exemplars of each type and familiarity (this also helped to avoid pseudoreplication; Kroodsma, Byers, Goodale, Johnson, & Liu, 2001). Also, we randomized the gap between playback calls so as not to create expectations of the listener. As a consequence, although we processed the largest amount of data analyzed to date in acoustic recognition experiments, signs of habituation were only found when playing back males' songs.

Some questions regarding vocal recognition by zebra finches remain still open. For example, in our experiment evidence for recognition of group members was limited. Therefore, how many different individuals can be distinguished and memorized remains unclear (Wiley, 2013). Another open question is whether the voice (i.e. an individual specific combination of relations between spectral features) itself is recognized or whether there is a specific identity label of each call type. A third open question is how phylogenetically common recognition of soft unlearned contact calls is, which is why comparative studies are necessary. We think that the novelty of our study, the expedients used to avoid habituation together with the strict data analysis approach, will influence how future playback studies will be designed and interpreted.

Consistency and flexibility of calling interactions

Through recording new and established pairs of zebra finches non-stop for a week and analyzing all their vocal output, we found that paired birds answer each other mainly using an affiliative contact call, the stack. In established pairs, the pattern of antiphonal calling was consistent over time. Interestingly, new pairs developed a vocal relationship, consistent antiphonal calling, in parallel with the increase of other affiliative social behaviors. In fact, incipient pair members spent more and more time in physical contact which agrees with previous studies (Silcox & Evans, 1982). Meanwhile, we observed post pairing behavioral convergence in vocal behavior, in which the number of answers (calls following within half a second the call of a partner, see Ter Maat, Trost, Sagunsky, Seltmann, & Gahr, 2014) became similar between partners. The degree and speed of post pairing convergence can have fitness consequences or be a proxy for it (Laubu, Dechaume-Moncharmont, Motreuil, & Schweitzer, 2016). Whether this is the case for the turn taking vocalizations remains an open question.

We showed that the turn-taking exchange can be flexible, but also extremely repeatable. Established pairs in the same context showed highly predictable vocal exchanges: they used always the same proportion of calls to answer the partner day after day and also within the day. Such a high degree of stereotypy is rarely encountered in vertebrate behavior (Bell, Hankison, & Laskowski, 2009). The continuous production of stacks might be the simplest discrete signal, an implicit binary one (Hailman, 2008), to declare the permanent presence of the sender. On the other hand additional, not mutually exclusive, functions can be hypothesized. For example, repeatable stack production might show a bird's commitment to stay with their partner. Also, the antiphonal exchange within a pair might signal their relationship to other group members. In addition, the matching of the number of answered calls within a pair can have a meaning similar to the matching of spectral features shown by some parrot species: address an individual and show interest (Balsby & Bradbury, 2009; Hile, Plummer, & Striedter, 2000).

Turn-taking vocal behavior is present in several animals, humans included (Levinson, 2015). Its study in non-human primates has increased in recent years (Flack, 2013) because it is one of the language principles (Levinson, 2015). However, no study so far has considered comparisons with birds. To aid the understanding of the principle of turn-taking I advocate for comparative studies and further studies on birds, especially to understand whether and how it is learned.

If the stack exchange is more meaningful than just signaling the presence/movement of an individual, indexes derived from quantification of the turn-taking calling behavior will become a measure for pair compatibility. Pair compatibility, defined as the fitness benefits that arise from the union of two specific partners, was demonstrated to be important for mate choice and fitness in different systems (Schuett, Tregenza, & Dall, 2010), including zebra finches (Ihle, Kempenaers, & Forstmeier, 2015; Schuett, Dall, & Royle, 2011). In our experiment we measured only forced pairs, hence we expect their degree of pair compatibility to vary appreciably. Interestingly, we found some support for the hypothesis that vocal exchange can be an index of pair compatibility; the proportion of answers to the partner and the amount of time that birds spent in physical contact were positively correlated, suggesting a relationship between affiliative social and vocal behavior. Further study of the vocal interactions could thus contribute to understanding the mechanisms of pairing, monogamy and pair compatibility (Adkins-Regan & Tomaszycki, 2007; Goodson, Kabelik, Kelly, Rinaldi, & Klatt, 2009; Tomaszycki & Atchley, 2017). In this context it is also important that further studies clarify which partner influences the other in order to achieve the extremely repeatable and stereotyped calling pattern of established pairs.

Using an innovative method, the individual vocal recording, we possibly bridged the gap between the fields of communication and sexual selection by showing consistent associations between vocal exchanges and pair formation and maintenance (**chapter 2**).

Partners' relative spatial position clarifies zebra finch repertoire

The commendable quantitative description of the zebra finch repertoire by Elie and Theunissen (2015b) still contained some ambiguity in the definition of two affiliative call types, the *stack* and the *tet* calls. Despite different spectral features, *stack* and *tet* were found to be emitted in the same context and therefore lumped into one category. Integrating the synchronous recordings of backpack microphones (**chapter 1**) with knowledge about the zebra finch repertoire (**chapter 2** and **3**) and the continuous tracking of bird spatial position (**chapter 4**) we found that only one call type, the *stack*, is strictly connected with movement. This suggests a specific function of the *stack* connected with locomotion, whereas *tet* calls, despite emitted in the same context, must have a different function. I recorded the smallest social unit, the pair, but the study of a group of animals is necessary to generalize and extend my conclusions.

In a constant environment, *stack* exchange is flexible during pair formation and fixed in established pairs (**chapter 3**). In a group of zebra finches, the stack exchange might signal movement (and/or commitment, see above) to the partner and potentially also to the other group members. The fixed rate of answers characteristic of each individual (**chapter 3,** Ma, Maat, & Gahr, 2017) can create an expectation in the receiver brain that helps to keep contact without paying attention to the signal. A change of call type can create sudden arousal (Beckers & Gahr, 2012). While our description provides support for this scenario we also noted that many of the movements were overemphasized, birds did not simply hop from one perch to another but often flew high arcs over the perches and chose indirect trajectories. Possibly, the purpose of the hops combined with the calling activity is not only locomotion; they might be a stereotypical behavior or part of a couple display, a pair maintenance dance. Auditory-visual displays, common in the animal kingdom, increase the dimensions of the signals on which selection (natural and sexual) acts (Candolin, 2003). Combination of auditory and other signals is phylogenetically widespread in birds (Cooper & Goller, 2004; Dalziell et al., 2013), and, although quite concealed, is present in many

Estridild finches where their influence on mate choice is yet to be evaluated (Langmore & Bennett, 1999; Ota, Gahr, & Soma, 2015; Ullrich, Norton, & Scharff, 2016).

Altogether, to finally understand the function of the repeatable antiphonal stack calling, we need studies on wild populations, as well as comparative research within the Estrildid finches (family Estrildidae) investigating how the connection between hopping and calling behavior is distributed phylogenetically.

The study of calls in birds: status, and prospects

Although being almost neglected for a long time (Marler, 2004), the study of bird calls is on the rise in the last few years. A lot of effort has been recently devoted to unraveling different aspects of call production (Benichov et al., 2016; Ter Maat et al., 2014), perception (Elie & Theunissen, 2015a; Theunissen & Elie, 2014) and usage (**chapter 2, 3,** Engesser, Ridley, & Townsend, 2016; Gill et al., 2015; Ma et al., 2017; Suzuki, Wheatcroft, & Griesser, 2016). Several factors have contributed to this increased attention. In part it is due to improved methodologies to record vocalizations individually, process them and handle large datasets. Also it may be due to their involvement in interesting and cognitively demanding behaviors. For example, there is more and more evidence that the real “language” of birds uses the calls and not songs. In fact, only in call associations we find the combinatory syntax: the position of the vocalization within a combination is important for the meaning (Engesser et al., 2016; Griesser, Wheatcroft, & Suzuki, 2018). Moreover, mainly calls and not songs mediate social relationships and form interactive vocal networks (Gill et al., 2015). Despite the increasing interest in calls and their production, there are still many gaps in our knowledge. Is the combinatory syntax found in the distantly related Pied Babbler and Japanese tit a common feature for songbirds (Griesser et al., 2018)? Do all the populations of babbler and tit species develop the use of syntax in their calls or did it only evolve under a specific condition?

Moreover, the ontogeny of call usage in birds still deserves study. Call spectral features are unlearned but the purposeful use of calls might be a learned feature. For example, in a primate, the common marmoset (*Callithrix jacchus*), the turn-taking exchange of contact calls is learned behavior and parents provide an active role guiding the development of it

Definition of the repertoire

The amount and the precision of data collected with an on-board microphone and automatic recording of position provide an opportunity for better definitions of bird repertoires (e.g. **chapter 4**). One reason why defining the repertoire is important is that the size of the vocal repertoire, or the “functional repertoire size” (defined in Bradbury & Vehrencamp, 1998), has been often used as a proxy for vocal complexity (Freeberg & Krams, 2015; Holt, Barati, & McDonald, 2017). Social and vocal complexity have often been correlated and used as mutual explanation for their evolution (Freeberg, Dunbar, & Ord, 2012; Krams, Krama, Freeberg, Kullberg, & Lucas, 2012; Leighton, 2017; Wilkinson, 2003). But not all vocalization types have the same value with regard to communicative complexity (Fischer, Wadewitz, & Hammerschmidt, 2017). This is mainly because of one reason: vocalizations can have different numbers of functions. Certain vocalizations have one specific function, for example the *tsik* call of the common marmoset it is only emitted at the appearance of potential threat (Bezerra & Souto, 2008). Other vocalizations have multiple functions, for instance, the *phee* call of the common marmoset, stereotyped in acoustic structure, signals both the presence of an individual in absence of visual contact but it can also serve to coordinate the behavior of multiple individuals (Takahashi, Narayanan, & Ghazanfar, 2013).

Defining vocal complexity is difficult and should not be limited to repertoire size. I argue that to investigate the evolution and driving forces of social and vocal complexity rather than simply count the number of different vocalization it is more useful to study in detail closely related species and to try to correlate their differences in repertoire to different lifestyles. For example, among Estrildid finches there are examples of convergent evolution, related species with different life histories, varying degrees of sexual dimorphism or social propensity. For instance we could compare the strictly seedeater, male-only singing zebra finch and the omnivorous both sexes singing Blue-capped cordon-bleu, *Uraeginthus cyanocephalus*; or the zebra finch with its congener the double-barred finch, *Taeniopygia bichenovii* which has no sexual dimorphism (Clement, 1993). Quantifying their vocal behavior with the accuracy presented here and relating it to the different morphological, behavioral and ecological traits could clarify differences in the repertoire size and vocal usage, and help to make inferences concerning communicative complexity.

Neural control of call production

Unlearned vocalizations in vocal non learners are thought to be controlled only at the level of midbrain and medulla regions (Jarvis, 2006). Unlearned calls in vocal learners can also be produced just using the brain stem (Simpson & Vicario, 1990). However recent studies have found that different telencephalic areas are heavily involved in call perception and production (Benichov et al., 2016; Elie & Theunissen, 2015a; Ter Maat et al., 2014). These findings find parallels with the picture emerging in non-human primates, where innate calls are under extensive cognitive and volitional control (Hage, Gavrilov, & Nieder, 2013; Hage & Nieder, 2013). In zebra finches secondary auditory areas, such as the caudomedial nidopallium (NCM) of adult birds neurons code for specific calls (Elie & Theunissen, 2015a); interestingly NCM projects directly and indirectly to the motor pathway of the song control system that controls learned vocalizations (Nottebohm, Stokes, & Leonard, 1976; Vates, Broome, Mello, & Nottebohm, 1996). This motor pathway was recently shown to be activated also during the emission of unlearned calls (Ter Maat et al., 2014), and may be important for the timing of calling (Benichov et al., 2016). The function of the activation of the song motor pathway during call production is still unclear. It might control the timing of call production, integrating information about the individual recognition processed in the secondary auditory areas (Gentner, Hulse, & Ball, 2004) and/or the decision about which type of call to use in a given context. The studies of the neural control cannot be independent of a precise behavioral quantification. My thesis provides the behavioral tools and paradigms that can be investigated at a neural level; like recording the activity of the motor control pathway nuclei during the change of call rates that occurs when the birds form pair bonds.

Calls and the origin of vocal learning in birds

Hypotheses for the origin of vocal learning are multiple and often mixed with thoughts on the origin of language (reviewed in Nowicki & Searcy, 2014). Many times it has been argued that vocal learning should have evolved in social species and that the possible driving forces are: i) female preference for varied vocalizations, ii) adaptation to sound propagation in different environments, iii) kin selection for information sharing.

Parallels about vocal learning in mammals and birds are numerous and useful (Fitch, 2000; Hockett & Hockett, 1960; Janik & Slater, 2000). But with the increase in a number of the species studied I think it is time to focus also on the differences and on intra-taxon comparisons. Moreover, we need to study unlearned vocalization usage and its neural control in vocal learners, because learned vocalizations most certainly have evolved upon unlearned ones. I am going to argue that the understanding of call usage and its motor control might help us to trace back the origin of vocal learning in Passeriformes. The involvement of the song control system in behaviors other than singing (see the previous paragraph) is a fascinating finding when considered from an evolutionary point of view. The control of producing unlearned vocalizations - timing and context - during social encounters might have been the substrate on which the modification of vocalizations started. To verify these hypotheses two directions need to be explored:

i) Identification of the Passerines' song control system homologous areas of other vocal learners lineages (Gahr, 2000; Liu, Wada, Jarvis, & Nottebohm, 2013) and their pattern of activation during non-learned vocalization.

ii) Quantitative description in several species of the behaviors in which unlearned vocalizations are involved.

Once more species will have been studied, the findings in my thesis regarding the quantitative description of behavior can be interpreted from an evolutionary point of view.

Overall conclusions

My work on zebra finches expanded our knowledge of vocal exchanges in a social bird species and provided tools to improve precise quantification of behavior. My dissertation presents critical first evidence that zebra finches are capable of individual vocal recognition using unlearned, short, soft calls and that a calling relationship has to develop within a newly formed pair to reach the stereotypy of calling exchanges in established pairs. These findings highlight the importance of having new tools to answer old questions. The automation of behavioral recording and scoring that I describe in this work is a new way of compiling ethograms that reduces observer bias and improves accuracy. Moreover, they allow us to fuel the open discussions on broad fundamental questions such as the importance and mechanism of pair compatibility and the origin of vocal learning. Future studies can expand upon this groundwork to examine the specific mechanisms of pairing and individual vocal recognition at the neural level and in larger social groups. Finally, broadening the research to non-model species in comparative studies is necessary to provide an evolutionary explanation to my findings.

References

Adkins-Regan, E., & Tomaszycki, M. (2007). Monogamy on the fast track. *Biology Letters*, *3*(6), 617–9. https://doi.org/10.1098/rsbl.2007.0388

Anisimov, V. N., Herbst, J. A., Abramchuk, A. N., Latanov, A. V, Hahnloser, R. H. R., & Vyssotski, A. L. (2014). Reconstruction of vocal interactions in a group of small songbirds. *Nature Methods*, *11*(11), 1135–7. Retrieved from http://dx.doi.org/10.1038/nmeth.3114

Balsby, T. J. S., & Bradbury, J. W. (2009). Vocal matching by orange-fronted conures (Aratinga canicularis). *Behavioural Processes*, *82*(2), 133–139. https://doi.org/10.1016/j.beproc.2009.05.005

Beckers, G. J. L., & Gahr, M. (2012). Large-scale synchronized activity during vocal deviance detection in the zebra finch auditory forebrain. *The Journal of Neuroscience*, *32*(31), 10594–608. https://doi.org/10.1523/JNEUROSCI.6045-11.2012

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, *77*(4), 771–783. https://doi.org/10.1016/J.ANBEHAV.2008.12.022

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, *26*(3), 309–318. https://doi.org/10.1016/j.cub.2015.12.037

Bezerra, B. M., & Souto, A. (2008). Structure and usage of the vocal repertoire of *Callithrix jacchus*. *International Journal of Primatology*, *29*(3), 671–701. https://doi.org/10.1007/s10764-008-9250-0

Bousquet, C. A. H., Sumpter, D. J. T., & Manser, M. B. (2011). Moving calls: a vocal mechanism underlying quorum decisions in cohesive groups. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1711), 1482–1488. https://doi.org/10.1098/rspb.2010.1739

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*. Sinauer Associates

Sunderland.

Caccamise, D. F., & Hedin, R. S. (1985). An Aerodynamic Basis for Selecting Transmitter Loads in Birds. *The Wilson Bulletin*, *97*(3), 306–318. https://doi.org/10.2307/4162104

Candolin, U. (2003, November). The use of multiple cues in mate choice. *Biological Reviews of the Cambridge Philosophical Society*. Cambridge University Press. https://doi.org/10.1017/S1464793103006158

Chen, Y., Clark, O., & Woolley, S. C. (2017). Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1855), 1–9. https://doi.org/10.1098/rspb.2017.0054

Chiu, C., Xian, W., & Moss, C. F. (2009). Adaptive echolocation behavior in bats for the analysis of auditory scenes. *Journal of Experimental Biology*, *212*(9), 1392–1404. https://doi.org/10.1242/jeb.027045

Clement, P. (1993). *Finches & sparrows: an identification guide*. (C. Helm, Ed.). London; Princeton, New Jersey: Princeton University Press.

Cooper, B. G., & Goller, F. (2004). Multimodal Signals: Enhancement and Constraint of Song Motor Patterns by Visual Display. *Science*, *303*(5657), 544–546. https://doi.org/10.1126/science.1091099

Dalziell, A. H., Peters, R. A., Cockburn, A., Dorland, A. D., Maisey, A. C., & Magrath, R. D. (2013). Dance choreography is coordinated with song repertoire in a complex avian display. *Current Biology*, *23*(12), 1132–1135. https://doi.org/10.1016/j.cub.2013.05.018

Elie, J. E., & Theunissen, F. E. (2015a). Meaning in the avian auditory cortex: neural representation of communication calls. *The European Journal of Neuroscience*, *41*(5), 546–67. https://doi.org/10.1111/ejn.12812

Elie, J. E., & Theunissen, F. E. (2015b). The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition*, *19*(2), 285–315. https://doi.org/10.1007/s10071-015-0933-6

Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(21), 5976–81. https://doi.org/10.1073/pnas.1600970113

Fernandez, M. S. A., Vignal, C., & Soula, H. A. (2017). Impact of group size and social composition on group vocal activity and acoustic network in a social songbird. *Animal Behaviour*, *127*, 163–178. https://doi.org/10.1016/j.anbehav.2017.03.013

Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour*, *134*, 229–237. https://doi.org/10.1016/j.anbehav.2016.06.012

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267. https://doi.org/10.1016/S1364-6613(00)01494-7

Flack, J. C. (2013). Animal communication: Hidden complexity. *Current Biology*, *23*(21), R967–R969. https://doi.org/10.1016/j.cub.2013.10.001

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1597), 1785–801. https://doi.org/10.1098/rstb.2011.0213

Freeberg, T. M., & Krams, I. (2015). Does social complexity link vocal complexity and cooperation? *Journal of Ornithology*, *156*(S1), 125–132. https://doi.org/10.1007/s10336-015-1233-2

Gahr, M. (2000). Neural song control system of hummingbirds: Comparison to swifts, vocal learning (songbirds) and nonlearning (suboscines) passerines, and vocal learning (budgerigars) and nonlearning (dove, owl, gull, quail, chicken) nonpasserines. *Journal of Comparative Neurology*, *426*(2), 182–196. https://doi.org/10.1002/1096-9861(20001016)426:2<182::AID-CNE2>3.0.CO;2-M

Gentner, T. Q., Hulse, S. H., & Ball, G. F. (2004). Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *190*(12), 1001–10. https://doi.org/10.1007/s00359-004-0556-x

Gill, L. F., Goymann, W., Ter Maat, A., & Gahr, M. (2015). Patterns of call communication between group-housed zebra finches change during the breeding cycle. *eLife*, *4*, e07770. https://doi.org/10.7554/eLife.07770

Golabek, K. A., & Radford, A. N. (2013). Chorus-call classification in the southern pied babbler: Multiple call types given in overlapping contexts. *Behaviour*, *150*(7), 691–712. https://doi.org/10.1163/1568539X-00003081

Goodson, J. L., Kabelik, D., Kelly, A. M., Rinaldi, J., & Klatt, J. D. (2009). Midbrain dopamine neurons reflect affiliation phenotypes in finches and are tightly coupled to courtship. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(21), 8737–42. https://doi.org/10.1073/pnas.0811821106

Griesser, M., Wheatcroft, D., & Suzuki, T. N. (2018, June 1). From bird calls to human language: exploring the evolutionary drivers of compositional syntax. *Current Opinion in Behavioral Sciences*. Elsevier. https://doi.org/10.1016/j.cobeha.2017.11.002

Hage, S. R., Gavrilov, N., & Nieder, A. (2013). Cognitive Control of Distinct Vocalizations in Rhesus Monkeys. *Journal of Cognitive Neuroscience*, *25*(10), 1692–1701.

https://doi.org/10.1162/jocn_a_00428

Hage, S. R., & Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications*, *4*, 2409. https://doi.org/10.1038/ncomms3409

Hailman, J. P. (2008). *Coding and redundancy: man-made and animal-evolved signals*. Harvard University Press.

Hile, A. G., Plummer, T. K., & Striedter, G. F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Animal Behaviour*, *59*(6), 1209–1218. https://doi.org/10.1006/anbe.1999.1438

Hiryu, S., Shiori, Y., Hosokawa, T., Riquimaroux, H., & Watanabe, Y. (2008). On-board telemetry of emitted sounds from free-flying bats: compensation for velocity and distance stabilizes echo frequency and amplitude. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *194*(9), 841–51. https://doi.org/10.1007/s00359-008-0355-x

Hockett, C. F., & Hockett, C. D. (1960). The Origin of Speech. *Scientific American*. Scientific American, a division of Nature America, Inc. https://doi.org/10.2307/24940617

Holt, J., Barati, A., & McDonald, P. G. (2017). The complex acoustic repertoire of a highly social species, the Noisy Miner, Manorina melanocephala. *Emu - Austral Ornithology*, *117*(1), 19–30. https://doi.org/10.1080/01584197.2016.1252508

Ihle, M., Kempenaers, B., & Forstmeier, W. (2015). Fitness Benefits of Mate Choice for Compatibility in a Socially Monogamous Species. *PLoS Biology*, *13*(9), e1002248. https://doi.org/10.1371/journal.pbio.1002248

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1–11. https://doi.org/10.1006/ANBE.2000.1410

Jarvis, E. D. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science*, *5*(1), 5–14. https://doi.org/10.2326/osj.5.5

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1597), 1879–91. https://doi.org/10.1098/rstb.2011.0222

Kroodsma, D. E. (1986). Design of Song Playback Experiments. *Auk*, *103*(3), 640–642. https://doi.org/10.2307/4087148

Kroodsma, D. E., Byers, B. E., Goodale, E., Johnson, S., & Liu, W.-C. (2001). Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour*, *61*(5), 1029–1033. https://doi.org/10.1006/anbe.2000.1676

Langmore, N. E., & Bennett, A. T. D. (1999). Strategic concealment of sexual identity in an estrilid finch. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1419), 543–550. https://doi.org/10.1098/rspb.1999.0670

Laubu, C., Dechaume-Moncharmont, F.-X., Motreuil, S., & Schweitzer, C. (2016). Mismatched partners that achieve postpairing behavioral similarity improve their reproductive success. *Science Advances*, *2*(3), e1501013. https://doi.org/10.1126/sciadv.1501013

Leighton, G. M. (2017). Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings. Biological Sciences*, *284*(1868), 20171508. https://doi.org/10.1098/RSPB.2017.1508

Levinson, S. C. (2015). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Liu, W. C., Wada, K., Jarvis, E. D., & Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nature Communications*, *4*, 2082. https://doi.org/10.1038/ncomms3082

Ma, S., Maat, A. Ter, & Gahr, M. (2017). Power-law scaling of calling dynamics in zebra finches. *Scientific Reports*, *7*(1), 8397. https://doi.org/10.1038/s41598-017-08389-w

Marler, P. (2004). Bird Calls: Their Potential for Behavioral Neurobiology. *Annals of the New York Academy of Sciences*, *1016*(1), 31–44. https://doi.org/10.1196/annals.1298.034

McGregor, P. K. (2005). *Animal communication networks*. Cambridge University Press.

Miller, D. B. (1979). The acoustic basis of mate recognition by female Zebra finches (Taeniopygia guttata). *Animal Behaviour*, *27*, 376–380. https://doi.org/10.1016/0003-3472(79)90172-6

Nottebohm, F., Stokes, T. M., & Leonard, C. M. (1976). Central control of song in the canary,Serinus canarius. *The Journal of Comparative Neurology*, *165*(4), 457–486. https://doi.org/10.1002/cne.901650405

Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, *28*, 48–53. https://doi.org/10.1016/j.conb.2014.06.007

Ota, N., Gahr, M., & Soma, M. (2015). Tap dancing birds: The multimodal mutual courtship display of males and females in a socially monogamous songbird. *Scientific Reports*, *5*, 16614. https://doi.org/10.1038/srep16614

Perez, E. C., Elie, J. E., Soulage, C. O., Soula, H. A., Mathevon, N., & Vignal, C. (2012). The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? *Hormones and Behavior*, *61*(4), 573–81. https://doi.org/10.1016/j.yhbeh.2012.02.004

Reichard, D. G., & Anderson, R. C. (2015). Why signal softly? The structure, function and evolutionary significance of low-amplitude signals. *Animal Behaviour*, *105*, 253–265. https://doi.org/10.1016/j.anbehav.2015.04.017

Schuett, W., Dall, S. R. X., & Royle, N. J. (2011). Pairs of zebra finches with similar “personalities” make better parents. *Animal Behaviour*, *81*(3), 609–618. https://doi.org/10.1016/j.anbehav.2010.12.006

Schuett, W., Tregenza, T., & Dall, S. R. X. (2010). Sexual selection and animal personality. *Biological Reviews*, *85*(2), 217–246. https://doi.org/10.1111/j.1469-185X.2009.00101.x

Silcox, A. P., & Evans, S. M. (1982). Factors affecting the formation and maintenance of pair bonds in the zebra finch, *Taeniopygia guttata*. *Animal Behaviour*, *30*(4), 1237–1243. https://doi.org/10.1016/S0003-3472(82)80216-9

Simpson, H. B., & Vicario, D. S. (1990). Brain pathways for learned and unlearned vocalizations differ in zebra finches. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *10*(5), 1541–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2332796

Stowell, D., Gill, L., & Clayton, D. (2016). Detailed temporal structure of communication networks in groups of songbirds. *Journal of the Royal Society, Interface*, *13*(119), 20160296. https://doi.org/10.1098/rsif.2016.0296

Suzuki, R., Matsubayashi, S., Saito, F., Murate, T., Masuda, T., Yamamoto, K., … Okuno, H. G. (2018). A spatiotemporal analysis of acoustic interactions between great reed warblers (Acrocephalus arundinaceus) using microphone arrays and robot audition software HARK. *Ecology and Evolution*, *8*(1), 812–825. https://doi.org/10.1002/ece3.3645

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, *7*, 10986. https://doi.org/10.1038/ncomms10986

Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology*, *23*(21), 2162–2168. https://doi.org/10.1016/J.CUB.2013.09.005

Ter Maat, A., Trost, L., Sagunsky, H., Seltmann, S., & Gahr, M. (2014). Zebra finch mates use their forebrain song system in unlearned call communication. *PloS One*, *9*(10), e109334. https://doi.org/10.1371/journal.pone.0109334

Theunissen, F. E., & Elie, J. E. (2014). Neural processing of natural sounds. *Nature Reviews. Neuroscience*, *15*(6), 355–66. https://doi.org/10.1038/nrn3731

Tomaszycki, M. L., & Atchley, D. (2017). Pairing Increases Activation of V1aR, but not OTR, in Auditory Regions of Zebra Finches: The Importance of Signal Modality in Nonapeptide-Social Behavior Relationships. *Integrative and Comparative Biology*, *57*(4), 878–890. https://doi.org/10.1093/icb/icx043

Ullrich, R., Norton, P., & Scharff, C. (2016). Waltzing Taeniopygia: integration of courtship song and dance in the domesticated Australian zebra finch. *Animal Behaviour*, *112*, 285–300. https://doi.org/10.1016/J.ANBEHAV.2015.11.012

Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (Taenopygia guttata). *Journal of Comparative Neurology*, *366*(4), 613–642. https://doi.org/10.1002/(SICI)1096-9861(19960318)366:4<613::AID-CNE5>3.0.CO;2-7

Vignal, C., Mathevon, N., & Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature*, *430*(6998), 448–51. https://doi.org/10.1038/nature02645

Vignal, C., Mathevon, N., & Mottin, S. (2008). Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behavioural Processes*, *77*(2), 191–8. https://doi.org/10.1016/j.beproc.2007.09.003

Wiley, R. H. (2013). Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biological Reviews*, *88*(1), 179–195. https://doi.org/10.1111/j.1469-

185X.2012.00246.x

Wilkinson, G. S. (2003). Social and vocal complexity in bats. In *Animal social complexity: Intelligence, culture, and individualized societies*. (pp. 322–341). Cambridge, MA, US: Harvard University Press.

Zann, R. A. (1996). *The zebra finch: a synthesis of field and laboratory studies*. *Oxford Univ, Press*. Retrieved from http://books.google.com/books?id=5KO6cZH0WbEC&pgis=1

Author contributions

Chapter 1: L.F.G.*, P.B.D.* and M.N.A.* share the first authorship and the order in which the names appear was chosen at random. L.F.G.*, P.B.D.* and M.N.A.* carried out conception and study design. L.F.G.*, P.B.D.* and M.N.A.* performed experiments and analysed results. A.T. and M.C.G. advised on the analyses and on the interpretation of the results. A.T. and M.C.G. conceptualised the wireless microphones and drove their development. H.S provided key technical expertise. L.F.G.*, P.B.D.* and M.N.A.* wrote the manuscript with support from all authors.

………………………………….. ……………………………………………

Lisa Franziska Gill Mauricio Nicolas Adreani

Chapter 2: P.B.D. conceived the study. P.B.D., M.N.A. and M.K. designed the experiment. M.K. performed the experiment. M.K. processed the data. P.B.D. and M.K. analysed the data with input from all authors. P.B.D. and M.K. wrote the manuscript with input from all authors. A.T. and M.G. provided supervision throughout the project.

Chapter 3: PBD, LT & AT conceived the study and designed the experiment. PBD performed the experiment. PBD processed the data. PBD & AT analysed the data with input from all authors. PBD wrote the manuscript with input from all authors. AT & LT provided supervision during the entire project. All authors read and approved the final manuscript.

Chapter 4: PBDA and AtM conceived of the study design. PBDA collected data, analyzed data, and wrote the manuscript. AtM and MG supervised and contributed to project planning, data analyses, and write up.

………………………………….. ……………………………………………

Manfred Gahr Pietro Bruno D'Amelio
Doktorvater Doktorand

Acknowledgements

I am deeply grateful to the three supervisors of the present work. Manfred Gahr provided unwavering support, fruitful discussions and original point of views that improved my work and my science understanding. Andries ter Maat gave me constant supervision; he was always there when needed. His experience and wisdom thought me well beyond the scope of the present work. I am incalculably thankful to Lisa Trost, she was the first one to believe in me when I came to the Max Planck as a student; I owe her a lot for this amazing experience. Lisa and Andries had to endure my complains while struggling with failures in science; they had to teach me to focus on what I can learn from failures and not to what has been lost.

My friend and colleague Nico Adreani played a key role in the development of the present dissertation. Together we discussed, planned, failed, succeeded, failed again and argued about almost every step of this work, gracias Nic for that and your friendship.

I am indebted to Wolfgang Forstmeier, Wolfgang Goymann, Alfredo Sanchez-Tojar, Lisa Gill, Albertine Leitao, Chiel Vellema and Fränzi Korner-Nievergelt for fruitful discussions that improved the quality of the dissertation and me as a scientist. I am grateful to Robin Abbey-Lee, Erica Stuber and Maude Baldwin for teaching me and correcting my English in multiple occasions. If my English is still poor it’s on me, you did a great job.

Vielen dank Hannes Sagunsky, Markus Abels and Cristina Reusch for providing key technical expertise and support in the lab and in building up the set-ups of the different experiments. I learned something every day from my colleagues of the Gahr group, thank you for being always supportive. I am especially grateful to Nicole Fritz for being always helpful and to Susanne Rössel for the lessons, support and laughs.

I want to thank all the animal care takers and especially Roswitha Brighton for taking very good care of my experimental birds, being flexible and put up with my weird experimental schedules.

Thank you Daniel Piechowski and Maggie Hieber Ruiz for coordinating the IMPRS, it provided high quality training and even better fun. I also thank the IMPRS PAC members Steffen Hage and Katharina Riebel for insightful comments on my work.

My mental health during these years was preserved thank to the friendship of numerous people, among them: Robin Abbey-Lee, Nico Adreani, Rocìo Agar, Yimen Araya-Ajoy, Lisa Arvidsson, Chiara Bonifazi, Giacomo Bucchi, Martin Bulla, Francesco D'Amelio, Davide Dominoni, Arnaud Da Silva, Andrea Capperucci, Tommaso Cecchini, Carlo Colonelli, Mariana Diales da Rocha, Luke Eberhart-Phillips, Massimo Fedi, Antje Girndt, Stefan Greif, Elena Kappers, Maggie Ko, Klemen Koselj, Uli Knief, Marco Marasca, Lucia Mentesana, Amanda Monte, Marcos Perez, Paolo Politi, Rene Quispe, Ignas Safari-Mnganya, Alfredo Sánchez-Tójar, Jessica Schnell, Francesca Santostefano, Susanne Seltmann, Michele Spreafico, Elias Stitch, Erica Stuber, Lorenzo Vanni, Camilla Villavicencio, Daniel Zuniga and Giulia Pumbi Bambini.

Most of all I want to thank my parents, Grazia e Salvatore, sapere che posso sempre contare su di voi è un privilegio raro, voler ricambiare la vostra fiducia mi sprona ad essere una persona piú positiva con me stesso e il prossimo.

Pietro D'Amelio
Seewiesen, 4th March 2018

Curriculum vitae

Pietro Bruno D'Amelio

Max Planck Institute for Ornithology
Behavioural Neurobiology - Department Gahr
82319, Seewiesen
Tel: +49 (0)8157 932 263
Email: pdamelio@orn.mpg.de; pie.damelio@gmail.com
Date of Birth: 12th November, 1986

Education

- **08/2012 – Present: Max Planck Institute for Ornithology, Seewiesen, Germany**
  Ph.D. student, dissertation: "*Vocal communication in zebra finches: a focused description of pair vocal activity*"

- **10/2009-11/2011: Pisa University, Italy**
  M.Sc. in Conservation and Evolution (110 *cum laude*/110)
  Thesis: "*Development of the Zebra Finch communication system, from pair to colony*"

- **10/2005-3/2009**: **Pisa University, Italy**
  B.Sc. in Ecology and Biodiversity (110 *cum laude*/110)
  Thesis: "*Anti-angiogenic effects of somatostatin: a potential therapeutic effect*"

Publications

- **Gill L.F.***, **D'Amelio P.B.***, **Adreani N.M.***, Sagunsky H., Gahr C.M., ter Maat A. A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods Ecol Evol*. 2016 Jun; Available from: http://doi.wiley.com/10.1111/2041-210X.12610
  ***joint first authors**

- **D’Amelio P.B.,** Trost L., ter Maat A. Vocal exchanges during pair formation and maintenance in the zebra finch (*Taeniopygia guttata*). *Frontiers in Zoology*, 2017, 14(1), 13; Available from: http://doi.org/10.1186/s12983-017-0197-x
- **D’Amelio P.B.**, Klumb. M., Adreani N.M., Gahr C. M., ter Maat A. Individual recognition of opposite sex vocalizations in the zebra finch. *Scientific Reports*, 2017, 7(1), 5579; Available from: http://doi.org/10.1038/s41598-017-05982-x

Under review:

- Carbeck K., DeMoranville K., **D'Amelio P.,** Goymann G., Trost L., Pierce B., Bryła A, Dzialo M., Bauchinger U., McWilliams S. Environmental cues and dietary antioxidants affect breeding behavior and testosterone of male European Starlings (*Sturnus vulgaris*)

Conference Talks

2018

- Society for Integrative and Comparative Biology (SICB) annual meeting, San Francisco, USA

2017

- European Ornithological Union (EOU), Turku, Finland
- XXVII Convegno Societa’ Italiana di Etologia (SIE), Pisa, Italy

2015

- 25th International Bioacoustics Council (IBAC), Murnau, Germany

2014

- 26th International Ornithological Conference (IOC), Tokyo, Japan

Conference Posters

2017

- 6th international bio-logging science symposium (BLS), Konstanz, Germany

2015

- Behaviour2015, Cairns, Australia
- 11th German Neuroscience Society, Goettingen, Germany

2013

- 106th German Zoological Society (DZG), Munich, Germany

Honors and Awards

- International Max Planck Research School Student Travel Grant, 2015, 2017
- International Max Planck Research School Project Grant, 2015
- ERASMUS Student Mobility for Placements, 2011

Statutory declaration and statement

Ehrenwörtliche Versicherung

Ich versichere hiermit an Eides statt, dass die von mir vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den ...06.03.2018......................... ....
Pietro Bruno D'Amelio

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Im Weiteren erkläre ich, dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

München, den ...06.03.2018......................... ....
Pietro Bruno D'Amelio