Rhythm in Animals' Acoustic Signals: Novel Methods for the Analysis of Rhythm Production and Perception on the Example of Bats, Birds, and Whales

Kumulative Dissertation zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
im Fach Biologie

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

von
Lara Sophie Burchardt

Berlin im Mai 2021

Die vorliegende Doktorarbeit wurde von Oktober 2017 bis Mai 2021 am Institut für Biologie an der Freien Universität Berlin in der Arbeitsgruppe „Communication and Cognition in Bats“ unter der Leitung von PD. Dr. Mirjam Knörnschild angefertigt. Finanziert wurde die Arbeit durch ein Elsa-Neumann Stipendium der Graduiertenförderung des Landes Berlin.

1. Gutachterin: PD. Dr. Mirjam Knörnschild

2. Gutachterin: Prof. Constance Scharff, Ph.D.

Disputation am 29.07.2021

Danksagung

Die allerwenigsten Dinge schafft man allein, für Doktorarbeiten gilt das ganz bestimmt: von wem würde man lernen, wer soll einen anstacheln diese Analyse oder jenes Experiment doch durchzuführen? Mit wem sollte man diskutieren, stundenlang an dieser einen Abbildung feilen, oder diesen vermaledeiten Code endlich zum Laufen bringen? Wer soll einem die eigene Arbeit zwar konstruktiv, aber gründlich auseinandernehmen? Wie um Himmelswillen würde man all die Daten aufnehmen, die kleinteilige Arbeit schaffen und gleichzeitig noch Experte für Alpha und Omega sein? Wer sollte einen auf die ganzen Kommafehler hinweisen und mit wem soll man hin und wieder das wohlverdiente Feierabendbier genießen?

Aber das Ganze fängt noch viel früher an, erstmal muss man diese ganz besonderen Leute haben, die einem den Glauben an sich selbst mitgeben und die die Begeisterung wecken und aufrechterhalten. Man braucht die Menschen, mit denen man an sonnigen und regnerischen Tagen einfach nur das Leben genießen und die Arbeit vergessen kann. Ganz besonderes Glück hat man, wenn man diesen einen Menschen hat, zu dem man abends nach Hause kommt und einfach nur noch Mensch sein muss.

Ich hatte in den letzten fast vier Jahren und auch vorher all diese Menschen und mehr um mich und das macht mich dankbar und glücklich. Wenn ich vieles aus diesen vier Jahren vergessen sollte, die Menschen und was sie mir beigebracht haben ganz bestimmt nicht.

Mirjam, die mir anbot eine Doktorarbeit bei ihr zu schreiben und die mich erst mit diesem spannenden Thema von Rhythmen in der Kommunikation von Tieren bekannt gemacht hat. Ich habe so viel von dir gelernt (nur die richtige Kommasetzung immer noch nicht, ich bin wohl einfach ein hoffnungsloser Fall). Danke für dein beständiges Vertrauen und deine fortwährende Unterstützung!

Constance, meine Zweitgutachterin, die jedem zeigt, wie gute Wissenschaft funktioniert, dass man dabei auch für sich selbst einstehen und mutig seinen Weg gehen muss. Danke, für dein Interesse an Rhythmen, dass du Mirjam damit bekannt gemacht hast und so auch diese Doktorarbeit erst möglich gemacht hast.

Ahana, mit der ich Büros umgeräumt, stundenlang über Gott, die Welt und natürlich auch über die Arbeit diskutiert habe und die so viel mehr als eine Kollegin für mich geworden ist.

Martina und Anja. Niemand ist so konstruktiv kritisch wie Martina und ich hätte mir keine bessere Büromitbewohnerin als Anja wünschen können.

Ella, Leonie und Meike die alle drei substanziell dazu beigetragen haben, dass die ABR Experimente erfolgreich liefen. Genauso dazu beigetragen haben Sina, Arpik, Florian und Jasmin.

Meine fantastischen anderen Co-Autori*innen Oliver Behr, Elodie Briefer, Inga Geipel, Gloria Gessinger und Philipp Norton.

Pernille Tønnesen, Pernille Meyer und Peter Madsen, die mir ihre Daten überlassen und mich unterstützt haben. Ihr habt schon dazu beigetragen, dass mein Auslandssemester in Dänemark so großartig war. Tusind tak!

Andrea Ravignani und Julio Hechavarria für ihre Expertise und ihr Interesse an meiner Arbeit. Thank you so much.

Frau Schwarz, weil wir alle wissen, dass ohne Frau Schwarz im FB1 am MfN nicht mehr viel funktionieren würde.

Alle Kolleg*innen in der AG Knörnschild, der AG Scharff und der AG Vallentin, danke für die schöne und lustige Zeit.

Meine Lehrer*innen Frau Gödecke und Herrn Graeber, die mich von Anfang an gefördert haben.

Mama, Papa und Hilke. Dazu muss man nicht mehr sagen.

Opa, der mich schon als Kind für die Natur und die Biologie begeistert hat.

Cit. Du bist mein „partner in crime“, mein „partner in Corona“ und in allen anderen Lebenslagen.

Ich danke euch von Herzen für eure Zeit, euer Wissen, eure Kritik, eure geteilten Gedanken, Anregungen, Ideen und eure Unterstützung, euer Zuhören und Bestärken, selbst wenn ihr euch manchmal gefragt habt, was da wohl gerade in meinem Kopf vor sich geht. Ich danke euch für euer Lachen und eure Scherze und eure Herzlichkeit. Diese Arbeit und ich wären ohne euch nicht dieselbe.

Für meine Familie

„Das ist Rhythmus, Adalbert, das ist das Leben!“

Mimmelitt das Stadtkaninchen

Publication List and Author Contributions

At the time of thesis submission, two of the four Chapters had been published in peer-review journals, one individual paper had been submitted and one was in preparation.

Publication A: published

Burchardt, L. S.; Norton, P.; Behr, O.; Scharff, C.; Knörnschild, M. (2019) General isochronous rhythm in echolocation calls and social vocalizations of the bat Saccopteryx bilineata, In: Royal Society Open Science 6: 181076. https://dx.doi.org/10.1098/rsos.181076

Author Contribution

OB and MK collected the data. PN and CS developed the computational framework (GAT approach). LSB analysed the data and wrote the manuscript. CS and MK supervised the project. All authors discussed the results and contributed to the final manuscript. All authors gave final approval for publication.

Publication B: published

Burchardt, L. S.; Knörnschild, M. (2020): Comparison of methods for rhythm analysis of complex animals' acoustic signals. In: PLoS Computational Biology 16(4): e1007755. https://doi.org/10.1371/journal.pcbi.1007755

Author Contribution

MK collected and prepared the bat data. LSB prepared the sperm whale data, established, and conducted the analyses and wrote the manuscript. LSB and MK discussed the results, both authors contributed to the final manuscript. Both authors gave final approval for publication.

Publication C:

Burchardt, L. S.; Lattenkamp, E. Z.; Gessinger, G.; Geipel, I.; Knörnschild, M. (2021) Adjusted auditory brainstem response procedure to measure rhythm perception in small mammals.

Author Contribution

LSB developed the idea. EZL built the setup. EZL and LSB wrote the code to adjust the setup to measure rhythm perception. GG, IG and MK recorded bat echolocation calls. EZL and MK conducted Experiment 1 in Panama. LSB and MK conducted Experiment 2 in Germany. LSB analysed the data and wrote the manuscript. MK supervised the project. All authors worked on and approved the final manuscript.

Publication D: submitted to "Ecology and Evolution" (August 2021)

Burchardt, L. S.; Briefer, E. F.; Knörnschild, M. (2021): Novel ideas to further expand the applicability of rhythm analysis.

Author Contribution

EFB recorded the skylark data. LSB conducted the analysis, developed the method additions, and wrote the manuscript. MK supervised the work. All authors contributed to the final manuscript and approved publication.

CS	Constance Scharff
EFB	Elodie F. Briefer
EZL	Ella Z. Lattenkamp
GG	Gloria Gessinger
IG	Inga Geipel
LSB	Lara S. Burchardt
MK	Mirjam Knörnschild
OB	Oliver Behr
PN	Philipp Norton

Zusammenfassung

Rhythmen, diese regelmäßigen Abfolgen von Ereignissen über die Zeit, sind sehr wichtig für alle möglichen Aspekte von akustischer Kommunikation. Wir versuchen, ihre Bedeutung für Kommunikation in Bezug auf ihre Wahrnehmung und Produktion sowohl bei Menschen als auch bei anderen Tieren zu verstehen. Ich untersuchte diese Rhythmen in akustischen Signalen von Tieren und konzentrierte mich insbesondere auf die Evaluierung und Entwicklung von Methoden zur vergleichbaren und reproduzierbaren Analyse von rhythmischen Strukturen. Ich untersuchte die Rhythmusproduktion bei Fledermäusen, Vögeln und Walen und studierte die Rhythmuswahrnehmung bei verschiedenen Fledermausarten. Untersucht wurden vor allem isochrone Muster, das sind sehr einfache, Metronom-ähnliche Strukturen. Insgesamt wurden 17 Datensätze von 14 verschiedenen Arten mit 940 Lautsequenzen analysiert und exakte isochrone Beat Frequenzen (in Hertz, so wie in Beats pro Sekunde) berechnet, die die Lautsequenzen gut beschreiben. Ein weiterer wichtiger Parameter in diesem Zusammenhang sind Gütewerte, die angeben, wie gut eine berechnete Schwebung eine einzelne Sequenz beschreibt.

In einer ersten Studie wurden die isochronen Beat Frequenzen für drei verschiedene Sequenztypen der Großen Sackflügelfledermaus *Saccopteryx bilineata* analysiert. Ein Sequenztyp ist durch eindeutige Lautcharakteristika oder einen eindeutigen Verhaltenskontext definiert. Die Analysen erfolgten mit einem Generate-and-Test-Ansatz. Alle drei Sequenztypen haben einen gemeinsamen Rhythmus von ca. 6 bis 24 Hz, der mit den Flügelschlagfrequenzen (ca. 12 Hz) dieser Art in Verbindung gebracht werden kann. Dabei wurden zwei der drei untersuchten Sequenztypen geäußert während die Fledermäuse gar nicht flogen. Anschließend erstellte ich eine Anleitung, wie man die zeitliche Struktur – oder den Rhythmus – der akustischen Signale eines beliebigen Tieres mit Methoden analysiert, die für eine breite Palette von Signalen anwendbar sind und deren Ergebnisse leicht vergleichbar und interpretierbar sind. Diese Anleitung wird das Verständnis der Rhythmik in akustischen Signalen von Tieren verbessern und vor allem den Vergleich zwischen verschiedenen Arten erleichtern. Die Methoden, die einbezogen wurden, reichen von einfachen Verteilungs- und visuellen Analysen bis hin zu höherer Mathematik wie der Fourier-Analyse. Alle Analysen basieren auf Inter-Onset-Intervallen, also den Zeitintervallen

zwischen dem Beginn eines Lautelements und dem nächsten Lautelement in einer bestimmten Sequenz. Für diese Anleitung wurden verschiedene bereits etablierte, sowie von mir neu entwickelte Methoden an drei Datensätzen getestet: Isolationsrufe der beiden Fledermausarten *S. bilineata* und *Carollia perspicillata* sowie Echoortungs-Sequenzen eines weiblichen Pottwals (*Physeter macrocephalus*). Ein wesentliches Ergebnis dieser Arbeit war darüber hinaus die Entwicklung eines universellen Gütewerts, der angibt, wie gut eine Beat Frequenz in Hertz eine beliebige Elementsequenz beschreibt. Er ist für verschiedene Methoden anwendbar und leicht vergleichbar. Er wurde an allen oben genannten Datensätzen sowie an komplexen Fluggesängen der Feldlerche *Alauda arvensis* getestet. Die Fluggesänge dienten auch dazu, die Verwendung der Fourier-Analyse für die Rhythmusanalyse komplexerer Signale anzupassen, sowie die Verwendung sogenannter Recurrenceplots zur Identifikation von Substrukturen in einer komplexen Lautfolge anhand ihrer zeitlichen Struktur zu veranschaulichen.

Darüber hinaus wurde das sogenannte „Auditory Brainstem Response"-Verfahren – eine etablierte Methode, die akustisch evozierte Feldpotentiale im auditorischen Hirnstamm misst – angepasst, um die Rhythmuswahrnehmung bei kleinen Säugetieren zu untersuchen. Seine allgemeine Anwendbarkeit konnte in dieser Arbeit bestätigt werden. Am Beispiel von 12 in Mittelamerika heimischen Fledermausarten wurden Unterschiede in der Wahrnehmungsstärke in Abhängigkeit von der Stimulus-Präsentationsrate für künstliche und natürliche Reize bei untrainierten Tieren aus freier Wildbahn und in Gefangenschaft gefunden. Niedrigere Stimulus Präsentationsraten, d.h. langsamere isochrone Rhythmen, lösten durchweg höhere Reaktionen aus als schnellere Stimulus Präsentationsrate. Die Wahrnehmungsrhythmen konnten teilweise mit den Produktionsrhythmen von Echoortungssequenzen der 12 untersuchten Arten abgeglichen werden, die ebenfalls isochrone Rhythmen aufwiesen.

Diese Arbeit kann zu einem besseren Verständnis von Rhythmen in akustischen Signalen von Tieren beitragen und bei der Suche nach adaptiven Funktionen und der Evolution von Rhythmus in der akustischen Kommunikation sowohl beim Menschen als auch bei anderen Tieren helfen. Sie kann darüber hinaus dazu beitragen, die Forschung zur Evolution der Sprache sowie der Musik zu fördern. Zudem trägt sie zum allgemeinen Wissen über die verschiedenen Aspekte und ihre Bedeutung der akustischen Kommunikation bei Tieren bei.

Index

Figure Index

Table Index

General Introduction

Chapter II

Chapter II - Supplements

Chapter III - Supplements

General Discussion

Summary

Rhythms, these systematic patterns of events in time, are very important for all kinds of aspects of communication. We are trying to understand their importance for communication with regards to their perception and production in both humans and other animals. My thesis investigated these rhythms in animals' acoustic signals and especially focused on the evaluation and development of method for the comparable and reproducible analysis of rhythmic structures. I investigated the rhythm production in bats, birds and whales and studied rhythm perception in various bat species. Under study were mostly isochronous patterns, which are very simple, metronome-like structures. A total of 17 datasets from 14 different species including 940 sound sequences were analysed and exact isochronous beat frequencies (in Hertz as in beats per second) calculated, that describe the sound sequences well. Isochronous beat frequencies were analysed for three different sequence types (i.e., uttered in different distinct contexts) of the greater sac-winged bat *Saccopteryx bilineata*, where analysed with a generate-and-test approach. They share a common rhythm of around 6 to 24 Hz, that can be linked to the wingbeat (around 12 Hz) frequencies of that species, even though two of the three analysed sequence types were uttered while bats were not flying. I then establish a workflow on how to analyse the temporal structure – namely the rhythm – of any animals' acoustic signal with methods that are applicable for a wide range of signals and results that are easily comparable and interpretable. This workflow will enhance the understanding of rhythmicality in animals' acoustic signals as well as facilitate cross-species comparison. Methods that were included ranged from simple distributional and visual analysis to higher mathematics such as Fourier analysis. All analyses rely on Inter-Onset-Intervals, the duration between the beginning of one sound element and the next sound element in a given sequence. For the workflow different already established and newly developed methods were tested on three datasets: isolation calls of the two bat species *S. bilineata* and *Carollia perspicillata* and echolocation call sequences of a female sperm whale (*Physeter macrocephalus*). Furthermore, an auditory brainstem response procedure – a well-established method, measuring evoked field potentials in the auditory brainstem – was adjusted to measure rhythm perception in small mammals. Its general applicability could be confirmed in this thesis. Using 12 species of Central American bats as an example, we found differences in perception strength depending on the stimulus presentation rate for artificial and natural stimuli in untrained wild and captive bats.

Lower stimulus presentation rates, i.e., slower isochronous rhythms, consistently elicited higher reactions than faster stimulus presentation rates. Perception rhythms could in parts be matched to the production rhythms of echolocation call sequences of the 12 tested species, that also showed isochronous rhythms. A key finding of this work was the development of a universal goodness-of-fit value, indicating how well a beat frequency in Hertz describes any element sequence, it can be applied for various methods and is easily comparable. It was tested on all datasets mentioned above, as well as on complex flight songs of the skylark *Alauda arvensis*. The flight songs also served to adjust the use of Fourier analysis for the rhythm analysis of more complex signals, as well as to illustrate the use of so-called recurrence plots for the identification of substructures in a complex sound sequence using their temporal structure.

This work can contribute to a better understanding of rhythms in animals' acoustic signals and help in the quest to uncover adaptive functions and the evolution of rhythmicality in acoustic communication in humans and other animals alike, furthering research on the evolution of language as well as music, and the general knowledge about the different aspects and their importance of acoustic communication in animals.

General Introduction

Boom Boom Chack. Boom Boom Chack. Boom Boom Chack. The power of rhythm is apparent already in these few words, as many people likely do not only read them in the intended rhythm but also recognize the song, they are so iconic for: 'We will rock you' by Queen. While this is a very prominent example of a well-executed musical rhythm, rhythms are very important for all kinds of aspects of communication. We are trying to understand their importance for communication with regards to their perception and production in both humans and other animals. My thesis adds to that by focusing on acoustic rhythm production and perception in mammals and birds, and by discussing the advantages and disadvantages of different existing methods, while also introducing new approaches.

Acoustic communication

Communication is a major driver of life on this planet, it is a key component of human culture. But what is it exactly and through which modalities does it work? Communication is the sharing of information between two or more subjects. For this sharing of information to work, both sender and receiver need to be able to understand the information and intent. This can be ensured by using a signal repertoire mutually understood by all subjects taking part in the communication (Bradbury & Vehrencamp, 2011). These signals can be conveyed through different modalities, and while we also know various examples for olfactory (Eisenberg & Kleiman, 1972) or gestural communication (Call & Tomasello, 2020), my thesis focuses on acoustic communication.

Acoustic communication works over long distances, without light, in air and water, and without the necessity that sender and receiver see each other (Bradbury & Vehrencamp, 2011). Nevertheless, acoustic signals often contain the necessary information to recognize an individual (Charrier et al., 2009; Knörnschild et al., 2013; Knörnschild & von Helversen, 2008; Mathevon et al., 2010). The perks of acoustic communication signals are the risk of being overheard, masking effects that can hinder signal transmission, and physical processes that will degrade an acoustic signal in both the frequency and the temporal domain, as well as its amplitude over larger distances (Forrest, 1994; Ryan & Sullivan, 1989; Wiley & Richards, 1978).

Information that is transmitted via acoustic signals can range from mate-attraction signals, over courtship or territorial defence signals, contact calls, alarm calls and threat signals, to group cohesion signals, or parent-offspring interactions (Bradbury & Vehrencamp, 2011). Acoustic communication can differ between species with regards to production mechanisms and specific function. The anatomy of the vocal apparatus, body size and environment differ, resulting in different sounds with regards to pitch, timbre, entropy, duration or the temporal structure between sound elements to name but a few important parameters of acoustic signals (Bradbury & Vehrencamp, 2011). Acoustic communication in animals is most studied in mammals, birds, frogs, and insects.

A complete communication attempt can be divided into smaller units. One unit that many analyses work with is one element of sound within a sequence of sounds (Figure 1A & 1B). One element is clearly visible for example in a spectrogram or oscillogram (two ways to visualize sound) and framed by silence. 'Element' is used here as the term for this clearly distinguishable small sound unit (Figure 1B). Depending on the context or study species, "element" could describe a syllable, call, pulse, click, sound or else (syllable: i.e.,(Behr et al., 2006; E. Briefer et al., 2008; Hultsch & Todt, 1989); call: i.e.,(Charrier et al., 2009; Ratcliffe et al., 2013); pulse: (Bøttcher et al., 2018; Moss et al., 2006); click: i.e.,(Ladegaard et al., 2015a; Le Bot et al., 2013), sound: (Bolgan et al., 2018)). Different elements can have very different spectral and temporal properties and can be grouped into different element types according to these properties within one species. Several elements in succession are called an element sequence (Figure 1A) and sequences again can differ in a recognizable pattern, making it possible to distinguish different sequence types (also i.e., vocalization types (Knörnschild et al., 2006), motifs (Hyland Bruno & Tchernichovski, 2017; Norton & Scharff, 2016) or phrases (E. Briefer et al., 2008)). The structure and complexity of sequences can differ enormously between species. Sequences can be very simple, being comprised of only one element type and only a few elements being uttered in direct succession (i.e., isolation calls of the bat *Carollia perspicillata* (Knörnschild et al., 2013)), over one element type being uttered in very long sequences (i.e., echolocation call sequences of toothed whales (Bøttcher et al., 2018)) to very complexly structured sequences with a huge number of element types being combined in an elaborate syntax in very long sequences. One individual skylark or nightingale for example can

produce up to ~800 different element types and combine them in such a complex syntax (nightingales: i.e., (Hultsch & Todt, 1981), skylarks: .i.e., (Briefer, Osiejuk, et al., 2010)).

How to analyse these elements and sequences, their structure both temporally and syntactically depends very much on the question. Generally, questions could be aimed at understanding the biology of animals, their ecology, and how they are shaping and using their environment but could also be asked to shed light on vocal communication in humans or understand the underpinnings and evolution of human language or music better.

Language, believed to be a trait unique to humans, defines human nature. Uncovering its origins is intriguing. But as neither spoken nor signed language fossilize it is very difficult to investigate the evolution of language in humans directly (Fitch, 2010; Hauser et al., 2002). So, scientists – biolinguists to be precise – successfully use animal models to investigate certain aspects of human language. They are asking questions on common features of the acquisition of the vocal repertoire during ontogeny, which are for example indeed shared in huge parts for humans and bat pups (Fernandez, 2020; Knörnschild et al., 2006). Closely connected are questions on learning, to be more precise on vocal learning, which not many animal clades are capable of (Martins & Boeckx, 2020). Next to humans, we know of this ability from birds (songbirds: (Hultsch & Todt, 1989; Tchernichovski et al., 2001; Thorpe, 1958; Wilbrecht & Nottebohm, 2003), parrots: (Pepperberg, 2010) and hummingbirds: (Araya-Salas & Wright, 2013; Gaunt et al., 1994)), and some clades within mammals, namely: cetaceans (Janik, 2014), pinnipeds (Hiss, 1983; Janik, 2014; Reichmuth & Casey, 2014; Stansbury & Janik, 2019), bats (Knörnschild, 2014; Vernes & Wilkinson, 2019) and elephants (Poole et al., 2005).

Especially bird song is often compared to human music. And while the evolution of music is as difficult to retrace as the evolution of language (i.e., does not fossilize either (Honing et al., 2015), some aspects of musicality can again be very well investigated in different animal models. Musicality could be explained as the biological underpinnings of music. It comprises a couple of spontaneously occurring biological traits, constrained by the biology and cognition of a species (Ravignani et al., 2018; Wallin, 1991). Important musicality traits are for example pitch, sonic qualities called timbre, or rhythm (Honing et al., 2015). Focusing a little more on rhythm, we find that it probably has multiple evolutionary

backgrounds (Kotz et al., 2018). A widely accepted definition of rhythm is that it is a 'systematic patterning of sound in terms of timing accent and grouping' (Patel, 2008). In the context of animal's acoustic signal sequences, the rhythm could also generally be described as the temporal structure of these sequences. These temporal structures can be very stereotyped in a particular species and the same rhythmic patterning might be present in different tempi or beats, which would be referred to as meter in musicology.

The research on temporal structures of animals' acoustic signal sequences as well as research on their perception is interesting for both fields: biomusicology and biolinguistics. They are especially interesting and relevant because flawed temporal perception in humans is for example linked to stuttering (Wieland et al., 2015) while an intact auditory system is tuned to certain musical features such as pitch and rhythm already in infants in the same way it is in adults (Stefanics et al., 2009; Winkler et al., 2009). Studies in animals show that it is possible to induce impairments in vocal production on the temporal level. Mimicking genetic errors in model species that are responsible for heavy speech and motor impairments in affected humans can degrade vocalizations on the temporal scale (Norton et al., 2019). These and other observations suggest that the correct production and perception of temporal structures is an essential component of acoustic communication throughout the animal kingdom including humans.

Next to trying to understand the implications of impaired temporal sound production and perception, there are other reasons while it is interesting to study and understand the rhythms of animals' acoustic communication. Analysing the temporal structure of animals' acoustic signal is interesting with regards to questions about species discrimination, physiological correlates like couplings to movements or respiration, mating preferences, or arousal coding (Burchardt et al., 2019; David et al., 2003; Manser, 2001; McRae, 2020; Norton & Scharff, 2016). Other questions include questions on entangling underlying processes in duetting or the development of temporal structures during ontogeny (Pika et al., 2018; Sasahara et al., 2015; Yoshida & Okanoya, 2005). The temporal structure of choruses, duetting, or antiphonies and how they are maintained is yet enigmatic in most regards, and these group behaviours are argued to be a key element in understanding the adaptive functions of rhythm (Ravignani et al., 2014).

To be able to answer such questions clear methodologies are necessary, enabling many different researchers to a) analyse the temporal structures of their acoustic data in a meaningful and comparable way and b) study how their species of interest perceive tempos and rhythms. The following two sections will give more details about examples of rhythm production in animals, explain the key methods used in this thesis and some other commonly used approaches to analyse temporal structures. The same will be done for examples of rhythm perception, explanations again focusing on approaches used in this thesis.

Rhythm Production

Having established the importance of rhythm in animal communication, we will have a look at methods to assess these, show examples and report what is known about biological correlates and the relevance of specific rhythms.

Keeping the definition in mind that rhythm is the 'systematic patterning of sound in terms of timing, accent and grouping' (Patel, 2008), what options does that open for the analysis of rhythms?

Firstly, we can describe the timing of sounds in general and in various ways, many of which are detailed in the following paragraphs. Secondly, it opens options to investigate the temporal structure on different levels: for example, in pulsatile sounds, we can investigate the rhythm *within* an element (Figure 1C), then we can analyse the rhythm *between* elements (Figure 1B) and finally between sequences, thereby considering different possible groupings of elements within one sequence, or the grouping of sequences in a bigger excerpt of communication. A third path with interesting implications could be which part within an element is accentuated when, and how elements in a sequence or sequences in a communication attempt are accentuated. So far different methods exist to tackle some of these options, but not for all.

Temporal parameters have long been investigated, even the term "rhythm analysis" in the context of acoustic communication in animals was coined already by 1974 (Ishay et al., 1974). Nevertheless, especially the comparison of results of rhythm analysis or generally comparison of results of the analysis of temporal structures in animals' acoustic signals is often difficult as many methods exist to analyse temporal structures, but seldomly the same parameters are reported. The only parameters that are widely

being used and reported are call rates[1] or averaged inter-onset-intervals[2] (IOI, Figure 1D) (Bøttcher et al., 2018; Le Bot et al., 2013; Manser, 2001; Moss et al., 2006; Ravignani, 2018; Schneider & Mercado, 2018b; Sun & Narins, 2005). These themselves are problematic though, because a complex phenomenon is tried to be described with a single number, which works well in many cases, but not in all and without other numbers or the original data being reported and published, it is almost impossible to later determine whether it was fair to describe a temporal structure with just one number. This is especially the case in studies where temporal parameters are but a by-product of other analyses and are not the focus of a study themselves. A directly related issue concerns the reporting of variability parameters in the context of call rates or other measures. Here we have the problem that as a variability indicator most often the standard deviation is given. The standard deviation is dependent on the mean, and therefore not easily comparable between studies. A second parameter describing the variability that is sometimes given is the variance, which is both dependent on the mean and the number of observations, which makes it even less suitable to be compared.

[1] number of produced elements per minute
[2] duration between the start of one element and the start of the next element, mostly reported as the average IOI of a dataset/ sequence type.

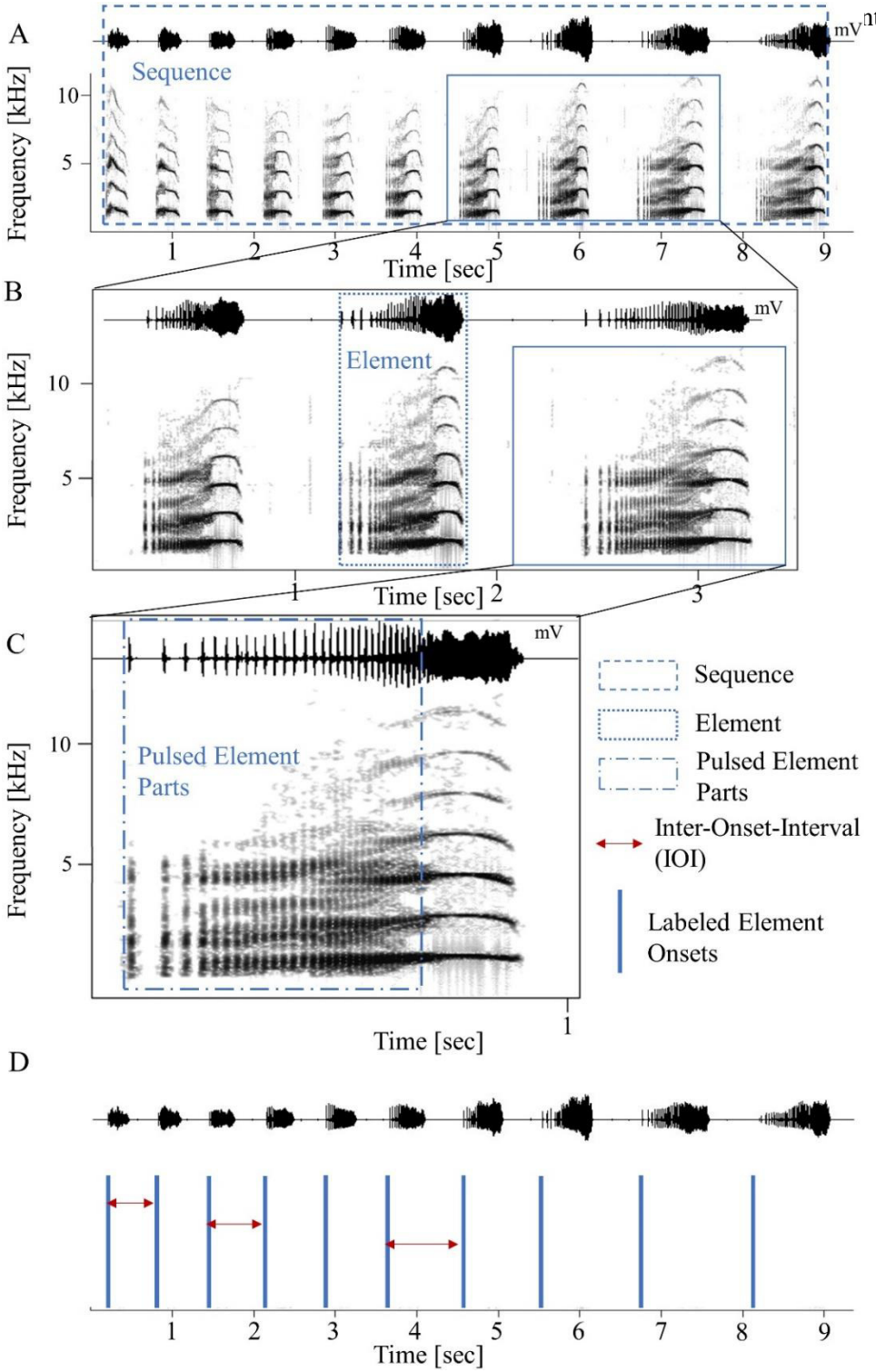


Figure 1: Analysis Levels for Rhythm Analysis
Basis for this figure is a territorial song of a male Saccopteryx bilineata individuum (Behr et al., 2006), that is shown in different amplification to explain some terminology and to demonstrate possible levels of analysis. (A) Oscillogram and spectrogram (FFT length: 1024, Hamming window, overlap 50%) of a territorial song sequence. (B) Amplification of the same territorial song, zooming in on three elements (FFT length: 2024, Hamming window, overlap: 75%). A rhythm could be analysed for elements in a sequence. (C) Detailed view of a single element from the territorial song to give an example for pulsed element parts (FFT length: 2024, Hamming window, overlap: 83.5%). Rhythms could also be analysed within such pulsed elements (i.e., (Bolgan et al., 2020; Picciulin et al., 2020). (D) The oscillogram of the territorial song was used to determine element onsets (blue bars), which are used to calculate Inter-Onset-Intervals (IOI, red arrows), i.e. the duration between the start of one element and the subsequent element.

Only in recent years more and more emphasis was put on finding clear and comparable parameters and methodologies to analyse the rhythm of animals' acoustic signals. This in term facilitates the use of rhythm analysis to distinguish between individuals or species and to use the results for cross-species comparison. The cross-species approach is highly valued in the field of rhythm analysis. Once there is enough knowledge about rhythm production in animals, (for example about specific rhythms being produced) and this information is available for a high number of species this can give rise to even more questions and open more fields of investigation. One example is to specifically explore the evolution of certain rhythmic patterns or properties, as it was recently done in woodpeckers (Garcia et al., 2020). Once clear methodologies are established which produce easily comparable results the ontogeny of rhythmicity in individuals or species could be investigated (Sasahara et al., 2015), or different correlates for the found rhythmic patterns could be proposed.

Rhythm Analysis

Important questions that rhythm analysis methods need to be able to answer include a) what the temporal structure of an element sequence is, including the question of which exact rhythms could describe a sequence best b) whether rhythmic patterns are similar or different between sequence types or individuals and c) how well an animal can produce or keep a certain rhythm.

The latest and most comprehensive overview paper on rhythm analysis methodologies was written by Ravignani and Norton (2017) and included methods to assess the temporal structure of sound sequences on different levels. A key question in the analysis of temporal structures always is whether a sequence can be described by an isochronous rhythm or whether it needs to be described by a heterochronous rhythm which has strong implications on the applicability of most methods used to date. An isochronous rhythm is a very simple, metronome-like rhythm, characterized by similar intervals between element onsets. Isochronous sequences can therefore be described by a single frequency in Hertz as in beats per second. A heterochronous sequence cannot be described by a single frequency. The focus of this thesis lies on analysis methods to quantify beats in isochronous sequences. Here we must make the important distinction between two scenarios. We can imagine a sequence where the signal itself is isochronous, like for example whale echolocation sequences ("signal isochrony" hereafter). On the other hand, we

could have sequences that are well described by an underlying isochronous beat, but there are "silent beats" of the theoretical underlying beat that are not accompanied by an element in the element sequence. The perception of isochrony is induced even though it is not there. This so-called "beat induction" is argued to be a fundamental musical trait in humans (Honing, 2012; Honing & Ploeger, 2012) and an important notion to keep in mind when analysing animals' acoustic signals ("induced isochrony" hereafter, (Ravignani, in press)). Examples for induced isochrony were found in male zebra finches' song or isolation calls of the greater sac-winged bat (Burchardt et al., 2019; Norton & Scharff, 2016). The methods to analyse the two different situations are the same, but some methods will fit better with sequences showing signal isochrony, others will describe sequences with induced isochrony better. The distinction is also important for interpreting results correctly.

One method to assess whether we can assume isochrony or not is the visualisation of the data in histograms. By plotting the distribution of so-called Inter-Onset-Intervals we can determine the general temporal structure of an element sequence. A unimodal distribution of IOIs translates to a small variability between IOIs, the steeper the distribution the more similar the IOIs in a distribution are. The more similar the intervals, the better are they described by an isochronous rhythm (Ravignani & Norton, 2017). Histograms are the perfect first analysis step, as they are a very good compromise between simplicity, effectiveness and most importantly applicability for data (Ioannidis, 2003). Another parameter to assess isochrony is the normalized pairwise variability index (nPVI). It was originally invented to assess temporal variability in human language (Grabe & Low, 2002a; Lin & Wang, 2007; Toussaint, 2012, 2013). It is a measure of variability between IOIs in a sequence. A small nPVI indicates little variation and a nPVI of 0 would mean that all IOIs are the same. In a sequence with induced isochrony, the nPVI might be quite high depending on the number of silent beats (or breaks). They would lie between 0 and ~50 as for theoretical clapping rhythms depending on the number of breaks, i.e., silent beats (Cameron et al., 2019; Duffy & Pearce, 2018). A high nPVI value is therefore not necessarily indicative of a random patterning, but signal isochrony can be excluded for higher values. Once isochrony is determined, it can be interesting to calculate exact beat frequencies that describe a single element sequence best. For example, an element sequence with one sound element roughly every 20 milliseconds could be well described by a frequency of 50 Hz (as 1/ 0.02 s = 50 Hz). Three different

methods exist to determine such exact beat frequencies: Fourier analysis (FFT), a generate-and-test approach (GAT) and IOI calculations. What was so far unclear was what kind of biological data could best be analysed with each of the three methods.

Fourier analysis is a widely used technique to decompose any continuous signal into its sinus components. It dates back to Joseph Fourier's studies on heat transfer in which he showed that these studies could be simplified by representing certain functions as the sum of a number of trigonometric functions (Fourier, 1822). Fourier analysis has been intensively used in ornithology and acoustics research to visualise sound since the late 1950's via so-called spectrograms, which were only developed in the 1940's by researchers at the Bell Telephone Laboratories (Marler & Slabbekoorn, 2004; Potter et al., 1947). In a spectrogram the frequency components of a sound are visualized across time. In the 1980's Fourier analysis was then firstly used, to not only assess the spectral parameters but also the temporal parameters of sounds (Dabbs, 1983). The methodology that is used nowadays to assess the rhythm of animals' acoustic signals with Fourier analysis is derived from a study on the development of zebra finch song (Saar & Mitra, 2008). For that approach the temporal information of a signal sequence has to be transformed into a binary sequence. To that end any event of interest, for example the start of elements, or the start of sequences are labelled as '1' while everything else is labelled as '0'. The sampling rate has to be considered. If two elements are separated by 1 second and we are measuring with a sampling rate of 1000, that would mean we have a thousand samples per second, so a '1' for the element onset would be followed by 999 zeros, to be followed by the second element onset encoded as '1' after 1 second. Afterwards a fast Fourier transformation (FFT), a discrete version of the Fourier analysis, which is much less time-consuming to be calculated, is calculated over the binary sequence (Saar & Mitra, 2008). As a Fourier analysis can decompose a signal into the sinus components of up to half of the sampling rate, in our example, we would find frequencies of up to 500 Hz describing the sequence. We end up with a power spectrogram, where the frequency with the highest peak would be the frequency describing an acoustic signal sequence best. Transferring this to the normal usage to analyse pitch components, this frequency would correspond to the loudest frequency. There was so far no goodness-of-fit value for this type of analysis.

A more intuitive way of finding a concrete beat frequency to describe an isochronous element sequence was developed in a GAT approach (Norton & Scharff, 2016). An element sequence consisting of only the temporal information of the sequence (start points, duration and endpoints of elements) is overlaid with beat frequencies in a certain frequency window with a certain frequency resolution (i.e., 10 to 100 Hz with a resolution auf 0.1 Hz). For every beat frequency the deviation from an element to the next closes beat is calculated and the root-mean-square deviation of all the deviations in a sequence is calculated and normalized for the frequency being tested. The resulting value, the frequency normalized root-mean-square deviation, short FRMSD, acts as the estimator for the best-fitting beat frequency (i.e., the frequency resulting in the smallest FRMSD) and simultaneously serves as a goodness-of-fit value, describing how good the fit is (Norton & Scharff, 2016; Ravignani & Norton, 2017).

A third method used in this thesis that was not described before[3] to calculates exact beat frequencies is merely a different form of reporting call rates or IOIs. For this form of IOI analysis (called IOI approach in the following), the inverse of the average IOI of a sequence is taken. As Hertz is 1/second, this gives us a frequency and therefore a value easily comparable to the two methods described above. A frequency as calculated with the IOI approach of 20 Hz would for example correspond to an average IOI duration of 50 ms (1/ 0.05s = 20 1/s). It is the fastest of the three methods and the only one not requiring any type of analysis code, but merely a calculator or even just mental arithmetic. As the method was not used before, there also was no suitable goodness-of-fit value so far.

These three methods are searching for isochronous beats, assuming an underlying isochronous patterning. It could therefore be argued that they are describing periodicity and different tempi rather than different rhythms. To consider this issue, the term “beat” is used to describe the different isochronous patterns throughout my thesis.

Apart from methods focusing on isochrony and describing exact beats, other scattered methods exist to analyse more complex rhythms. These methods mostly only extend to some specific question for example how to assess the similarity between two temporal structures, no matter how complex. This

[3] Briefly described here for a better overview of the methods used throughout this thesis, the method is established in Chapter II.

similarity could be analysed with cross-correlation methods, which can be described as a similarity measure of two sequences as a function of the displacement of one sequence relative to the other or even more generally as the similarity between two observations, where the similarity is described as a function of the time lag between the two observations (Hamilton, 1994a, 1994b). The approach was proposed for the use in rhythm analysis in animal communication by Ravignani and Norton (2017) but to the best of my knowledge has not yet been used in the described way to assess temporal pattern similarity between different sequences of animals' acoustic signals. Cross-correlation of temporal parameters was successfully being used in musicology studies (Percival & Tzanetakis, 2014) or time series analysis of human behaviour (Boker et al., 2002), though. Furthermore, cross-correlation has extensively been used in the form of spectrographic cross-correlation, where similarity between sequences or elements is assessed in the spectral rather than the temporal domain (i.e.,(Cortopassi & Bradbury, 2000; Khanna et al., 1997; Terry et al., 2001; Zsebők et al., 2021)). Complementary to cross-correlation approaches, auto-correlation is being used to assess the regularity within a single sequence, without defining a concrete beat or assuming an underlying isochronous pattern (Heinsohn et al., 2017a; Le Bot et al., 2013; Ravignani & Norton, 2017). Another option to analyse multiple parameters of a complex temporal pattern is the so-called multifractal analysis (Roeske et al., 2018). It can detect similarities on various levels and across different scales, through searching for self-similarity of a signal. Therefore, it could be said it detects predictability (as indicated through self-similarity) on different levels and consequently also unpredictable (unsimilar) states of a continuous signal. The method is a very powerful tool in the analysis of complex acoustic signals (Roeske et al., 2018).

Examples of Rhythm Analysis in Action

The following paragraphs are giving an overview about studies on rhythm analysis, dividing examples by the mechanism of sound production. Examples will be given for the sound production in tetrapod's (including amphibia and amniotes) as well as different examples in fish and insects and the use of objects (i.e., tools) for sound production in some birds and mammals. The focus lies on the analysis of isochronous patterns.

The sound production in tetrapods happens internally and is usually achieved by the vibration of soft tissue. This vibration is elicited by air flow and the final sound subsequently modulated through filtration and modulation in the vocal tract (Bradbury & Vehrencamp, 2011; Fitch, 2000). The following examples showcase rhythm in sounds produced in that way.

The song of zebra finch males has been intensively studied with regards to questions on vocal learning (i.e., (Slater et al., 1988; Tchernichovski et al., 2001)), the neural foundations of song (i.e., (Bolhuis et al., 2010; Scharff & Nottebohm, 1991; Simpson & Vicario, 1990)) or female preferences (i.e., (Riebel, 2009)). In a recent study the rhythm of male zebra finch's undirected and directed song was investigated. Individual males had a distinct isochronous rhythm which fitted element onsets better than expected by chance. Individual males produced distinct rhythms that ranged from 10 to 60 Hz, with no apparent differences between undirected and directed song. The rhythms were analysed with the custom-made GAT approach and the goodness-of-fit tested with the FRMSD parameter as well as with a modelling approach, to compare the goodness-of-fit between song sequences in the correct order and the goodness-of-fit of the best fitting rhythm in randomly combined sequence of song elements. The clear result was that the calculated rhythms fitted significantly better to actual song sequences than to the randomly combined song sequences (Norton & Scharff, 2016).

A different approach was used to quantify the development of rhythm in juvenile zebra finch song. First, a point process was calculated: the song sequence was transformed into a binary sequence, where element onsets were encoded as '1' and everything else as '0'. A fast Fourier transformation was calculated on this binary sequence, resulting in a "rhythm spectrogram". It was found that the fundamental frequency in this "rhythm spectrogram" corresponds to the motif duration, a motif being a repeated combination of a few different element types in the zebra finches' song. It is suggested that this method can be used to track the transition in song development from juvenile rhythms to adult rhythms, corresponding to changes in the motif composition and the crystallization of element types (Saar & Mitra, 2008).

In a purely descriptive study, the Inter-Onset-Intervals of a spontaneously vocalizing female harbour seal pup were reported. The clear aim of this study was to be able to balance the ecological relevance of

stimuli with experimental control in a follow-up playback experiment with the same female. The stimuli for the playback experiment, more precisely the rhythms of the stimuli, could therefore be individually adapted (Ravignani, 2018). This study gives a strong example for the relevance of knowledge about temporal structures of animal's acoustic sequences even when the temporal structure is not itself being studied.

A very well-known example using call rates is the study on the alarm calls of the social meerkat *Suricata suricatta*. In that species different call rates in an alarm call context encode urgency. Alarm calls with a low urgency to all predator types, aerial or terrestrial, are produced with relatively long IOIs. For alarm calls in response to terrestrial predators, the call rate then for example increased with urgency (Manser, 2001).

The change of call rates is not only indicative of urgency but can also be found in anurans as a response to anthropogenic noise. In response to airplane flyby noise and low-frequency motorcycle sound playbacks different pond-edge frog species in central Thailand changed their call rates. While three of the present species significantly decreased their call rates, one species significantly increased their calling rate (Sun & Narins, 2005). This study is giving a good example for the potential of rhythm analysis to track behavioural changes. Changes can occur in the temporal domain of acoustic signals, not only in response to anthropogenic influences, but might also happen in response to other species or naturally occurring changes in the environment.

Sound production mechanisms are very diverse in fish (Fine & Parmentier, 2015; Ladich, 2014) and in many cases the temporal structure within a sound element carries meaning, showcasing another example of the importance of the analysis of temporal structures in animal's acoustic signals on various levels. Many vocal fish species produce highly stereotyped, pulsatile sounds, generated by rhythmic contractions of muscles, acting under a rhythmically active vocal motor network (Bass & McKibben, 2003). As such, the temporal structure *within* fish sound elements (for another example see Figure 1C) can be indicative of neurophysiological processes (Fine & Parmentier, 2015; Sprague, 2000). The measurement and the analysis of temporal features within single fish sound elements are routinely included in fish bioacoustics studies and are mostly reporting rates. These rates were found to be able

to code for species identity (Parmentier et al., 2009) as well as for specific behavioural patterns or motivational states (Bolgan et al., 2020; Hawkins & Amorim, 2000; Picciulin et al., 2020). Furthermore, within-sound temporal features can be influenced by environmental conditions such as water temperature, as this can impact the performance of the morpho-physiological processes underlying sound production (Kever et al., 2015; Ladich, 2018). They can also be affected by anthropogenic disturbances (Picciulin et al., 2012) same as reported above for anurans (Sun & Narins, 2005).

Yet another sound production mechanism, found in many insect species, is stridulation. Here, sounds are produced by rubbing two movable parts of the cuticle or other bony body parts against each other. The rhythm of the produced sounds can for example be indicative of the communicating species: the qualitative assessment of rhythms in *Gryllus* specimen led to the discovery of two cryptic species of *Gryllus assimilis*, following the concept proposed for cryptic species in 1964 (Walker, 1964). Way earlier, already in 1932, a key for species identification for Orthopteras was written, only using sound features as identifying characteristics including the rhythmic (or arhythmic) structure of sounds (Fulton, 1932). In element sequences produced through stridulation again the temperature has an important effect on the exact rhythms found. In a study on *Gryllus bimaculates* the repetition rates of different element types increased linearly with increasing temperatures between 15 and 24°C (higher temperatures had no effect). The phenomenon is linked to phonotaxis in females of the respective species. Females responded best to synthetic songs matching the rhythmic properties of the male's calling at the temperature the female was tested in (15, 24, or 30 °C). This suggests an interdependence between the physiological mechanisms underlying pattern generation and pattern recognition in crickets (Doherty, 1985).

To further illustrate in how many examples and on how many levels the temporal structure of sounds plays an important role, we can look at the rhythm of sounds produced by tools or external objects. A group of chimpanzees (*Pan troglodytes*) in Fongoli, Senegal, was recorded when cracking baobab fruit. Fruits are cracked by repeatedly slapping them on a hard surface like the ground or a tree branch. Four females and ten males that were recorded showed significant differences in the slapping intervals. Individual differences are hypothesized to facilitate the recognition of unseen companions (Merguerditchian et al., 2018).

A second animal species using tree parts to drum is the palm cockatoo *Probosciger aterrimus*. It is reported that the drumming shares key fragments of human instrumental music. The reported fragments are: manufacturing a tool for sound production, performing the drumming in a consistent context, individual styles as well as the regular production of beats and the repetition of components. Used for the analysis of the rhythmic patterns are shape parameters of the IOI distribution and autocorrelation (Heinsohn et al., 2017a). This study highlights the potential of rhythm analysis for biomusicology related questions.

An already mentioned example of yet another "drumming" example is the analysis of pecking patterns in various woodpecker species. Here the rhythmic patterns were described with 22 parameters including 10 parameters describing the temporal structure for example the "species-specific median drum pulse rate", the "minimum time interval between 2 pulses" and the "maximum time interval between 2 pulses" or the "Interval nPVI". The parameters were used to define drumming types and eventually to reconstruct the evolution of drumming signals in woodpeckers (Garcia et al., 2020).

One important aspect that has almost always been neglected up to this point is the question of how well an animal keeps a beat. Certain variability parameters (such as low standard deviation and variance) are used to argue for a very consistent and therefore rhythmic sound production, but the quantitative evidence (for example shown through goodness-of-fit values and their comparison between patterns) for that is frankly just missing. The only study introduced here that specifically asked and answered this question was the study by Norton & Scharff on isochronous song production in male zebra finches (Norton & Scharff, 2016). One reason for this lack of information on the goodness-of-fit is a lack of methods to specifically analyse it for different methods. For two of the three methods to extract exact beat frequencies, we do not have reasonable suggestions on how to determine the goodness-of-fit, that is for Fourier analysis and the IOI approach. Only for the GAT approach, the method-specific value FRMSD was introduced (Norton & Scharff, 2016) so far. I attempt to close these gaps in this thesis.

All examples showcased here describe individual rhythms, as these are also under study in this thesis. Nevertheless, it is to be mentioned that the individual level, though interesting and fundamental for the general understanding of the temporal structures of animals' acoustic signals, is not the only level being

analysed. Rhythms play a fundamental role in turn-taking situations, such as in chorus or duetting situations (Pika et al., 2018; Yoshida & Okanoya, 2005) as well. It is argued that especially individual rhythms produced in such group behaviour settings need to be understood better. Group behaviour settings focusing on mammals and birds that were under investigation so far include, for example, turn-taking and duetting in great apes and monkeys (Pougnault et al., 2020; Takahashi et al., 2013), sperm whales (Schulz et al., 2008), vampire bats (Carter et al., 2008) or birds (Dahlin & Benedict, 2014; Gochfeld, 1978; Hultsch & Todt, 1982; Thorpe, 1975). As it is argued, that in these contexts individual rhythms are shaped by group dynamics, they are argued to being key to understand the adaptive functions of rhythms (Ravignani et al., 2014).

The question of possible adaptive functions and why signals are produced in a rhymical fashion is discussed in many of the mentioned studies. Next to being shaped by group behaviour settings, another hypothesis exists on how and why rhythmic production of signals is frequently found. For example, different motor correlates could be assumed to shape rhythms, i.e., depending on different sound production mechanisms. A known correlate between a specialized form of sound production and motor patterns is found in bats. The production of echolocation calls is coupled to wingbeat and the respiratory cycle in many bat species (Kalko, 1994; Schnitzler, 1971; Suthers et al., 1972; Wong & Waters, 2001). During flight, the respiratory cycle follows the wingbeat, as each wing stroke passively ventilates the lungs. To not break this energy-efficient coupling, echolocation call production is also matched to this pattern. Inspiration corresponds to the upward wing stroke, while the expiration phase correlates to the downward stroke of the wing (Freiherr von Saalfeld, 1939). Echolocation calls are then produced around the reversal point between ex- and inspiration. As production rhythms of echolocation calls need to be very flexible this coupling between wingbeat and echolocation call production nevertheless only holds true for situations where the sensory needs allow it. The pattern is for example broken for the feeding buzz or in environments with a lot of clutter, in both of which situations echolocation call rates can increase dramatically (Moss et al., 2006; Moss & Surlykke, 2001; Ratcliffe et al., 2013). A coupling between wingbeat and sound production was also found in the tongue-clicking pteropodid bat *Rousettus aegyptiacus*, indicating that strong coupling of wing beat, respiration, and sonar emission is widespread in bats regardless of the specific sound production mechanism. The question remains what rhythms

would be found in situations where bats are perched and coupling between wingbeat and sound production would not be expected.

Another important aspect to consider to be able to answer the "why" is the perception of the acoustic signals, which is also relevant in turn-taking situations and is furthermore connected to the fact, that rhythmic signals might be easier perceived (i.e., (Jones et al., 1981; Norton & Scharff, 2016). This point will be detailed in the next section on 'Rhythm perception'.

Rhythm Perception

The reasons to produce rhythms of a certain beat could be connected to neural or motor correlates, as discussed above. An important question that remains for most examples is whether these production rhythms have an ecological meaning, whether they carry meaning themselves and how they are perceived. Some of the studies presenting different production rhythms already included a biological reasoning, a connection to the perception of signals (Doherty, 1985), or were meant as preparatory work to set up behavioural experiments in a more informed way (Ravignani, 2018). But what else do we know about the perception of rhythms and how could we further analyse it?

One of the hypotheses why signals are produced in a rhythmic fashion to begin with is that rhythmic signals are easier to perceive and that rhythmic signals facilitate signal perception. The concept of 'rhythmic attention' (sensu (Jones et al., 1981)) is important to mention in that context. This 'rhythmic attention' should help receivers to decode periodic signals not only easier but also faster (Rohenkohl et al., 2012). A rhythmic signal is better predictable than an arhythmic signal produced in a random temporal succession. To further facilitate this effect, it is to be noted that the attention of receivers cycles in an oscillatory way, but only when a rhythm is present (i.e., (Barnes & Jones, 2000; Large & Jones, 1999)). That way, most 'attentional energy' can be provided, when a stimulus is expected. As the cognitive capacities of any species are limited, (Shapiro et al., 1997), the optimized allocation of attentional energy is not only beneficial but necessary. Concrete examples for that can be found in humans. When asked to assess the differences in the pitch of two focal tones, separated by regularly timed tones, individuals could assess the difference better when the second focal tone followed the same regular timing as the separating tones in between. The assessment of differences was worse when the

second focal tone was temporally slightly displaced from the regular timing (Jones et al., 2002). A similar phenomenon could be found for visual stimuli in macaques. There neuronal oscillations in the primary visual cortex entrain to a visual stimuli stream when the stream is rhythmic, a mechanism resulting in decreased reaction time and an increase in the response gain for events that are task-relevant (Lakatos et al., 2008).

Getting back to concrete examples of the assessment of rhythm perception in the acoustic communication in animals on a behavioural level: the first instance of biologically relevant rhythm perception in a non-human mammal was found in the northern elephant seal. Males of that species can discriminate the sounds of familiar and unfamiliar male opponents. A study found that this discrimination ability is driven strongly by the temporal structure combined with the timbre of the sounds (Mathevon et al., 2017).

The question remains whether this discrimination ability with regards to temporal structures can be generalized to other contexts in northern elephant seals, or whether the perception of differences is limited to the species-specific context. The overall question is whether this or other examples where production rhythms play a role imply an overall sensitivity for rhythmic patterns or not (ten Cate et al., 2016). Another example where changed rhythms could alter the behavioural response, was the "cooing" of doves but it was unclear whether this ability could be generalized (Slabbekoorn & ten Cate, 1999). Especially in birds, various experiments were run to test the general ability to discriminate between temporal patterns and how well discrimination could be generalized. Even though zebra finches for example can distinguish between isochronous and irregular stimuli, the ability for generalization between different tempos of isochronous patterns is not given. It is suggested that zebra finches distinguish the isochronous and irregular stimuli using local temporal features rather than the regularity of the stimulus in contrast to the irregularity in the other stimuli as the discriminator, which in turn would enable them to generalize between tempos (van der Aa et al., 2015). Such a generalization was found in the European Starling (*Sturnus vulgaris*) though. Individuals learned to discriminate between regular and irregular temporal sequences, and the discrimination was maintained for different modifications of the temporal structure, like changing either the tone duration or the IOI duration while keeping the other parameter constant. Overall, this suggested the discrimination of rhythmic versus arhythmic patterns in

the European Starling is based on a qualitative pattern attribute, namely the rhythmicity, rather than being based only on local temporal features (Hulse et al., 1984a, 1984b; Humpal & Cynx, 1984). The same ability was found in jackdaws (*Corvus monedula*), which also showed evidence for using the relative temporal structure to discriminate between sound sequences (Reinert, 1965).

The clear need to test more species with various patterns to reveal underlying mechanisms and patterns of how birds perceive temporal structures was expressed together with the suggestion that current evidence could hint at an underlying graded scale for the sensitivity of regularity (ten Cate & Spierings, 2019). This graded scale (limited to birds) would then encompass pigeons on one end, which only showed very little abilities to discriminate between temporal patterns (Hagmann & Cook, 2010), and parrot species, that are even able to use a perceived rhythm to induce movements synchronized to a given beat[4], on the other end (Patel et al., 2009a, 2009b; Schachner et al., 2009).

Moving from behavioural studies to rhythm perception and signal detection thresholds on a neuronal level, the focus shifts to studying the processes underlying acoustic perception, auditory processing, or the involvement of specific brain areas in these processes. Ultimately many of these studies also investigate how acoustic information is encoded and decoded (e.g., (García-Rosales, Beetz, et al., 2018; Land et al., 2016; Linnenschmidt & Wiegrebe, 2019; Portfors & Wenstrup, 1999; Wetekam et al., 2020).

A very common method in the field of analysing signal detection thresholds are auditory brainstem response experiments (ABRs). They are a powerful method to measure hearing thresholds and define audiograms and were developed in the 1970s by Jewett and Williston. In an ABR the neuronal activity in the brainstem part of the auditory pathway in response to an acoustic stimulus is measured as field potentials (Jewett & Williston, 1971). ABRs are intensively used in animal models as well as humans, to assess hearing statuses, determine sensitivity thresholds, or recognize mental disorders (Hecox & Galambos, 1974; Juselius Baghdassarian et al., 2018; Källstrand et al., 2014; Land et al., 2016; Obrist & Wenstrup, 1998; Szymanski et al., 1999). The auditory pathway includes five brainstem nuclei: the cochlear nucleus, the superior olivary complex, the nucleus of the lateral lemniscus, the inferior

[4] Entrainment: like the human capacity to dance in tune to a specific rhythm. More generally: the ability to synchronize movements to an external rhythmic stimulus.

colliculus, and the medial geniculate (Claesdotter-Hybbinette et al., 2016; Purves et al., 2012). While in a classical ABR experiment the sensitivity in response to differently pitched stimuli is measured, different reasons hint at the fact that ABRs can be adjusted to measure sensitivity with regards to beat or tempo perception as well. For example, one of the auditory nuclei, the Inferior Colliculus (IC) shows a very high temporal precision in electrophysiological recordings of Seba's short-tailed bat *Carollia perspicillata* (Phyllostomidae) (Macias et al., 2016), which leads to the assumption that temporal perception is in part already happening in the auditory brainstem. Furthermore, specialized neurons in the IC of the mustached bat, *Pteronotus parnellii* (Mormoopidae), facilitate response strength by echo delay-tuning (Portfors & Wenstrup, 1999). Moreover, it is known that and how the IC contributes to an ABR signal in a small mammal (Land et al., 2016). ABR procedures where stimulus presentation rates, and therefore the temporal structure of the stimuli, were changed instead of the pitch frequency of stimuli, were intensively used to study adaptation processes in the auditory brainstem of various animals, such as cats, gerbils, mice, chicken, humans, and echolocating bottle-nosed and common dolphins before (Burkard et al., 1994; Burkard & Sims, 2002; Burkard et al., 1997; Burkard et al., 2017; Burkard et al., 1996a; Burkard et al., 1996b; Burkard & Voigt, 1989; Ridgway et al., 1981). This could be an interesting way to study beat perception as well.

By adjusting an ABR procedure to measure beat perception, the advantages of the ABR method compared to behavioural tests could be utilized. In an ABR experiment, it is possible to directly measure the spontaneous reaction to a specific easily interchangeable stimulus of an animal on an individual level, while most parameters can be controlled for. It is a minimally invasive technique and there is no need for unnecessary stress due to long periods of isolation and training for the animals. On the other hand, behavioural experiments are likely to produce more sensitive results (Heffner et al., 2008). In contrast to the ABR approach, which of course neglects the importance of attention for stimulus perception as animals are anesthetized for the procedure, to exclude to measure brainstem activity caused by movement, sight, smell, or other unwanted processes, behavioural tests do include the attention of an individual and can indicate whether a stimulus or the discrimination between different stimuli is relevant for the animal. An ideal scenario could use preliminary ABR experiments to be able to set up a subsequent behavioural test in a more informed way, to minimize the number of animals

needed or the time spent training an animal, because certain stimuli could already be excluded in the behavioural setup due to the results of the ABR experiments.

Study animals

Different study species, with very different sound sequence characteristics, were used in this thesis. The study species included bats, cetaceans, and songbirds. The biggest amount of data used were various bat vocalizations from a total of 12 species from six families. Furthermore, echolocation sequences from long deep dives of one female sperm whale *Physeter macrocephalus* were analysed as well as the flight song of 14 different individuals of male skylarks. Species and the corresponding sequence types that were analysed are shown in Table 1.

Different sequences served different purposes: while for some sequences it was the genuine interest to describe their temporal structure, other sequences served the purpose to establish and test new methodological approaches, because of certain properties of the sequences or the sequence type already being intensively studied, so that results could be easily linked to interpretations.

Bats

All 12 bat species for which element sequences were analysed in this thesis occur in Central America, and all element sequences analysed were recorded there. For 10 of 12 species only echolocation call sequences were analysed that were uttered in a search flight context, the type of echolocation that is most often coupled to wingbeat frequencies (Kalko, 1994; Moss et al., 2006; Ratcliffe et al., 2011; Schnitzler, 1971). The species differ in their life history spanning from nectarivorous bats, over frugivorous bats to insectivorous bats, and the blood-eating vampire bat *Desmodus rotundus*, resulting in very different echolocation behaviours, showcasing at least some of the variety found in the echolocation behaviour of the second-largest mammal order. As the temporal information of received echolocation echoes is important for the bats to navigate, they are an interesting candidate taxon to ask questions about both rhythm production and perception. ABRs are known to work very well in bats (Burkard & Moss, 1994; Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019; Obrist & Wenstrup, 1998; Wetekam et al., 2020).

The two species *Saccopteryx bilineata* and *Carollia perspicillata*, for which we analysed more than one sequence type, are introduced in more detail below. *Carollia perspicillata* also had a prominent role in the rhythm ABR experiments that were carried out.

Saccopteryx bilineata

The greater sac-winged bat *Saccopteryx bilineata* has a rich vocal repertoire (Behr & Helversen, 2004) and is capable of vocal production learning (Knörnschild et al., 2010), which makes it an interesting species to investigate in terms of rhythm production, as rhythms can be compared between individuals and the various sequence types as well as in learned and innate contexts. The distinct sequence types are uttered in different contexts. Used were three sequence types: next to the already mentioned search flight echolocation, we also analysed territorial song and isolation calls. All three types are multi-element sequences (i.e., consists of at least three elements in a row) with clear element onsets. Both characteristics are prerequisites to be able to analyse their temporal structure. Isolation calls are produced by pups to solicit maternal care and by adult males to appease dominant conspecifics, the calls are innate (Fernandez & Knörnschild, 2017; Knörnschild, Nagy, et al., 2012; Knörnschild & von Helversen, 2008). An isolation call sequence is built with different element types and can be up to 2 seconds long (Knörnschild, Nagy, et al., 2012; Knörnschild & von Helversen, 2008). Adult males produce territorial songs to attract mating partners and repeal rivals (Behr & Helversen, 2004; Knörnschild et al., 2017). Territorial songs are learned during ontogeny by imitation of conspecific's song (Knörnschild et al., 2010). As well as the isolation calls, echolocation calls are produced by males and females. They are used for orientation, navigation, and insect prey capture (Jung, Kalko, & Helversen, 2007) but also serve a function facilitating social communication among group members (Knörnschild, Jung, et al., 2012). These three sequence types were chosen to get insights into the periodicity of innate signals (echolocation) as well as signals directly produced after birth but being changed during ontogeny through vocal production learning (isolation calls (Knörnschild, Jung, et al.,2012)) and learned sound signals (territorial song), while also being able to investigate potential age differences in periodicity.

Carollia perspicillata

Seba's short-tailed bat *Carollia perspicillata* also produces different sequence types for communication. Analysed for this thesis were the short isolation call sequences, which in contrast to isolation calls of *S. bilineata* consist of only one element type. They are also used to elicit maternal care and carry an individual signature strong enough for mothers to recognize their pup (Knörnschild et al., 2013). *Carollia perspicillata* is also a model organism often used in electrophysiological studies, and ABRs are known to work well in this species. (Hechavarría et al., 2013; Macias et al., 2016; Wetekam et al., 2020)

Sperm Whale *Physeter macrocephalus*

The echolocation call sequences of the sperm whale *Physeter macrocephalus* are very long and stereotyped in both the spectral and the temporal dimension. They have been analysed intensively in many studies and were therefore used to test methods on sequences where we knew quite well what we expected the results to be. The analysed data came from one single individual: the female sperm whale Sophocles, recorded by the Dominican sperm whale project on 24. April 2014 (for details on study site and recordings see (Bøttcher et al., 2018; Tønnesen et al., 2018).

Skylark, *Alauda arvensis*

The third animal group sounds were analysed from were songbirds. After aerial mammals, vocalizing in movement and stationery, and aquatic animals, producing sound also during movements, we decided to use bird song also produced in movement: the flight song of male skylarks (*Alauda arvensis*). Skylark song is very complex: one individual can incorporate more than 300 different syllables in its song which can be combined in various ways, giving rise to a lot of variation also with regards to the temporal structure (Aubin, 1982; E. Briefer et al., 2008; Elodie Briefer et al., 2008). That way these songs were optimal to test and adjust methods after the initial development with sequences we knew what to expect (i.e., sperm whale echolocation sequences). Furthermore, we could test how the methods would cope with more complex temporal structures.

Table 1: Studied Species and Analysed Sequence Types

Animal Group	Family	Species	Sequence types	Sequences (individuals if known)
Bats	Emballonuridae	*Rhynchonycteris naso*	Echolocation call sequences	8 (4)
	Emballonuridae	*Saccopteryx bilineata*	Echolocation call sequences Territorial songs Isolation Calls	33 (33) 142 (14) 500 (25)
	Emballonuridae	*Saccopteryx leptura*	Echolocation call sequences	15 (12)
	Molossidae	*Molossus molossus*	Echolocation call sequences	15 (15)
	Mormoopidae	*Pteronotus parnellii*	Echolocation call sequences	15 (4)
	Phyllostomidae	*Carollia perspicillata*	Echolocation call sequences Isolation calls	10 (4) 47 (5)
	Phyllostomidae	*Desmodus rotundus*	Echolocation call sequences	13 (7)
	Phyllostomidae	*Glossophaga soricina*	Echolocation call sequences	8 (3)
	Phyllostomidae	*Lonchorhina aurita*	Echolocation call sequences	17 (9)
	Phyllostomidae	*Phyllostomus hastatus*	Echolocation call sequences	14 (4)
	Thyropteridae	*Thyroptera tricolour*	Echolocation call sequences	19 (6)
	Vespertilionidae	*Myotis nigricans*	Echolocation call sequences	10 (7)
Birds	Alaudidae	*Alauda arvensis*	Flight song	14 (14)
Toothed Whales	Physeteridae	*Physeter macrocephalus*	Echolocation call sequences	60 (1)

Thesis Aim and Outline

There are some major gaps in the currently established methodological framework for rhythm analysis of animals' acoustic signals. Especially how good certain descriptors are in describing a temporal pattern could not be answered for Fourier analysis or simple rate and frequency calculations using the IOI approach. Comparison or reasonable quantitative assessment about which method might depict a temporal pattern best in any given situation was all but impossible. Furthermore, a clear methodology with exemplary biological data to illustrate possible results and how to interpret them was missing. An aspect that is unanswered for many studies is whether and in what way different beats are perceived and how this reflects biological relevance for the animals.

Reproducibility, interpretation biases, p-hacking (the distortion or manipulation of results through data mining), and apophenia (the tendency to see a pattern in random data) are key issues in all research fields. Defining clear methodologies with open access to code and data is one way of tackling those issues (Munafó et al., 2017) and, therefore, and overarching aim of this thesis. The concrete topics to be addressed in this thesis all aim to further the field of rhythm analysis of animals' acoustic signals with regards to different aspects: 1) the description of temporal structures of animals' acoustic signal sequences, 2) the extension of clear workflows and the establishment of interpretation examples for biological data, 3) the development of parameters to clearly and comparably assess the goodness-of-fit of a given beat frequency (in Hz) for a specific element sequence and 4) the assessment of the perception of different isochronous beats.

The thesis begins with a first study on isochronous rhythms in the bat *Saccopteryx bilineata* (**Chapter I, Publication A: Burchardt et al, 2019**). The main aim here was to describe the temporal patterns of three sequence types (territorial songs, isolation calls, and echolocation call sequences) of this species. With these three sequence types, we tested rhythmicity in learned (territorial songs), sequences being produced from birth but being changed through ontogeny (isolation calls) and innate sequence types (echolocation call sequences) as well as in pups and adults. Sequences were analysed with the GAT approach and the best fitting isochronous beats were reported. All sequence types shared common

isochronous beats of around 6 to 24 Hz, which correlate to the wingbeat frequencies of that species. Acoustic signals were uttered in those beats independent of whether a bat was flying or perched.

After this initial study, it became apparent that a clear workflow taking biological data into account was missing. We set out to establish one using the GAT approach used in Chapter I as well as Fourier analysis and the for the first time applied IOI approach (**Chapter II, Publication B: Burchardt and Knörnschild, 2020**). Furthermore, additional parameters were developed to assess the goodness-of-fit of beats for the Fourier analysis and IOI analyses. All methods were calculated for three different datasets: isolation calls of the bat *Carollia perspicillata*, isolation calls of the bat *Saccopteryx bilineata,* and echolocation call sequences of the toothed whale *Physeter macrocephalus*, all of which showed an isochronous structure of their vocalizations with different consistency. Depending for example on the duration of sequences certain methods prove to be better suited than others, to analyse specific best fitting beats. A clear workflow is provided, stating which analysis path is suited best for which kind of data.

In **Chapter III (Publication C, Burchardt et al, submitted)** the perception of rhythms is in focus. The main question was whether different isochronous beats are differently well perceived by different bat species. One hypothesis was that it might be possible to find a correlation between production rhythms and perception rhythms. The ABR approach was adjusted to measure rhythm perception, by varying stimulus presentation rates instead of stimulus pitch. The procedure was tested on a total of 98 individuals of 12 Central American bat species in the wild and in captivity with artificial and natural stimuli. Different beats indeed elicited differently strong reactions in the auditory brainstem, with slow rhythms (i.e., 6 Hz) consequently eliciting higher reactions than faster rhythms (i.e., 100 Hz). The perception could in parts be correlated to production rhythms. The results of this study are in any case important for future considerations on how to set up ABR experiments in terms of which stimulus type and presentation rates to use.

Open questions remain on actual comparability and on the applicability of aforementioned methods for complex acoustic signals, and how rhythm analysis could further be improved. **Chapter IV (Publication D, Burchardt, Briefer & Knörnschild, in prep.)** deals with these questions and presents

a universal goodness-of-fit value (ugof) and expansions of the Fourier analysis. The ugof value allows to assess the fit of any beat in Hertz to any element sequence and a method is described to calculate a corresponding p-value. For the first time the fit of single elements in a sequence can be analysed, which is an important addition in terms of research on accentuation and grouping and can be interesting in other research fields like heart-beat analysis as well. The addition to established Fourier analysis was the idea to report more than one beat to describe a complex sequence and possibilities for the interpretation of these beats are given. Moreover, a third idea is presented to find grouping in the temporal structure through recurrence plots, which plot a sequence of IOIs as their differences, building a raster showing the differences between every IOI with every n^{th} IOI as colour-coded squares. which provide a visual representation of the sequence's temporal structure, by depicting a sequence as colour-coded squares, each square representing the comparison between an IOI pairing in the sequence. Ideas are tested on datasets from Chapters I-III and a new dataset of complex flight song of the skylark *Alauda arvensis*.

References

Araya-Salas, M., & Wright, T. (2013). Open-ended song learning in a hummingbird. *Biol Lett, 9*(5), 20130625. https://doi.org/10.1098/rsbl.2013.0625

Aubin, T. (1982). Habituation au chant territorial chez l'alouette des champs (*Alauda arvensis* L.). *Biology of Behaviour, 7*, 353-362.

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention and Time. *Cognitive Psychology*(41), 254-311. https://doi.org/10.1006/cogp.2000.0738

Bass, A. H., & McKibben, J. R. (2003). Neural mechanisms and behaviours for acoustic communication in teleost fish. *Progress in Neurobiology, 69*(1), 1-26. https://doi.org/10.1016/S0301-0082(03)00004-2

Behr, O., & Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioural Ecology and Sociobiology*(56), 106-115. 10.1007/s00265-004-0768-7

Behr, O., Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., & Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behavioural Ecology*(17), 810-817. https://doi.org/10.1093/beheco/arl013

Boker, S. M., Xu, M., Rotondo, J. L., & King, K. (2002, Sep). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioural time series. *Psychological Methods, 7*(3), 338-355. https://doi.org/10.1037/1082-989x.7.3.338

Bolgan, M., Amorim, M. C. P., Fonseca, P. J., Di Iorio, L., & Parmentier, E. (2018, Jul 12). Acoustic Complexity of vocal fish communities: a field and controlled validation. *Scientific Reports, 8*(1), 10559. https://doi.org/10.1038/s41598-018-28771-6

Bolgan, M., Crucianelli, A., Mylonas, C. C., Henry, S., Falguière, J. C., & Parmentier, E. (2020). Calling activity and calls' temporal features inform about fish reproductive condition and spawning in three cultured Sciaenidae species. *Aquaculture, 524*, 735243. https://doi.org/10.1016/j.aquaculture.2020.735243

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience, 11*(11), 747-759. https://doi.org/10.1038/nrn2931

Bøttcher, A., Gero, S., Beedholm, K., Whitehead, H., & Madsen, P. T. (2018). Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification. *The Journal of the Acoustical Society of America, 144*(1), 365-374. https://doi.org/10.1121/1.5047657

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication* (2nd ed.). Sinauer Associates Oxford University Press.

Briefer, E., Aubin, T., Lehongre, K., & Rybak, F. (2008). How to identify dear enemies: the group signature in the complex song of the skylark *Alauda arvensis*. *Journal of Experimental Biology, 211*, 317-326. https://doi.org/10.1242/jeb.013359

Briefer, E., Osiejuk, T. S., Rybak, F., & Aubin, T. (2010). Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. *Journal of Theoretical Biology, 262*(1), 151-164. https://doi.org/10.1016/j.jtbi.2009.09.020

Briefer, E., Rybak, F., & Aubin, T. (2008). When to be a dear enemy: flexible acoustic relationships of neighbouring skylarks, *Alauda arvensis*. *Animal Behaviour, 76*, 1319-1325. https://doi.org/10.1016/j.anbehav.2008.06.017

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science, 6*. http://doi.org/10.1098/rsos.181076

Burkard, R., Jones, S., & Jones, T. (1994). Conventional and cross-correlation brain-stem auditory evoked responses in the white leghorn chick: rate manipulations. *The Journal of the Acoustical Society of America, 95*(4), 2136-2144. https://doi.org/10.1121/1.408675

Burkard, R., & Moss, C. F. (1994). The brain-stem auditory-evoked response in the big brown bat (*Eptesicus fuscus*) to clicks and frequency-modulated sweeps. *The Journal of the Acoustical Society of America, 96*(2).

Burkard, R., & Sims, D. (2002). The Human Auditory Brainstem Response to High Click Rates. *American journal of audiology, 10*(2), 53-61. https://doi.org/10.1044/1059-0889(2001/008)

Burkard, R., Trautwein, P., & Salvi, R. (1997). The effects of click level, click rate, and level of background masking noise on the inferior colliculus potential (ICP) in the normal and carboplatin-treated chinchilla. *The Journal of the Acoustical Society of America, 102*(6), 3620-3627. 10.1121/1.420149

Burkard, R. F., Finneran, J. J., & Mulsow, J. (2017). The effects of click rate on the auditory brainstem response of bottlenose dolphins. *The Journal of the Acoustical Society of America, 141*(5), 3396. https://doi.org/10.1121/1.4983447

Burkard, R. F., McGee, J., & Walsh, E. J. (1996a). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. I. Peak latencies. *The Journal of the Acoustical Society of America, 100*(2), 978-990. https://doi.org/10.1121/1.416209

Burkard, R. F., McGee, J., & Walsh, E. J. (1996b). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. II. Peak amplitudes. *The Journal of the Acoustical Society of America, 100*(2), 991-1002. https://doi.org/10.1121/1.416210

Burkard, R. F., & Voigt, H. F. (1989). Stimulus dependecies of the gerbil brain-stem auditory-evoked response (BAER). I: Effects of click level, rate, and polarity. *The Journal of the Acoustical Society of America, 85*(6), 2514-2525.

Call, J., & Tomasello, M. (2020). *The Gestural Communication of Apes and Monkeys* (1st ed.). Psychology Press.

Cameron, D. J., Zioga, I., Lindsen, J. P., Pearce, M. T., Wiggins, G. A., Potter, K., & Bhattacharya, J. (2019, Aug). Neural entrainment is associated with subjective groove and complexity for performed but not mechanical musical rhythms. *Exp Brain Res, 237*(8), 1981-1991. 10.1007/s00221-019-05557-4

Carter, G. G., Skowronski, M. D., Faure, P. A., & Fenton, B. (2008, 2008/10/01/). Antiphonal calling allows individual discrimination in white-winged vampire bats. *Animal Behaviour, 76*(4), 1343-1355. https://doi.org/10.1016/j.anbehav.2008.04.023

Charrier, I., Pitcher, B. J., & Harcourt, R. G. (2009). Vocal recognition of mothers by Australian sea lion pups: individual signature and environmental constraints. *Animal Behaviour, 78*(5), 1127-1134. https://doi.org/10.1016/j.anbehav.2009.07.032

Claesdotter-Hybbinette, E., Matti Cervin, M., & Akerlun, S. (2016). Gender specific differences in auditory brain stem response in young patients with ADHD. *Neuropsychiatry, 06*(01). https://doi.org/10.4172/Neuropsychiatry.1000114

Cortopassi, K. A., & Bradbury, J. W. (2000). The Comparison of harmonically rich Sounds using pectrographic Cross-correlation and Principal Coordinates Analysis. *Bioacoustics, 11*(2), 89-127. https://doi.org/10.1080/09524622.2000.9753454

Dabbs, J. M. (1983, August 23-27). *Fourier analysis of the rhythm of conversation* Annual Meeting of the American Psychological Association, Washington DC. ERIC Document Reproduction Service No. ED 222 959. https://eric.ed.gov/?id=ED222959

Dahlin, C. R., & Benedict, L. (2014, 2014/01/01). Angry Birds Need Not Apply: A Perspective on the Flexible form and Multifunctionality of Avian Vocal Duets [https://doi.org/10.1111/eth.12182]. *Ethology, 120*(1), 1-10. https://doi.org/10.1111/eth.12182

David, J. A. d. O., Zefa, E., & Fontanetti, C. S. (2003). Cryptic species of *Gryllus* in the light of bioacoustic (Orthoptera: Gryllidae). *Neotropical Entomology, 32*, 75-80. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-566X2003000100010&nrm=iso

Doherty, J. A. (1985). Temperature Coupling and 'Trade-Off' Phenomena in the acoustic Communication System of the Cricket, *Gryllus Bimaculates* De Geer (Gryllidae). *The Journal of Experimental Biology, 114*, 17-35.

Duffy, S., & Pearce, M. (2018). What makes rhythms hard to perform? An investigation using Steve Reich's Clapping Music. *PLoS One, 13*(10), e0205847. 10.1371/journal.pone.0205847

Eisenberg, J. F., & Kleiman, D. G. (1972). Olfactory Communication in Mammals. *Annual Review of Ecology and Systematics, 3*(1), 1-32. https://doi.org/10.1146/annurev.es.03.110172.000245

Fernandez, A. A. (2020). Vocal ontogenetic processes in bat pups, From babbling behaviour to the interplay of social and vocal complexity [Doctoral thesis, Freie Universität Berlin].

Fernandez, A. A., & Knörnschild, M. (2017). Isolation Calls of the Bat *Saccopteryx bilineata* Encode Multiple Messages. *Animal Behaviour and Cognition, 4*(2), 169-186. https://doi.org/10.12966/abc.04.05.2017

Fine, M. L., & Parmentier, É. (2015). Mechanisms of Fish Sound Production. In *Sound communication in fishes* (pp. 77-126). Springer.

Fitch, W. T. (2000, Jul). 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

Fitch, W. T. (2010). *The Evolution of Language*. Cambridge University Press.

Forrest, T. G. (1994). From Sender to Receiver: Propagation and Environmental Effects on Acoustic Signals. *American Zoologist, 34*(6), 644-654. https://doi.org/10.1093/icb/34.6.644

Fourier, J.-B. J. (1822). *Théorie analytique de la chaleur*. F. Didot.

Freiherr von Saalfeld, E. (1939). Untersuchungen der Fledermaus-Atmung. *Zeitschrift für vergleichende Physiologie*(26(2)), 242-252. https://doi.org/10.1007/BF00593038

Fulton, B. B. (1932). North Carolina's singing Orthoptera. *Journal of the Elisha Mitchell Scientific Society, 47*, 55-69. http://entnemdept.ufl.edu/Walker/buzz/i00lf32.pdf

García-Rosales, F., Beetz, M. J., Cabral-Calderin, Y., Kössl, M., & Hechavarria, J. C. (2018). Neuronal coding of multiscale temporal features in communication sequences within the bat auditory cortex. *Communications Biology, 1*(1), 200. https://doi.org/10.1038/s42003-018-0205-5

Garcia, M., Theunissen, F., Sebe, F., Clavel, J., Ravignani, A., Marin-Cudraz, T., Fuchs, J., & Mathevon, N. (2020). Evolution of communication signals and information during species radiation. *Nature Communications, 11*(1), 4970. https://doi.org/10.1038/s41467-020-18772-3

Gaunt, S. L. L., Baptista, L. F., Sanchez, J. E., & Hernandez, D. (1994). Song learning as evidenced from song sharing in two hummingbird species (*Colibri coruscans* and *C. thalassinus*). *The Auk, 111*(1), 87-103. https://doi.org/10.2307/4088508

Gochfeld, M. (1978, 1978/01/12). Intraspecific Social Stimulation and Temporal Displacement of Songs of the Lesser Skylark, *Alauda gulgula*. *Zeitschrift für Tierpsychologie, 48*(4), 337-344. https://doi.org/10.1111/j.1439-0310.1978.tb00264.x

Grabe, E., & Low, E. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology, 7*, 515-546.

Hagmann, C. E., & Cook, R. G. (2010). Testing meter, rhythm, and tempo discriminations in pigeons. *Behavioural Processes, 85*(2), 99-110. https://doi.org/10.1016/j.beproc.2010.06.015

Hamilton, J. D. (1994a). Difference Equations. In *Time Series Analysis* (pp. 1-24). Princeton University Press. https://doi.org/ https://doi.org/10.1515/9780691218632-002

Hamilton, J. D. (1994b). Lag Operators. In *Time Series Analysis* (pp. 25-42). Princeton University Press. https://doi.org/https://doi.org/10.1515/9780691218632-007

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science, 298*(5598), 1569. https://doi.org/10.1126/science.298.5598.1569

Hawkins, A., & Amorim, M. C. (2000). Spawning Sounds of the Male Haddock, *Melanogrammus aeglefinus*. *Environmental Biology of Fishes, 59*, 29-41. https://doi.org/10.1023/A:1007615517287

Hechavarría, J. C., Macías, S., Vater, M., Voss, C., Mora, E. C., & Kössl, M. (2013). Blurry topography for precise target-distance computations in the auditory cortex of echolocating bats. *Nature Communications, 4*, 2587. http://dx.doi.org/10.1038/ncomms3587

Hecox, K., & Galambos, R. (1974). Brain Stem Auditory Evoked Responses in Human Infants and Adults. *Archives of Otolaryngology, 99*(1), 30-33. https://doi.org/10.1001/archotol.1974.00780030034006

Heffner, H. E., Koay, G., & Heffner, R. S. (2008). Comparison of behavioural and auditory brainstem response measures of threshold shift in rats exposed to loud sound. *The Journal of the Acoustical Society of America, 124*(2), 1093-1104. https://doi.org/10.1121/1.2949518

Heinsohn, R., Zdenek, C. N., Cunningham, R. B., Endler, J. A., & Langmore, N. E. (2017). Tool-assisted rhythmic drumming in palm cockatoos shares key elements of human instrumental music. *Science Advances, 3*(6), e1602399. https://doi.org/10.1126/sciadv.1602399

Hiss, A. (1983, January 3). Hoover. *The New Yorker*. https://www.newyorker.com/magazine/1983/01/03/hoover

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*, 85-91. https://doi.org/10.1111/j.1749-6632.2011.06402.x

Honing, H., & Ploeger, A. (2012). Cognition and the evolution of music: pitfalls and prospects. *Topics in Cognitiv Sciences, 4*(4), 513-524. https://doi.org/10.1111/j.1756-8765.2012.01210.x

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B, 370*(1664).

Hulse, S. H., Humpal, J., & Cynx, J. (1984a). Discrimination and Generalization of Rhythmic and Arrhythmic Sound Patterns by European Starlings (*Sturnus vulgaris*). *Music Perception, 1*(4).

Hulse, S. H., Humpal, J., & Cynx, J. (1984b). Processing of Rhythmic Sound Structures by Birds. *Annals of the New York Acadmey of Science, 423*(1), 407-419.

Hultsch, H., & Todt, D. (1981). Repertoire sharing and song-post distance in nightingales (*Luscinia megarhynchos B.*). *Behavioural Ecology and Sociobiology, 8*(3), 183-188. https://doi.org/10.1007/BF00299828

Hultsch, H., & Todt, D. (1982). Temporal performance roles during vocal interactions in nightingales (Luscinia megarhynchos B.). *Behavioural Ecology and Sociobiology, 11*(4), 253-260. 10.1007/BF00299302

Hultsch, H., & Todt, D. (1989). Context Memorization in the Song-Learning of Birds. *Naturwissenschaften, 76*, 584-586. https://doi.org/10.1007/bf00462873

Humpal, J., & Cynx, J. A. (1984). Discrimination of Temporal Components of Acoustic Patterns by Birds [https://doi.org/10.1111/j.1749-6632.1984.tb23466.x]. *Annals of the New York Academy of Sciences, 423*(1), 600-602. https://doi.org/10.1111/j.1749-6632.1984.tb23466.x

Hyland Bruno, J., & Tchernichovski, O. (2017). Regularities in zebra finch song beyond the repeated motif. *Behavioural Processes*. https://doi.org/10.1016/j.beproc.2017.11.001

Ioannidis, Y. (2003). The History of Histograms (abridged). In J.-C. Freytag, P. Lockemann, S. Abiteboul, M. Carey, P. Selinger, & A. Heuer (Eds.), *Proceedings 2003 VLDB Conference* (pp. 19-30). Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-012722442-8/50011-2

Ishay, J., Motro, A., Gitter, S., & Brown, M. B. (1974). Rhythms in Acoustical Communication by the Oriental Hornet, *Vespa orientalis Animal Behaviour, 22*, 741-744. https://doi.org/10.1016/s0003-3472(74)80026-6

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology, 28*, 60-65. https://doi.org/10.1016/j.conb.2014.06.010

Jewett, D. L., & Williston, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain, 94*(4), 681-696. https://doi.org/10.1093/brain/94.4.681

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evicence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7*(5), 1059-1073.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*(13), 313-319. https://doi.org/10.1111/1467-9280.00458

Jung, K., Kalko, E. K. V., & Helversen, O. (2007). Echolocation calls in Central American emballonurid bats: signal design and call frequency alternation. *Journal of Zoology,* 272, 125-137.

Juselius Baghdassarian, E., Nilsson Markhed, M., Lindström, E., Nilsson, B. M., & Lewander, T. (2018). Auditory brainstem response (ABR) profiling tests as diagnostic support for schizophrenia and adult attention-deficit hyperactivity disorder (ADHD). *Acta Neuropsychiatrica, 30*(3), 137-147. https://doi.org/10.1017/neu.2017.24

Kalko, E. K. V. (1994). Coupling of sound emission and wingbeat in naturally foraging european pipistrelle bats (Microchiroptera: Vespertilionidae). *Folia Zoologica*, 43(4), 363-376.

Källstrand, J., Lewander, T., Baghdassarian, E., & Nielzén, S. (2014). A new method for analysing auditory brain-stem response waveforms using a moving-minimum subtraction procedure of digitized analog recordings. *Neuropsychiatric Disease and Treatment, 10*, 1011-1016. https://doi.org/10.2147/NDT.S59178

Kever, L., Boyle, K. S., & Parmentier, E. (2015). Effects of seawater temperature on sound characteristics in *Ophidion rochei* (Ophidiidae). *Journal of Fish Biology, 87*(2), 502-509. https://doi.org/10.1111/jfb.12730

Khanna, H., Gaunt, S. L. L., & McCallum, D. A. (1997). Digital Spectrographic Cross-Correlation: Tests of Sensitivity. *Bioacoustics, 7*(3), 209-234. https://doi.org/10.1080/09524622.1997.9753332

Knörnschild, M. (2014). Vocal production learning in bats. *Current Opinion in Neurobiology, 28*, 80-85. https://doi.org/10.1016/j.conb.2014.06.014

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behaviour in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften,* 93, 451-454.

Knörnschild, M., Bluml, S., Steidl, P., Eckenweber, M., & Nagy, M. (2017). Bat songs as acoustic beacons - male territorial songs attract dispersing females. *Scientific Reports, 7*(1), 13918. https://doi.org/10.1038/s41598-017-14434-5

Knörnschild, M., Feifel, M., & Kalko, E. K. V. (2013). Mother-offspring recognition in the bat *Carollia perspicillata*. *Animal Behaviour, 86*, 941-948. https://doi.org/10.1016/j.anbehav.2013.08.011

Knörnschild, M., Jung, K., Nagy, M., Metz, M., & Kalko, E. K. V. (2012). Bat echolocation calls facilitate social communication. *Proceedings of the Royal Society B, 279*, 4827-4835. https://doi.org/10.1098/rspb.2012.1995

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters, 6*(2), 156-159.

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Animal Behaviour, 84*, 761-769. http://dx.doi.org/10.1016/j.anbehav.2012.06.029

Knörnschild, M., & von Helversen, O. (2008). Non-mutual vocal mother-pup recognition in the sac-winged bat, *Saccopteryx bilineata*. *Animal Behaviour, 76*(3), 1001-1009. https://doi.org/10.1016/j.anbehav.2008.05.018

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences, 22*(10), 896-910. https://doi.org/10.1016/j.tics.2018.08.002

Ladegaard, M., Jensen, F. H., de Freitas, M., Ferreira da Silva, V. M., & Madsen, P. T. (2015, Oct). Amazon river dolphins (*Inia geoffrensis*) use a high-frequency short-range biosonar. *Journal of Experimental Biolology, 218*(19), 3091-3101. https://doi.org/10.1242/jeb.120501

Ladich, F. (2014). Fish bioacoustics. *Current Opinion in Neurobiology, 28*, 121-127. https://doi.org/10.1016/j.conb.2014.06.013

Ladich, F. (2018). Acoustic communication in fishes: Temperature plays a role. *Fish and Fisheries, 19*(4), 598-612. https://doi.org/10.1111/faf.12277

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*, 110-113. https://doi.org/10.1126/science.1154735

Land, R., Burghard, A., & Kral, A. (2016). The contribution of inferior colliculus activity to the auditory brainstem response (ABR) in mice. *Hearing Research, 341*, 109-118. https://doi.org/10.1016/j.heares.2016.08.008

Large, E. W., & Jones, M. R. (1999). The dynamics of attenting: how people track time varving events. *Psychological Review*(106), 119-159.

Lattenkamp, E. Z., Nagy, M., Drexl, M., Vernes, S. C., Wiegrebe, L., & Knornschild, M. (2021). Hearing sensitivity and amplitude coding in bats are differentially shaped by echolocation calls and social calls. *Proceedings of the Royal Society B, 288*, 20202600. https://doi.org/10.1098/rspb.2020.2600

Le Bot, O., Bonnel, J., Mars, J. I., & Gervaise, C. (2013). Odontocete click train deinterleaving using a single hydrophone and rhythm analysis. *The Journal of the Acoustical Society of America, 133*(5), 3312-3312. https://doi.org/10.1121/1.4805506

Lin, H., & Wang, Q. (2007). Mandarin rhythm: An acoustic study. *Journal of Chinese Language and Computing, 17*(3), 127-140.

Linnenschmidt, M., & Wiegrebe, L. (2019). Ontogeny of auditory brainstem responses in the bat, *Phyllostomus discolour*. *Hearing Research, 373*, 85-95. https://doi.org/10.1016/j.heares.2018.12.010

Macias, S., Hechavarria, J. C., & Kossl, M. (2016). Temporal encoding precision of bat auditory neurons tuned to target distance deteriorates on the way to the cortex. *Journal of Comparative Physiology A, 202*(3), 195-202. https://doi.org/10.1007/s00359-016-1067-2

Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B, 268*, 2315-2324. https://doi.org/10.1098/rspb.2001.1773

Marler, P., & Slabbekoorn, H. (2004). *Nature's Music The Science of Birdsong*. Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-473070-0.X5000-2

Martins, P. T., & Boeckx, C. (2020). Vocal learning: Beyond the continuum. *PLoS Biol, 18*(3), e3000672. https://doi.org/10.1371/journal.pbio.3000672

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals Voices. *Current Biology, 27*(15), 2352-2356.e2352. https://doi.org/10.1016/j.cub.2017.06.035

Mathevon, N., Koralek, A., Weldele, M., Glickman, S., E., & Theunissen, F. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecology, 10*(9). https://doi.org/10.1186/1472-6785-10-9

McRae, T. R. (2020). A review of squirrel alarm-calling behaviour: What we know and what we do not know about how predator attributes affect alarm calls. *Animal Behaviour and Cognition, 7*(2), 168-191. https://doi.org/10.26451/abc.07.02.11.2020

Merguerditchian, A., Vuillemin, A., & Pruetz, J. D. (2018). *Identifying the ape beat in the wild: rhythmic indvidual signatures from sounds of manual fruit cracking in Fongoli Chimpanzees* The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12),

Moss, C. F., Bohn, K., Gilkenson, H., & Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *PLOS Biology, 4*(4), e79. https://doi.org/10.1371/journal.pbio.0040079

Moss, C. F., & Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *The Journal of the Acoustical Society of America, 110*(4), 2207-2226. https://doi.org/10.1121/1.1398051

Munafó, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *nature human behaviour, 1*. https://doi.org/10.1038/s41562-016-0021

Norton, P., Barschke, P., Scharff, C., & Mendoza, E. (2019). Differential Song Deficits after Lentivirus-Mediated Knockdown of FoxP1, FoxP2, or FoxP4 in Area X of Juvenile Zebra Finches. *The Journal of Neuroscience, 39*(49), 9782-9796. https://doi.org/10.1523/jneurosci.1250-19.2019

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

Obrist, M. K., & Wenstrup, J. J. (1998). Hearing and hunting in red bats (*Lasiurus borealis*, Vespertilionidae): audiogram and ear properties. *Journal of Experimental Biology, 201*, 143-154.

Parmentier, E., Lecchini, D., Frederich, B., Brie, C., & Mann, D. (2009). Sound production in four damselfish (Dascyllus) species: phyletic relationships? *Biological Journal of the Linnean Society, 97*(4), 928-940. https://doi.org/10.1111/j.1095-8312.2009.01260.x

Patel, A. D. (2008). *Music, Language and the Brain*. NY: Oxford University Press.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009a,). Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology, 19*(10), 827-830. https://doi.org/10.1016/j.cub.2009.03.038

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009b). Studying Synchronization to a Musical Beat in Nonhuman Animals. *Annals of the New York Academy of Sciences, 1169*(1), 459-469. https://doi.org/10.1111/j.1749-6632.2009.04581.x

Pepperberg, I. M. (2010). Vocal learning in Grey parrots: A brief review of perception, production, and cross-species comparisons. *Brain and Language, 115*(1), 81-91. https://doi.org/10.1016/j.bandl.2009.11.002

Percival, G., & Tzanetakis, G. (2014). Streamlined Tempo Estimation Based on Autocorrelation and Cross-correlation With Pulses. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 22*(12), 1765-1776. https://doi.org/10.1109/TASLP.2014.2348916

Picciulin, M., Fiorin, R., Facca, C., & Malavasi, S. (2020). Sound features and vocal rhythms as a proxy for locating the spawning ground of *Sciaena umbra* in the wild. *Aquatic Conservation: Marine and Freshwater Ecosystems, 30*(7), 1299-1312. https://doi.org/10.1002/aqc.3340

Picciulin, M., Sebastianutto, L., Codarin, A., Calcagno, G., & Ferrero, E. A. (2012). Brown meagre vocalization rate increases during repetitive boat noise exposures: a possible case of vocal compensation. *The Journal of the Acoustical Society of America, 132*(5), 3118-3124. https://doi.org/10.1121/1.4756928

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B, 285*, 20180598. https://doi.org/10.1098/rspb.2018.0598

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Elephants are capable of vocal learning. *Nature, 434*, 455-456. https://doi.org/10.1038/434455a

Portfors, C. V., & Wenstrup, J. J. (1999). Delay-Tuned Neurons in the Inferior Colliculus of the Mustached Bat: Implications for Analyses of Target Distance. *Journal of Neurophysiology, 82*(3), 1326-1338.

Potter, R. K., Kopp, G. A., & Kopp, H. C. G. (1947). *Visible Speech* D. Van Nostrand Company. https://books.google.de/books?id=qSAByAEACAAJ

Pougnault, L., Levréro, F., Mulot, B., & Lemasson, A. (2020). Breaking conversational rules matters to captive gorillas: A playback experiment. *Scientific Reports, 10*(1), 6947. https://doi.org/10.1038/s41598-020-63923-7

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., & White, L. E. (2012). *Neuroscience* (5th Edition ed.). Sinauer Associates Inc.

Ratcliffe, J. M., Elemans, C. P., Jakobsen, L., & Surlykke, A. (2013). How the bat got its buzz. *Biology Letters, 9*(2). https://doi.org/10.1098/rsbl.2012.1031

Ratcliffe, J. M., Jakobsen, L., Kalko, E. K. V., & Surlykke, A. (2011). Frequency alternation and an offbeat rhythm indicate foraging behaviour in the echolocating bat, *Saccopteryx bilineata*. *Journal of Comparative Physiology A*. https://doi.org/10.1007/s00359-011-0630-0

Ravignani, A. (2018). Spontaneous rhythms in a harbor seal pup calls. *BMC research notes, 11*(1). https://doi.org/10.1186/s13104-017-3107-6

Ravignani, A. (in press). Isochrony, vocal learning and the acquisition of rhythm and melody. *Behavioural and Brain Science, in press*.

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01118

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speec, movement and animal vocalizations. *Journal of Language Evolution*. https://doi.org/10.1093/jole/lzx002

Ravignani, A., Thompson, B., & Filippi, P. (2018). The Evolution of Musicality: What Can Be Learned from Language Evolution Research? *Frontiers in Neuroscience, 12*(20). https://doi.org/10.3389/fnins.2018.00020

Reichmuth, C., & Casey, C. (2014, 2014/10/01/). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology, 28*, 66-71. https://doi.org/10.1016/j.conb.2014.06.011

Reinert, J. (1965). Takt- und Rhythmusunterscheidung bei Dohlen. *Zeitschrift für Tierpsychologie, 22*(6). https://doi.org/10.1111/j.1439-0310.1965.tb01683.x

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., & Galambos, R. (1981). Auditory brainstem response in dolphins. *Proceedings of the National Academy of Science, 78*(3). https://doi.org/10.1073/pnas.78.3.1943

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

Roeske, T., Kelty-Stephen, D., & Wallot, S. (2018). Multifractal analysis reveals music-like dynamic structure in songbird rhythms. *Scientific Reports,* 8. https://doi.org/10.1038/s41598-018-22933-2

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *The Journal of Neuroscience,* 32, 8424-8428. https://doi.org/10.1523/JNEUROSCI.0804-12.2012

Ryan, M. J., & Sullivan, B. K. (1989). Transmission effects on Temporal Structure in the Advertisement Calls of Two Toads, *Bufo woodhousii* and *Bufo valliceps*. *Ethology, 80*, 182-189. https://doi.org/10.1111/j.1439-0310.1989.tb00738.x

Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One, 3*(1), e1461. https://doi.org/10.1371/journal.pone.0001461

Sasahara, K., Tchernichovski, O., Takahasi, M., Suzuki, K., & Okanoya, K. (2015). A rhythm landscape approach to the developmental dynamics of birdsong. *Journal of the Royal Society Interface, 12*(112). https://doi.org/10.1098/rsif.2015.0802

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology, 19*(10), 831-836. https://doi.org/10.1016/j.cub.2009.03.061

Scharff, C., & Nottebohm, F. (1991). A comparative study of the behavioural deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *The Journal of Neuroscience, 11*(9), 2896. https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991

Schneider, J. N., & Mercado, E. (2018). Characterizing the rhythm and tempo of sound production by singing whales AU - Schneider, Jennifer N. *Bioacoustics*, 1-18. https://doi.org/10.1080/09524622.2018.1428827

Schnitzler, H. U. (1971). Fledermäuse im Windkanal. *Zeitschrift für vergleichende Physiologie*, 73, 209-221.

Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Animal Behaviour, 76*(6), 1977-1988. https://doi.org/10.1016/j.anbehav.2008.07.032

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Science, 1*, 291-295.

Simpson, H. B., & Vicario, D. S. (1990). Brain Pathways for Learned and Unlearned Vocalizations Differ in Zebra Finches. *The Journal of Neuroscience, 10*(5), 1541-1556.

Slabbekoorn, H., & ten Cate, C. (1999). Collared Dove Responses to Playback: Slaves to the Rhythm. *Ethology, 105*(5), 377-391. https://doi.org/10.1046/j.1439-0310.1999.00420.x

Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song Learning in Zebra Finches (*Taeniopygia guttata*): Progress and Prospects. In J. S. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater (Eds.), *Advances in the Study of Behaviour* (Vol. 18, pp. 1-34). Academic Press. https://doi.org/https://doi.org/10.1016/S0065-3454(08)60308-3

Sprague, M. W. (2000). The single sonic muscle twitch model for the sound-production mechanism in the weakfish, *Cynoscion regalis*. *The Journal of the Acoustical Society of America, 108*(5), 2430-2437. https://doi.org/10.1121/1.1315296

Stansbury, A. L., & Janik, V. M. (2019). Formant Modification through Vocal Production Learning in Gray Seals. *Current Biology, 29*(13), 2244-2249.e2244. https://doi.org/10.1016/j.cub.2019.05.071

Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology, 120*(2), 304-308. https://doi.org/10.1016/j.clinph.2008.11.020

Sun, J. W. C., & Narins, P. M. (2005). Anthropogenic sounds differentially affect amphibian call rate. *Biological Conservation, 121*(3), 419-427. https://doi.org/10.1016/j.biocon.2004.05.017

Suthers, R. A., Thomas, S. P., & Suthers, B. J. (1972). Respiration, Wing-beat and Ultrasonic Pulse Emission in an Echo-locating Bat. *The Journal of Experimental Biology, 56*, 37-48.

Szymanski, M., Bain, D., Kiehl, K., Pennington, S., Wong, S., & Henry, K. R. (1999). Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioural audiograms. *The Journal of the Acoustical Society of America, 106*(2), 1134-1141.

Takahashi, Daniel Y., Narayanan, Darshana Z., & Ghazanfar, Asif A. (2013). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology, 23*(21), 2162-2168. 10.1016/j.cub.2013.09.005

Tchernichovski, O., Mitra, P. P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science, 291*, 2564-2569. https://doi.org/10.1126/science.1058522

ten Cate, C., & Spierings, M. (2019). Rules, rhythm and grouping: auditory pattern perception by birds. *Animal Behaviour, 151*, 249-257. https://doi.org/10.1016/j.anbehav.2018.11.010

ten Cate, C., Spierings, M., Hubert, J., & Honing, H. (2016). Can Birds Perceive Rhythmic Patterns? A Review and Experiments on a Songbird and a Parrot Species. *Frontiers in Psychology, 7*, 730. https://doi.org/10.3389/fpsyg.2016.00730

Terry, A. M. R., McGregor, P. K., & Peake, T. M. (2001). A Comparison of some Techniques used to assess vocal Individuality. *Bioacoustics, 11*(3), 169-188. https://doi.org/10.1080/09524622.2001.9753461

Thorpe, W. H. (1958). The Learning of Song Patterns by Birds, with especial Relevance to the Song of the Chagginch *Fringilla coelebs*. *Ibis, 100*, 535-570. https://doi.org/10.1111/j.1474-919x.1958.tb07960.x

Thorpe, W. H. (1975). The biological significance of duetting and antiphonal song. *Acta Neurobiologiae Experimentalis, 35*(5-6), 517-528.

Tønnesen, P., Gero, S., Ladegaard, M., Johnson, M., & Madsen, P. T. (2018). First-year sperm whale calves echolocate and perform long, deep dives. *Behavioural Ecology and Sociobiology, 72*(10), 165. https://doi.org/10.1007/s00265-018-2570-y

Toussaint, G. T. (2012, 23.07-28.07). The Pairwise Variability Index as a Tool in Musical Rhythm Analysis. In E. Cambouroploulos, C. Tsougras, P. Mavromatis, & K. Pastiadis, Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Science of Music, Thessaloniki, Greece.

Toussaint, G. T. (2013). *The geometry of musical rhythm: what makes a "good" rhythm good?* CRC Press.

van der Aa, J., Honing, H., & ten Cate, C. (2015). The perception of regularity in an isochronous stimulus in zebra finches (*Taeniopygia guttata*) and humans. *Behavioural Processes, 115*, 37-45. https://doi.org/10.1016/j.beproc.2015.02.018

Vernes, S. C., & Wilkinson, G. S. (2019). Behaviour, biology and evolution of vocal learning in bats. *Philosophical Transactions B, 375*(1789), 20190061. https://doi.org/10.1098/rstb.2019.0061

Walker, T. J. (1964). Cryptic Species Among Sound-Producing Ensiferan Orthoptera (Gryllidae and Tettigoniidae). *The Quarterly Review of Biology, 39*(4), 345-355. https://doi.org/10.1086/404325

Wallin, N. L. (1991). *Biomusicology - Neurophysiological, Neuropsychological, and Evolutionary Perspectives on the Origins and Purpose of Music*. Pendragon Press.

Wetekam, J., Reissig, C., Hechavarria, J. C., & Kössl, M. (2020). Auditory brainstem responses in the bat *Carollia perspicillata*: threshold calculation and relation to audiograms based on otoacoustic emission measurement. *Journal of Comparative Physiology A, 206*(1), 95-101. https://doi.org/10.1007/s00359-019-01394-6

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and Language, 144*, 26-34. https://doi.org/10.1016/j.bandl.2015.03.008

Wilbrecht, L., & Nottebohm, F. (2003). Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Reviews, 9*(3), 135-148. https://doi.org/10.1002/mrdd.10073

Wiley, R. H., & Richards, D. G. (1978). Physical Constraints on Acoustic Communication in the Atmosphere: Implications for the Evolution of Animal Vocalizations. *Behavioural Ecology and Sociobiology, 3*, 69-94. https://doi.org/10.1007/bf00300047

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences, 106*(7), 2468. https://doi.org/10.1073/pnas.0809035106

Wong, J. G., & Waters, D. A. (2001). The Synchonisation of signal emission with wingbeat during the approach phase in Soprano Pipistrelles (*Pipistrellus Pygmaeus*). *The Journal of Experimental Biology*(204), 575-583.

Yoshida, S., & Okanoya, K. (2005). Evolution of Turn-Taking: A Bio-Cognitive Perspective. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 12*(3), 153-165. https://doi.org/10.11225/jcss.12.153

Zsebők, S., Schmera, D., Laczi, M., Nagy, G., Vaskuti, É., Török, J., Zsolt Garamszegi, L., & Ellison, A. (2021). A practical approach to measuring the acoustic diversity by community ecology methods. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210x.13558

Chapter I

General Isochronous Rhythm in Echolocation Calls and Social Vocalizations of the Bat *Saccopteryx bilineata*

Adapted from: Burchardt, L.S.; Norton, P.; Behr, O.; Scharff, C.; Knörnschild, M. (2019) General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*,

In: Royal Society Open Science 6: 181076. http://dx.doi.org/10.1098/rsos.181076

Abstract

Rhythm is an essential component of human speech and music but very little is known about its evolutionary origin and its distribution in animal vocalizations. We found a regular rhythm in three multisyllabic vocalization types (echolocation call sequences, male territorial songs, and pup isolation calls) of the neotropical bat *Saccopteryx bilineata*. The intervals between element onsets were used to fit the rhythm for each individual. For echolocation call sequences, we expected rhythm frequencies around 6-24 Hz, corresponding to the wingbeat in *S. bilineata* which is strongly coupled to echolocation calls during flight. Surprisingly, we found rhythm frequencies between 6 Hz and 24 Hz not only for echolocation sequences but also for social vocalizations, e.g. male territorial songs and pup isolation calls, which were emitted while bats were stationary. Fourier analysis of element onsets confirmed an isochronous rhythm across individuals and vocalization types. We speculate that attentional tuning to the rhythms of echolocation calls on the receivers' side might make the production of equally steady rhythmic social vocalizations beneficial.

Introduction

Music is widespread in all human cultures but its evolutionary origin is poorly understood (Honing et al., 2015). The field of biomusicology attempts to answer questions on the origin and purpose of music by focusing on the physiological, psychological, behavioural, and evolutionary aspects of music in a comparative approach. That approach includes not only human music but musicality as a term for different traits that occur spontaneously and are based on and constrained by biology and cognition in animal vocalizations (Ravignani et al., 2018; Wallin, 1991). Music contains several key components – that can be separately investigated as musicality traits – such as pitch (governing melody and harmony), rhythm (defining temporal structure), and sonic qualities named timbre (Honing et al., 2015). Our study focuses on rhythm as a musicality trait, likely with multiple evolutionary backgrounds (Kotz et al., 2018).

Rhythm can be defined as a "systematic patterning of sound in terms of timing, accent, and grouping" (Patel, 2008). Overall, our intuitive understanding of rhythm concerns periodicity, which is the expectation of a recurrent event. One special kind of periodic rhythm is an isochronous beat, as produced e.g. by a metronome. In an isochronous beat, all beats have the same length and all beat-to-beat intervals have the same length (Patel, 2008). When it comes to analysing animal vocalizations for rhythmicity, two questions need to be answered. (a) How well can an animal produce a certain rhythm and (b) are rhythmic patterns similar or different between vocalization types and between individuals? Another interesting comparison not regarded in this project would be between species. Furthermore, the relevance and biological constraints shaping an existing rhythm need to be discussed.

In a recent study on rhythm in song of zebra finches (*Taeniopygia guttata*) both questions were answered. Individual males had a distinct isochronous rhythm which fitted syllable onsets better than expected by chance. Distinct rhythms between individual males ranged from 10 to 60 Hz (Norton & Scharff, 2016). Other examples of animals producing rhythmic signals include the Palm Cockatoo (*Probosciger aterrimus*) which uses tools to 'drum' a quasi-isochronous beat on branches in a consistent context (the rhythm frequencies were not analysed in detail) (Heinsohn et al., 2017b) or chimpanzees cracking baobab fruits in a fashion probably eligible to generate individual signatures, which might help to recognize unseen companions (Merguerditchian et al., 2018). A subsequent question would be whether animals can distinguish between rhythms, isochronous or otherwise. Rats for example are able to discriminate between different isochronous rhythms in a habituation-dishabituation experiment (Celma-Miralles & Toro, 2018b) while European Starlings are able to discriminate between rhythmic and arrhythmic patterns (Hulse et al., 1984a). Moreover, the first instance for a biologically relevant rhythm in non-human mammalian vocalizations was found in the Northern Elephant Seal, where males can discriminate between familiar and unfamiliar male opponents using the temporal structure of vocalizations. Rhythms apparently differ between individuals in a way that facilitates discrimination of individuals (Mathevon et al., 2017). Nevertheless, compared to other aspects of vocal communication, studies on rhythmicality in animals are still sparse.

Our study aims to broaden the knowledge of rhythm in animal vocalizations by investigating whether isochronous rhythms can be found in different vocalization types of bats. Specifically, we investigated how well different vocalizations of bats fit an isochronous beat and whether the patterns are similar between individuals or vocalization types.

We studied the Neotropical greater sac-winged bat *Saccopteryx bilineata* which has a rich vocal repertoire (Behr & Helversen, 2004) and is capable of vocal production learning (Knörnschild et al., 2010). The species' vocal repertoire consists of distinct vocalization types that are uttered in different behavioural contexts. In this study, we focused on echolocation call sequences, isolation calls, and territorial songs, all of which are multisyllabic vocalizations with clear syllable onsets. Isolation calls are produced by pups to solicit maternal care and by adult males to appease dominant conspecifics (Fernandez & Knörnschild, 2017; Knörnschild, Nagy, et al., 2012; Knörnschild & von Helversen, 2008). With a length of up to 2 seconds and a multisyllabic structure, isolation calls of *S. bilineata* are amongst the most acoustically complex bat isolation calls described (Knörnschild, Nagy, et al., 2012; Knörnschild & von Helversen, 2008). Territorial songs are produced by adult males to repel rivals and attract mating partners (Behr & Helversen, 2004; Knörnschild et al., 2017). They are acquired by imitating conspecifics' song during ontogeny (Eckenweber & Knörnschild, 2013; Knörnschild et al., 2017; Knörnschild et al., 2010). Echolocation calls are produced by male and female *S. bilineata* for orientation, navigation, and insect prey capture (Jung, Kalko, & von Helversen, 2007); in addition to their primary function, echolocation calls facilitate social communication among group members (Knörnschild, Jung, et al., 2012a). We chose those three vocalization types to get insight into rhythmicity

in both innate vocalizations (isolation calls, echolocation call sequences) and learned ones (territorial songs) as well as to investigate potential age differences in rhythmicity (in pup isolation calls).

The individual rhythms found in zebra finch song were discussed to be potentially advantageous for anticipating events, i.e. song syllables. Tuning attention to rhythmic production could reduce 'attentional energy' (*sensu*: (Bermeitinger & Frings, 2015)) and increase signal perception (Norton & Scharff, 2016).Correspondingly, rhythmicity in bat vocalizations might be adaptive for saving metabolic energy since flight is energetically costly. In many bat species, echolocation calls are coupled to wingbeat and respiratory cycle (e.g. (Kalko, 1994; Schnitzler, 1971; Suthers et al., 1972; Wong & Waters, 2001)), which is thought to be energy efficient. Moreover, not only behavioural correlates can be found but neuronal correlates: wingbeat and echolocation calls in *Roussettus aegyptiacus* are tightly coupled around theta frequencies (5 – 12 Hz, (Yartsev & Ulanovsky, 2013)), brain wave frequencies which are known to play a role in active movement and stimulus intake (Colgin, 2013). Consistently, preliminary data on *S. bilineata* suggests a wingbeat of around 6-12 Hz (pers. communication H.-U. Schnitzler). During search flight one or two echolocation calls might be uttered per wingbeat, which corresponds to echolocation call intervals of 6 to 24 Hz (wingbeat frequencies of around 6-12 Hz) found in other studies on *S. bilineata* (Bayefsky-Anand, 2006; Jung, Kalko, & von Helversen, 2007; Ratcliffe et al., 2011).

Because of the coupling of wingbeat and echolocation pulses, we predicted periodic, isochronous pulses (following (Ravignani et al., 2014)) with frequencies between approximately 6 to 24 Hz in echolocation call sequences of *S. bilineata*. We assumed that echolocation call sequences would fit a specific isochronous rhythm significantly better than random vocal sequences would. Moreover, we expected this rhythm to be similar between individuals due to common physiological constraints. Since social vocalizations (pup isolation calls and male territorial songs) are uttered by perched bats in the day roost, not coupled to wingbeat, we predicted to find individually different rhythms that might support vocal discrimination of different individuals, as previous research suggests.

Methods

Labelling of vocalization types

We analysed three different vocalization types of *S. bilineata*, namely isolation calls, territorial songs, and echolocation call sequences (Figure 1). Isolation calls and territorial songs are multi-component vocalization types containing four different element types each, while echolocation call sequences are series of one element type with alternating frequencies.

For each vocalization, the on- and offset of its elements and the duration of the silent gaps between elements was determined for subsequent analyses. For isolation calls and territorial songs, element on- and offsets were determined manually based on oscillograms (see (Knörnschild, Nagy, et al., 2012) and (Behr et al., 2006) for details). For echolocation call sequences, we used an

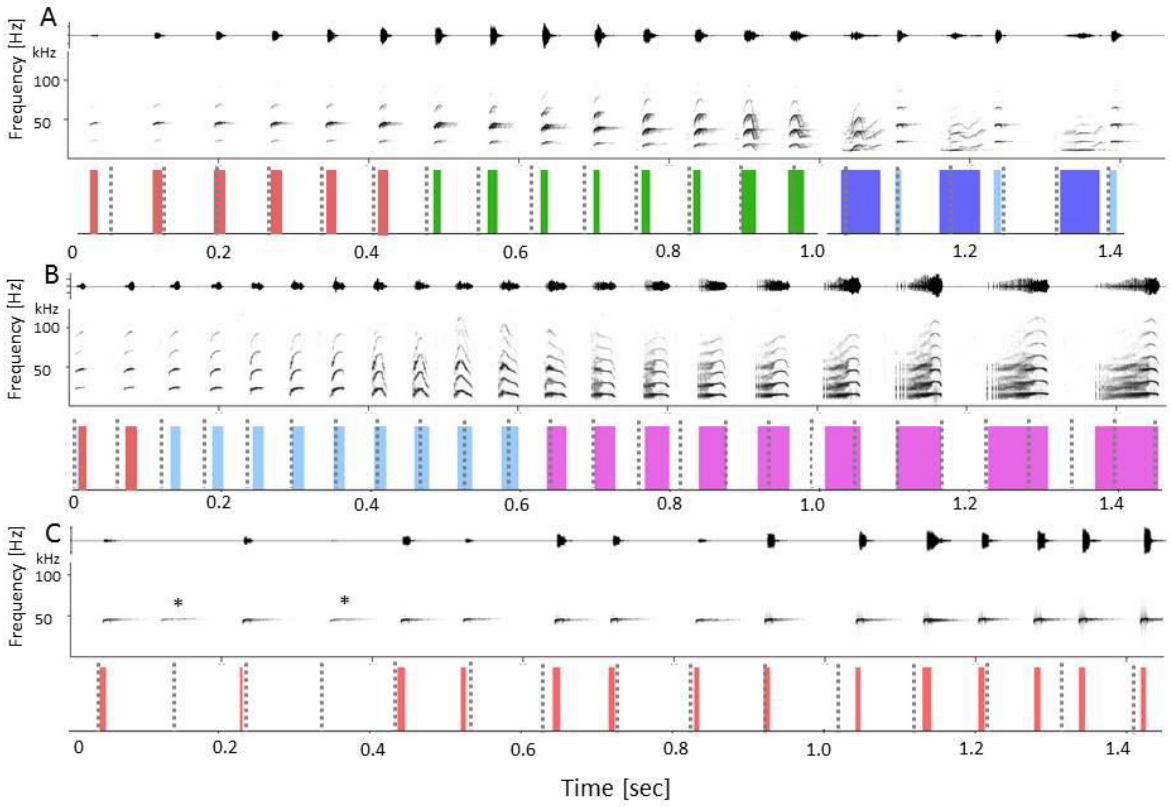


Figure 2: Rhythm^S Fits Well on Three Vocalization Types
Oscillograms (top rows in A–C) and spectrograms (middle rows) of vocalizations (A): isolation call, (B): territorial song, (C): echolocation call sequence) with fitted rhythmS as dotted lines in the bottom row. Element durations are indicated by coloured bars, measured from the oscillograms. Note that echoes visible in the spectrograms may make the elements appear longer than they are in the oscillograms. Different colours indicate different element types (described in earlier studies (Knörnschild et al., 2008;Behr et al., 2006). (A) Introductory elements, simple variable elements followed by composite elements and simple stereotyped elements in an alternating order. (B) Echolocation-like calls (comparable to the introductory elements in (A), short tonal elements and buzz elements. (C) Echolocation calls. (*) indicates two elements not being labelled due to a low amplitude.

automatized procedure in Avisoft SASLab Pro (based on amplitude detection threshold; - 20 dB relative to the call's peak frequency) to determine element on- and offsets.

We analysed isolation calls from 25 pups (10 males, 13 females, 2 not sexed) belonging to a population of wild *S. bilineata* in Costa Rica (see (Knörnschild, Nagy, et al., 2012) for details on study site and sound recordings). Each isolation call contained 5 – 26 elements (14 ± 3.5, mean ± SD) and was composed of 2 – 4 different element types (mean: 3 element types). We followed the nomenclature introduced in an earlier study (Knörnschild & von Helversen, 2008) and labelled the element types (a-d) as introductory elements (a), simple variable elements (b), composite elements (c), and simple stereotyped elements (d). Data for each pup consisted of 20 isolation calls, recorded at two ontogenetic

stages (non-volant and volant; 10 isolation calls each). Only one call per pup and day was selected to minimize temporal dependence among vocalizations.

We analysed territorial songs of 14 adult males belonging to a population of wild *S. bilineata* in Costa Rica (see (Behr et al., 2006) for details on study site and sound recordings). Data for each male consisted of 10 – 11 songs, which were recorded on different days. Each song contained 6 – 46 elements (20 ± 8.0 mean ±SD) and was composed of 1 – 5 five different element types (mean: 3 different element types). We followed the nomenclature introduced in an earlier study (Behr et al., 2006) and labelled the element types (a-e) as short tonal elements (a), buzz elements (b), trills (c), noise bursts (d) and echolocation-like calls (e) (Figure 1).

Sequences of echolocation calls were recorded from 33 wild Costa Rican *S. bilineata* (15 males, 18 females) when they were released after capture, i.e. in a non-foraging context. Calls of known individuals were recorded in standardized release situations in relatively open space (e.g. at a forest clearing). Recorded calls resembled normal search calls (see (Knörnschild, Jung, et al., 2012a) for details on study sites and sound recordings). Echolocation call sequences consisted of 11 – 38 elements (21± 6.95, mean ± SD) with no further differentiation into different element types. One echolocation call sequence per bat was used for further analysis.

Assessment of best-fitting rhythms

Simply analysing inter-onset intervals of social vocalizations, as is often done for echolocation call sequences (e.g. (Bayefsky-Anand, 2006; Jung, Kalko, & von Helversen, 2007; Ratcliffe et al., 2011)), is problematic since this would oversimplify the temporal structure of multisyllabic social vocalizations with strongly varying syllable durations. Other approaches to analyse temporal structure of animal vocalizations include generate-and-test approaches or Fourier Analysis (Ravignani & Norton, 2017). We chose a generate-and-test approach (GAT approach) originally developed for rhythm analysis in zebra finch song (Norton & Scharff, 2016). The GAT approach allowed us to find an isochronous rhythm (i.e. a pattern with equal time intervals) that best fitted the onsets of elements in a given sequence. We named this best fitting rhythm 'signal-derived rhythm' or rhythm^{S} (same as pulse^{S} in (Norton & Scharff, 2016)). The GAT approach was performed by a custom MATLAB program (see (Norton & Scharff, 2016) section 3.7 for more details). It creates isochronous pulse trains in 0.01 Hz increments in a predefined frequency window of 5-100 Hz (i.e. 5-100 pulses per second). The lower range of rhythm frequencies was determined by expected values (Bayefsky-Anand, 2006; Jung, Kalko, & von Helversen, 2007; Ratcliffe et al., 2011) the upper range experimentally by testing different ranges. 100 Hz was deemed appropriate because, when testing for up to 200 Hz only very few values for best fitting rhythms lay above 100 Hz. Restricting the frequency window was a question of minimizing computing time. For each rhythm, temporal deviations of each element to the nearest pulse gave an overall root-mean-square deviation (RMSD). Pulses were offset (+ one phase in 1 ms steps, see (Norton & Scharff, 2016)) to minimize the RMSD. Since RMSD is negatively correlated with rhythm frequency (i.e. faster rhythms generally result in lower RMSD values; see Figure 2, bottom), we normalized the RMSD by multiplying

it by the respective rhythm frequency, yielding a measure for deviation relative to the rhythm period; it describes the average temporal deviation as a fraction of a full cycle. The resulting frequency-normalized RMSD (FRMSD) was used to assess the goodness-of-fit for each rhythm: the lowest FRMSD indicated the best-fitting rhythm frequency. This way the slowest isochronous rhythm, coinciding best with element onsets, was found (Figure 3).

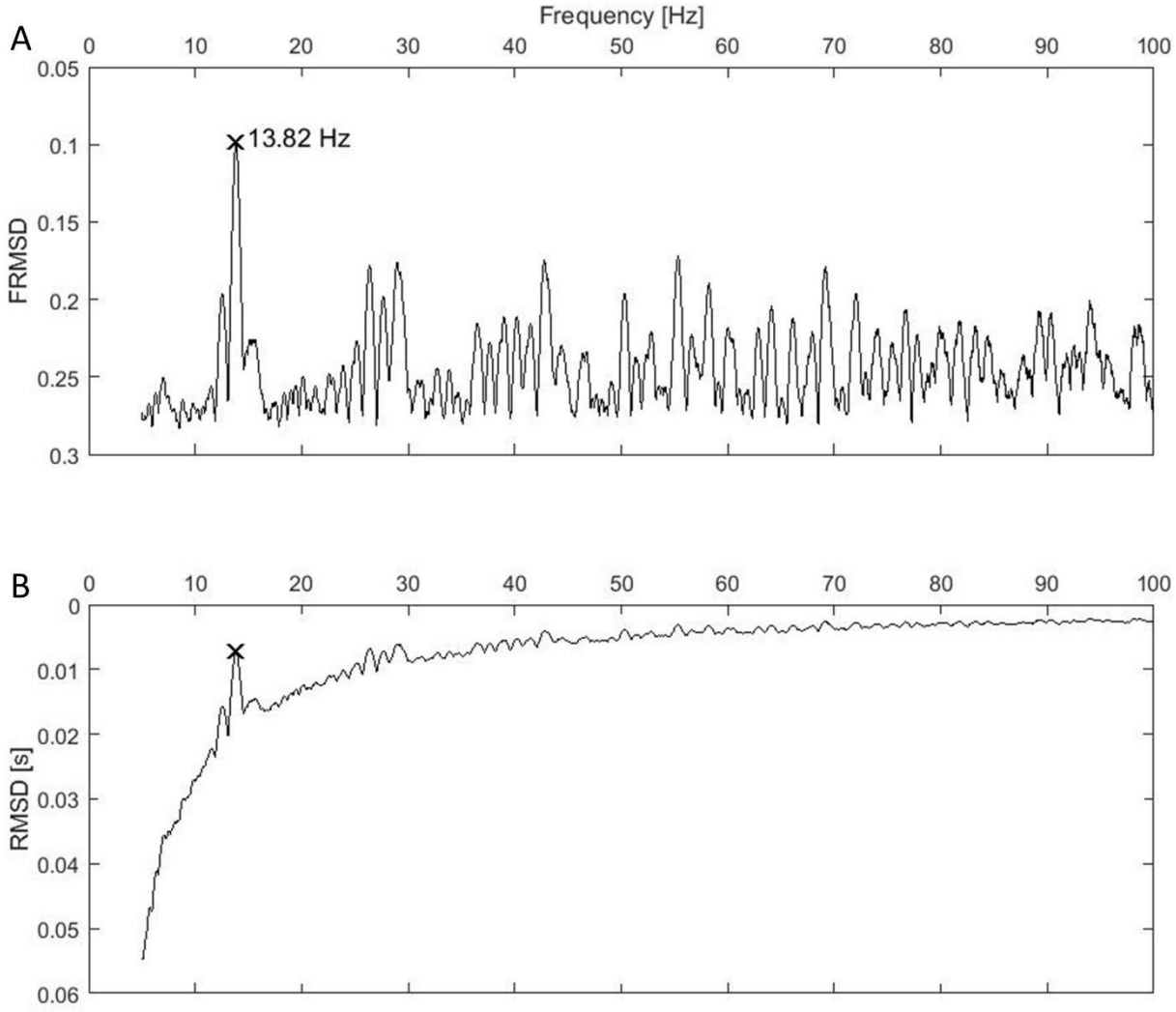


Figure 3: Optimization Process
Best-fitting rhythms were found by selecting the rhythm with the lowest corresponding FRMSD (black cross with corresponding rhythm^S), the frequency-normalized root-mean-square deviation (A); (B) shows the corresponding RMSD values.

Clustering

A visual examination of the resulting best-fitting rhythm^S indicated an accumulation of certain frequency values for each individual and vocalization type. Rhythm frequencies showed a strong right skewness, which is why common measures such as mean, or median would have been inaccurate. Therefore, we performed a cluster analysis to assess whether specific rhythm frequency clusters existed. We applied an agglomerative, hierarchical clustering algorithm which used the group average of frequency distances as a dissimilarity measure (dissimilarity threshold was set to 0.05 for all data sets). The frequency data were log10-transformed before clustering because an earlier study (Norton & Scharff, 2016) showed

that log10-transformation resulted in comparable clusters for different frequencies since these clusters had the least frequency-dependent standard deviation.

Modelling

To confirm that the rhythm frequencies obtained by the GAT approach are an inherent property of the respective vocalization type and cannot be found in arbitrary element sequences, we created artificial temporal vocalization patterns based on the previously measured element and gap durations, assessed their FRMSD values and compared them to the FRMSD values of the original vocalization types. We created two different types of artificial vocalization patterns that were used in different models: In Model 1 we used artificial vocalization patterns with randomized element and gap duration but intact sequence information (i.e. the correct order of consecutive elements). In Model 2, we used artificial vocalization patterns where each element and gap were replaced with a random duration, irrespective of element type and sequence. Model 2 did not apply to echolocation call sequences because they consisted of only one element type repeated in series, thus making the dismissal of sequence information pointless. Element and gap durations for both models were drawn randomly out of the pool of original recorded durations of the same type from all individuals (elements a–e and gaps following elements a–e respectively). The respective pool from which durations were drawn contained only element and gap durations of the vocalization type (isolation calls, territorial songs, or echolocation call sequences) to be modelled.

For each vocalization, we ran both Model 1 and 2 (not for echolocation call sequences, see above) 50 times. For every iteration, a new FRMSD value was obtained. We calculated the means of all model FRMSD values per individual and compared them to the means of all original FRMSD values per individual.

Fourier Analysis

Results of the GAT approach were compared to FFT analyses of all sequences (following (Norton & Scharff, 2016; Saar & Mitra, 2008)). Timestamps of element onsets were used to form a binary point process. We created strings with a time resolution of 5 ms in which only events (i.e. element onsets) were represented by ‘1’, everything else in the string was represented by ‘0’. The higher the temporal resolution of the input data, the lower the frequency resolution of the FFT output will be. With the sequence lengths available to us, a time resolution of 5 ms proved to be the best compromise between the two constraints. After calculating a fast Fourier analysis, frequencies of maximum power were selected as ‘best fitting rhythms and the pattern compared to GAT-results. A customized Matlab script was used for the analysis.

Statistics

Data distribution was assessed using a Shapiro-Wilks test for all datasets. Artificial data from both randomizations (1 and 2) were compared to original data with repeated measures ANOVA (Tukey’s post hoc comparison) for isolation calls and territorial songs. Echolocation call sequences were tested against randomization 1 via a Welch-corrected t-test because variances differed significantly. A paired t-test was used to compare the results of different ontogeny stages in isolation calls. Statistical

differences were considered significant for P<0.05 (*P<0.05, **P<0.01, ***P<0.001). When random numbers were needed those were generated using the R-function 'runif'.

Software

For analyses and preparing figures, we used Matlab (Version 2016b & 2015b), R (Version 3.5.1), GraphPad Version 5 and Avisoft SASLab Pro Version 5.2.10. Customized Matlab programs written by Philipp Norton (PN) and Lara Sophie Burchardt (LSB) were adjusted and used for the rhythm optimization (PN), model calculations (PN & LSB), FFT analysis (LSB) and cluster visualization (PN).

Results

Isochronous rhythm

For each vocalization, we found an isochronous rhythm (rhythm^S) that coincided best with the onsets of elements (supplementary audio files A1-3). A rhythm^S between 6 – 20 Hz dominated across individuals as well as across vocalization types: 49.4% of isolation calls (247 out of 500 calls), 41% of territorial songs (59 out of 143 songs) and 57% of echolocation call sequences (19 out of 33 sequences) had a best fitting rhythm of 6-20 Hz (Figure 4).

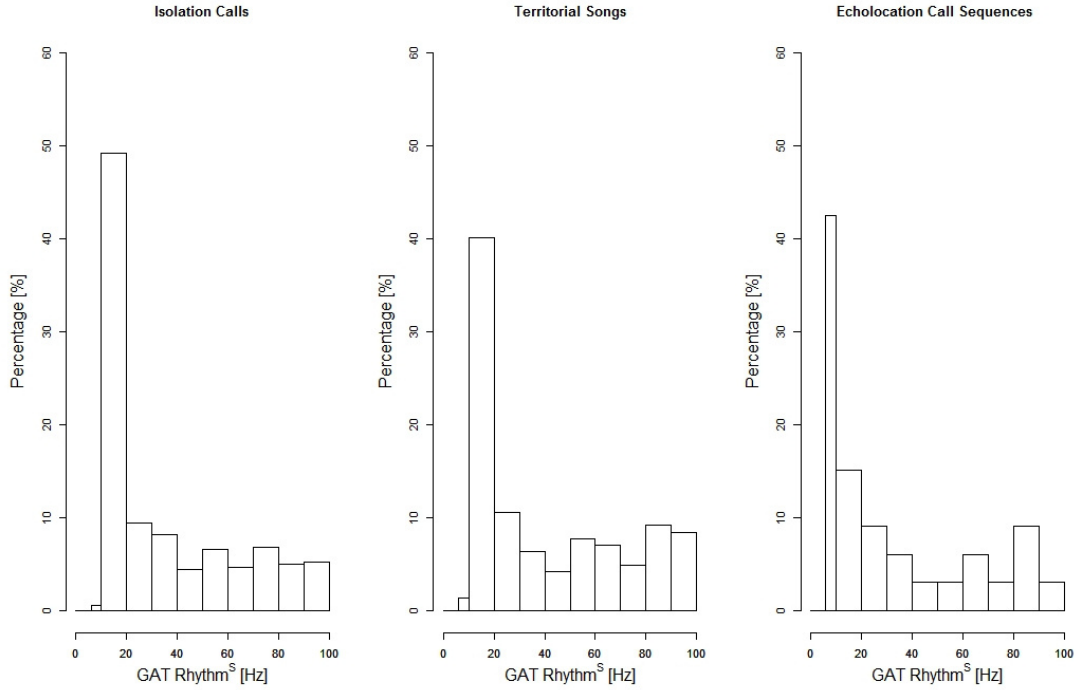


Figure 4: GAT Analysis
Regular rhythms in S. bilineata vocalizations. The relative majority of calls/songs occurred in rhythm frequencies below 20 Hz for all vocalization types.

Corresponding results were obtained when focusing on individuals instead of vocalization types. 20 out of 25 pups produced isolation calls which clustered predominantly in the frequency range of 10-20 Hz; the largest clusters contained 25-70% of calls per pup (Figure 5A). 9 out of 14 males produced territorial songs which clustered predominantly in the frequency range of 10-20 Hz; the largest clusters contained 30-60% of songs (Figure 5B). We considered clusters with their mean falling into the range between

10-20 Hz and the cluster comprising at least 25% of data (clusters are marked in red in Figure 5A-C). Echolocation call sequences clustered predominantly in the

frequency range of 6-20 Hz, 39% of sequences making up the strongest cluster (between 6 and 10 Hz), adding up to 57% between 6 and 20 Hz. Note that echolocation call sequences were pooled over all individuals (Figure 5C, see Methods).

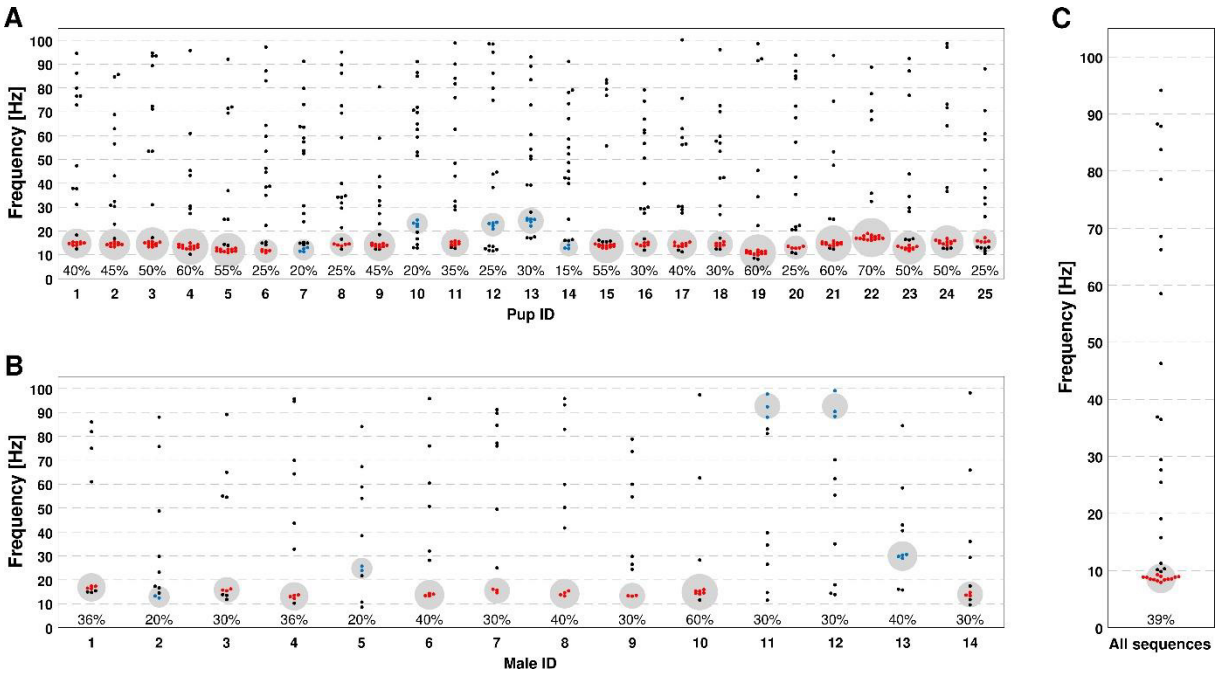


Figure 5: Isochronous Beat in Bat Vocalization
(A) rhythm clusters in isolation calls of S. bilineata pups; (B) rhythm clusters in territorial songs of S. bilineata males; (C) rhythm clusters in echolocation call sequences of S. bilineata adults. Marked in red are the data belonging to the largest cluster containing at least 25% of songs/calls, within the range of 6–20 Hz. Marked in blue are the data belonging to the largest cluster that were not considered. The percentage of data in the largest cluster is shown at the bottom of each column. The area of circles is scaled to the percentage of calls/songs in the respective clusters.

Comparison to artificially randomized vocalizations

To confirm that the observed element onsets in *S. bilineata* vocalizations aligned to an isochronous rhythm well and more closely than expected by chance, we compared the FRMSD values of artificial vocalization types to the FRMSD values of the original vocalization types. All artificial vocalization types had randomized element and gap durations; sequence information, i.e. the consecutive order of elements was either preserved (Model 1) or ignored (Model 2).

As expected, original vocalizations had significantly lower FRMSD values than artificial model 1 or model 2 vocalizations (Repeated Measures ANOVA: isolation calls: $F=71.17$, $df= 74$, $p<0.0001$; territorial songs: $F=30.38$, $df= 41$, $p<0.0001$; unpaired t-test (with Welch correction): echolocation call sequences: $t=2.35$,$df=33$, $p= 0.0023$), indicating that the element onsets of original vocalizations matched an isochronous rhythm more closely than expected by chance (Figure 6).

Fourier analysis

Results of the fast Fourier analysis of a binary point process string where element onsets were represented by '1' resulted in the same if not stronger picture at the level of vocalization types. 55.4%

of isolation calls, 47.8% of territorial songs and 66% of echolocation call sequences showed a dominant rhythm between 6 and 20 Hz (54% between 6-10Hz) (Supplements, Figure 7).

Ontogeny effect

Furthermore, we ran statistical analyses to investigate the effect of ontogeny on rhythmicity for pup isolation calls. For each pup, we compared the frequencies of isochronous rhythms of the first and last two isolation calls recorded during ontogeny (non-volant phase and volant phase). $Rhythm^S$ frequencies in isolation calls did not change significantly during the pups' ontogeny (paired t-test: t=1.31, df= 49 p=0.20, Supplements, Figure 8).

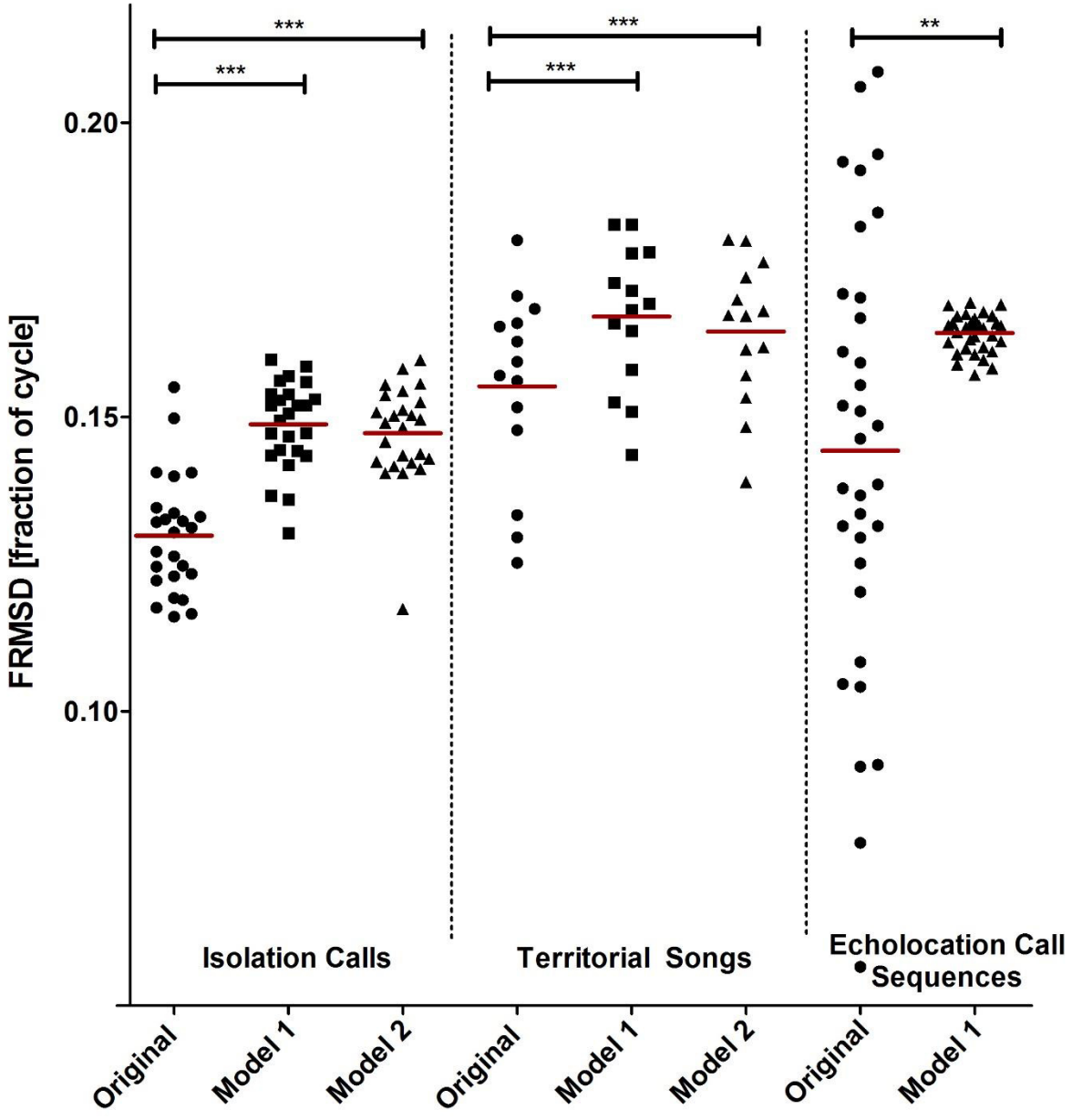


Figure 6: Model Validation
Mean values for FRMSD, comparing original data to 'bat-like' artificial data (Model 1: intact sequence information, Model 2: random sequences). Original data showed significantly lower deviations (*p , 0.05; **p , 0.01; ***p , 0.001). Depicted are means per individual for isolation calls

and territorial song, and best-fitting rhythms of single sequences for echolocation call sequences, explaining the higher spread. Red lines indicate the respective mean of a dataset.

Discussion

The novel aspect presented in this study is the documentation of isochronous rhythm patterns in different vocalization types of the bat *S. bilineata*. With a generate-and-test approach (GAT) as well as an FFT analysis, vocalizations were analysed to find a best fitting rhythm over a wide frequency range of 5-100 Hz (i.e. pulses per second). Even though the three analysed vocalization types (pup isolation calls, male territorial songs, and echolocation call sequences) differed in their acoustic structure and the behavioural situation they were produced in, their best-fitting rhythms fell in a quite narrow frequency window. Element onsets coincided best with rhythm frequencies between 6-20 Hz, independent of vocalization type and vocalizing individual. Analyses showed that rhythm frequencies were most abundant between 6-10 Hz for echolocation call sequences and between 10-20 Hz for territorial songs and isolation calls. The same picture was found with an FFT analysis at the level of vocalization types.

Therefore, the best fitting rhythms were comparatively similar across vocalization types and vocalizing individuals in *S. bilineata*, with social communication signals showing rhythms twofold of echolocation call sequences. Other studies on rhythmicality in animal vocalizations so far did show patterns that differed between individual animals (Norton & Scharff, 2016), and temporal structure, namely the rhythm, may be used by conspecifics for individual discrimination (Mathevon et al., 2017). A biological constraint shaping rhythms to be more alike between individuals is not apparent. Since there are not many comparable studies yet our results might prove to be the rule rather than an exception.

Nevertheless, the pattern of rhythm^{S} in the analysed vocalizations, could be caused by physiological constraints and/or mechanisms to save energy. The production of echolocation calls when a bat is searching for prey items but has not detected anything yet is correlated with respiration which, in turn, is tightly coupled to wing beat. For many bat species, a 1:1 relation has been found (e.g. (Suthers et al., 1972)). The soprano pipistrelle (*Pipistrellus pygmaeus*), for example, produces one or two echolocation calls per wingbeat and respiratory cycle (Wong & Waters, 2001). In other pipistrelle bats (*P. pipistrellus, P. kuhlii, P. nathusii* (Kalko, 1994)), greater horseshoe bats (*Rhinolophus ferrumequinum*), little brown bats (*Myotis lucifugus*), Parnell's mustached bats (*Pteronotus parnellii rubiginosus*) and Seba's short-tailed bats (*Carollia perspicillata* (Schnitzler, 1971)) wingbeat and echolocation calls are also coupled. Coupling was also found in the tongue-clicking Pteropodid bat *Rousettus aegyptiacus* , indicating that a strong coupling of wing beat, respiration and sonar emission is widespread in bats regardless of sound production mechanism.

In *S. bilineata,* respiratory cycle and wing beat are between 6 and 12 Hz during search flight (pers. communication H.-U. Schnitzler). Our results suggest that in the release situation the echolocation call sequences were recorded, bats mainly uttered one call per wingbeat, which fits the low sensory needs in the relatively open space in which releases took place. In a situation with higher sensory needs,

expected rhythm frequencies should be doubled, i.e. lie between 12 and 24 Hz, most of which overlaps strongly with the rhythm frequencies found in the social vocalizations. Therefore, we argue that the rhythm frequencies most abundant in social vocalizations (10-20 Hz) and in echolocation call sequences during search flight (6-10 Hz and, to a lesser degree, 10-20 Hz) can be regarded as comparatively similar.

During prey capture, however, echolocation call sequences contain not only search flight calls but also approach flight calls (when prey has been detected and is approached) and a so-called final buzz (immediately before prey capture, very short and broadband echolocation calls with extremely short IOIs are produced), which enhances the sensory information available for the foraging bat. Even though wing beat, respiratory cycle and sonar emission are tightly coupled during search flights (likely to increase energy efficiency), this might not provide sufficient sensory information during prey capture, i.e. in a situation where high temporal resolution is needed (per wing beat and respiratory cycle up to 10-15 pulses can be emitted (Kalko, 1994). A larger ratio between wing beat, respiratory frequency, and emitted echolocation calls could result in a weaker rhythmic pattern in our analyses. In the approach phase, the number of echolocation calls per wing beat can vary widely, depending on the current sensory needs of a foraging bat. Therefore, it seems reasonable to assume that echolocation call sequences during prey capture do not follow any clear rhythm but strongly depend on the bats' current sensory needs. This could easily be tested on echolocation call sequences recorded in foraging situations. Correspondingly, a previous study on the big brown bat *Eptesicus fuscus* showed that the strict 1:1 synchronization of wing beat, respiration, and call emission was not found during complex navigation tasks, where freely behaving individuals had to search for prey (tethered mealworms, suspended at about 1.5 m height) in a flight room, equipped with various obstacles, such as artificial houseplants (Moss et al., 2006). During search flights, however, metabolic needs, e.g. being energy efficient, may play a more important role (Suthers & Fattu, 1973). To investigate the task/situation dependence of the coupling of wing beat, respiration, and call emission it would be worthwhile to analyse rhythm^{S} of echolocation call sequences produced in a feeding context in bat species in which a strict 1:1 coupling has been found during search flight.

The determination of rhythm^{S} (method developed by (Norton & Scharff, 2016)) could be a valuable addition to currently used methods since it is not dependent on a laboratory setting. Knowledge of wing beat, and/or respiratory rates could be combined with analyses of rhythm^{S} of echolocation call sequences and social vocalizations recorded from freely behaving, wild bats to gain insights on coupling relations in natural situations. Especially for more complex vocalization types with variable element and gap durations, the GAT approach and FFT analysis provide a more detailed picture than simply assessing IOIs. The latter method ignores the sequential structure of vocalizations and their variable element durations, potentially concealing higher order regularity.

To assess the goodness-of-fit for our analyses of rhythm^{S}, we compared deviations from rhythm^{S} of original and artificially created vocalizations that were randomly drawn from a pool of typical element and gap durations. Original vocalizations deviated significantly less from rhythm^{S} than did artificial

vocalizations (i.e. element onsets of original vocalizations coincided significantly better with an isochronous rhythm than artificial vocalizations), indicating that the rhythm^S found in *S. bilineata* vocalizations was not an artifact of the typical duration and sequence of this species' vocalizations.

One aspect worthy of discussion is the relation between rhythm frequencies of echolocation call sequences produced by *S. bilineata* during search flight (which were coupled to wing beat frequencies) and social vocalizations produced by individuals hanging in their day-roost (pup isolation calls and male territorial songs). We doubt that rhythm frequencies of isolation calls and territorial songs are caused by a coupling of sound emission to respiration since echolocation calls produced by roosting bats can occur at any point in the respiratory cycle (Suthers et al., 1972). Taken this into account, it seems reasonable to assume that social calls can be emitted at any point in the respiratory cycle as well. Nevertheless, as stated before, we argue there is a relation between the dominant frequencies of the three vocalization types, and we regard them as being comparatively similar. The similarity of rhythm frequencies could suggest a common evolutionary background, which might be the coupling between respiration, wingbeat and echolocation call emission. However, increasing evidence suggests that flight preceded echolocation (Simmons & Geisler, 1998; Speakman, 2001), which would indicate that vocal communication preceded echolocation as well (assuming that bats' predecessors communicated with social calls, as many small mammals do). It is therefore possible that social calls, despite being probably phylogenetically older than echolocation, adopted the rhythm frequencies of echolocation calls at some point.

It is interesting to compare the strength of rhythms between isolation calls and territorial songs since isolation calls are produced within minutes after birth (Knörnschild & von Helversen, 2008) while territorial songs are learned during ontogeny (Knörnschild et al., 2010). Generally, a higher variability in rhythm^S may be expected when comparing learned vocalizations to innate vocalizations. In our study, rhythm frequencies predominantly clustered between 6-20 Hz, but cluster strength of individuals was on average lower in territorial song than in isolation calls (37% in territorial song compared to 44.7% in isolation calls; GAT approach). This difference in individual cluster strength resembled the overall difference between both vocalization types, since only 41.6% of all territorial songs had rhythm frequencies between 6-20 Hz, while 49.8% of isolation calls did.

Rhythmic properties of echolocation could represent the same neuronal correlates underlying production of social vocalizations. In the Egyptian fruit bat (*R. aegyptiacus)* wingbeat and tongue clicks are tightly coupled around 10 Hz (Yartsev & Ulanovsky, 2013), as we found for *S. bilineata*. These rhythm frequencies show a resemblance to the frequency of theta brain waves. Thought to be important for movements, spatial memory and active stimulus intake(Colgin, 2013) amongst others, theta waves might be a promising neural correlate explaining the production of the detected rhythms.

It might be advantageous to produce rhythmic vocalizations because 'rhythmic attention' (*sensu* (Jones et al., 1981)) helps receivers to decode rhythmic signals easier and faster (Rohenkohl et al., 2012). The attention of receivers cycles in an oscillatory way when a rhythm exists (e.g. (Barnes & Jones, 2000; Large & Jones, 1999)). Since rhythmic signals are predictable, 'rhythmic attention' enables receivers to

provide most 'attentional energy' at a time point where the next stimulus is to be expected. This is advantageous because cognitive capacities are limited (Shapiro et al., 1997) and an optimization of attention timing is helpful to not miss relevant stimuli. For example, when humans were asked to assess the difference in pitch of two focal tones separated by regularly timed tones, the assessment of pitch difference was better when the second focal tone followed the regular timing of the separating tones than when was slightly displaced from the regular timing (Jones et al., 2002). Another example from macaques shows that neuronal oscillations in the primary visual cortex entrain to a stimuli stream (visual stimuli) when the stream is rhythmic, a mechanism resulting in decreased reaction time and an increase in the response gain for events that are task relevant (Lakatos et al., 2008a). Bats' attention as well as the auditory system collectively could be tuned to echolocation rhythms, because bats are exposed to those rhythms for large parts of their lives(Fenton, 2003). Therefore, it might be advantageous to produce vocalizations in the same frequency window to increase detection by receivers. At the moment, we do not know whether rhythmic attention plays a role in *S. bilineata*. Playback experiments violating expected rhythmic patterns in social vocalizations or direct assessment of the animals' rhythm perception would be a valuable avenue for future research. A switch from a rhythm determined by physiological constraints to a rhythm decoupled from its original production constraints but still with an adaptive function (e.g. rhythmic attention) might have been one step during evolution that paved the way to develop music as we know it.

In summary, this study demonstrates an isochronous rhythm in three bat vocalization types in which metabolic constraints leading to rhythmic patterns are more (echolocation calls) or less (isolation calls, territorial songs) likely. The two methods used in this study (GAT and FFT) enable the analysis of best fitting rhythms in a corresponding way. Future studies should profit by complementary use of both methods in addition to IOI assessment. To further study the coupling or decoupling of wing beat, respiration, and sound emission in animals as well as its biological relevance, it would be highly beneficial to compare different species of bats and birds which sing in flight as well as other echolocating mammals. Such a comparative approach could provide valuable insights into the origin and relevance of rhythmicality in animals (Kotz et al., 2018).

Ethics

All experiments and protocols for capturing and handling bats comply with the current laws of Costa Rica. Permit numbers are given in the original publications from which the data were drawn (Behr et al., 2006; Knörnschild, Jung, et al., 2012a; Knörnschild, Nagy, et al., 2012).

Data Accessibility

The dataset supporting this article has been uploaded as part of the electronic supplementary material. The code (GAT approach) with detailed explanations was already published in a previous publication (Ravignani & Norton, 2017).

References

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention and Time. *Cognitive Psychology*(41), 254-311. https://doi.org/10.1006/cogp.2000.0738

Bayefsky-Anand, S. (2006). Echolcation calls of the Greater Sac-winged Bat (*Saccopteryx bilineata*) in different amounts of clutter. *Bat Research News, 47*(1), 7-10.

Behr, O., & Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioural Ecology and Sociobiology, 56*, 106-115. https://doi.org/10.1007/s00265-004-0768-7

Behr, O., Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., & Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behavioural Ecology, 17*, 810-817. https://doi.org/10.1093/beheco/arl013

Bermeitinger, C., & Frings, C. (2015). Rhythm and attention: Does the beat position of a visual or auditory regular pulse modulate T2 Detection in the attentional blink? *Frontiers in Psychology*(6). https://doi.org/10.3389/fpsyg.2015.01847

Celma-Miralles, A., & Toro, J. M. (2018). Beat perception in a non-vocal learner: rats can identify isochronous beats. Proceedings of the 12th International Conference on the Evolution of Language,

Colgin, L. L. (2013). Mechanisms and functions of theta rhythms. *Annual Review of Neuroscience, 36*, 295-312. https://doi.org/10.1146/annurev-neuro-062012-170330

Eckenweber, M., & Knörnschild, M. (2013). Social influences on territorial signaling in male greater-sac winged bats. *Behavioural Ecology and Sociobiology, 67*(4), 639-648. https://doi.org/10.1007/s00265-013-1483-z

Fenton, M. B. (2003). Eavesdropping on the echolocation and social calls of bats. *Mammal Review, 33*, 193-204. https://soi.org/10.1046/j.1365-2907.2003.00019.x

Fernandez, A. A., & Knörnschild, M. (2017). Isolation Calls of the Bat *Saccopteryx bilineata* Encode Multiple Messages. *Animal Behavior and Cognition, 4*(2), 169-186. https://doi.org/10.12966/abc.04.05.2017

Heinsohn, R., Zdenek, C. N., Cunningham, R. B., Endler, J. A., & Langmore, N. E. (2017, Jun). Tool-assisted rhythmic drumming in palm cockatoos shares key elements of human instrumental music. *Sci Adv, 3*(6), e1602399. 10.1126/sciadv.1602399

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B, 370*(1664). http://dx.doi.org/10.1098/rstb.2014.0088

Hulse, S. H., Humpal, J., & Cynx, J. (1984). Discrimination and Generalization of Rhythmic and Arrhythmic Sound Patterns by European Starlings (Sturnus vulgaris) [10.2307/40285272]. *Music Perception: An Interdisciplinary Journal, 1*(4), 442. http://mp.ucpress.edu/content/1/4/442.abstract

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evicence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7*(5), 1059-1073.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*(13), 313-319. https://doi.org/10.1111/1467-9280.00458

Jung, K., Kalko, E. K. V., & von Helversen, O. (2007). Echolocation calls in Central American emballonurid bats: signal design and call frequency alternation. *Journal of Zoology, 272*, 125-137. https://doi.org/10.1111/J.1469-7998.2006.00250.X

Kalko, E. K. V. (1994). Coupling of sound emission and wingbeat in naturally foraging european pipistrelle bats (Microchiroptera: Vespertilionidae). *Folia Zoologica, 43*(4), 363-376.

Knörnschild, M., Bluml, S., Steidl, P., Eckenweber, M., & Nagy, M. (2017). Bat songs as acoustic beacons - male territorial songs attract dispersing females. *Scientific Reports, 7*(1), 13918. https://doi.org/10.1038/s41598-017-14434-5

Knörnschild, M., Jung, K., Nagy, M., Metz, M., & Kalko, E. K. V. (2012). Bat echolocation calls facilitate social communication. *Proceedings of the Royl Society B*(279(1748)), 4827-4835.

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters, 6*(2), 156-159.

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Animal Behaviour, 84*, 761-769. http://dx.doi.org/10.1016/j.anbehav.2012.06.029

Knörnschild, M., & von Helversen, O. (2008). Non-mutual vocal mother-pup recognition in the sac-winged bat, *Saccopteryx bilineata*. *Animal Behaviour, 76*(3), 1001-1009. https://doi.org/10.1016/j.anbehav.2008.05.018

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences, 22*(10), 896-910. https://doi.org/10.1016/j.tics.2018.08.002

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*(320(5872)), 110-113. 10.1126/science.1154735

Large, E. W., & Jones, M. R. (1999). The dynamics of attenting: how people track time varving events. *Psychological Review, 106*, 119-159.

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals Voices. *Current Biology, 27*(15), 2352-2356.e2352. https://doi.org/10.1016/j.cub.2017.06.035

Merguerditchian, A., Vuillemin, A., & Pruetz, J. D. (2018). *Identifying the ape beat in the wild: rhythmic indvidual signatures from sounds of manual fruit cracking in Fongoli Chimpanzees* The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12),

Moss, C. F., Bohn, K., Gilkenson, H., & Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *PLOS Biology, 4*(4), e79. https://doi.org/10.1371/journal.pbio.0040079

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

Patel, A. D. (2008). *Music, Language and the Brain*. NY: Oxford University Press.

Ratcliffe, J. M., Jakobsen, L., Kalko, E. K. V., & Surlykke, A. (2011). Frequency alternation and an offbeat rhythm indicate foraging behavior in the echolocating bat, *Saccopteryx bilineata*. *Journal of Comparative Physiology A*. https://doi.org/10.1007/s00359-011-0630-0

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01118

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speech, movement and animal vocalizations. *Journal of Language Evolution*. https://doi.org/10.1093/jole/lzx002

Ravignani, A., Thompson, B., & Filippi, P. (2018). The Evolution of Musicality: What Can Be Learned from Language Evolution Research? *Frontiers in Neuroscience, 12*(20). https://doi.org/10.3389/fnins.2018.00020

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *The Journal of Neuroscience, 32*, 8424-8428. https://doi.org/10.1523/JNEUROSCI.0804-12.2012

Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One, 3*(1), e1461. https://doi.org/10.1371/journal.pone.0001461

Schnitzler, H. U. (1971). Fledermäuse im Windkanal. *Zeitschrift für vergleichende Physiologie, 73*, 209-221.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Science, 1*, 291-295.

Simmons, N., & Geisler, J. H. (1998). Phylogenetic relationships of Icaronycteris, Archaeonycteris, Hassianycteris, and Palaeochiropteryx to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the American Museum of Natural History, 235*, 4-169.

Speakman, J. R. (2001). The evolution of flight and echolocation in bats: another leap in the dark. *Mammal Review, 31*(2), 111-130. doi:10.1046/j.1365-2907.2001.00082.x

Suthers, R. A., & Fattu, J. M. (1973). Mechanisms of Sound Production by Echolocating Bats. *American Zoologist, 13*(13), 1215-1226.
Suthers, R. A., Thomas, S. P., & Suthers, B. J. (1972). Respiration, Wing-beat and Ultrasonic Pulse Emission in an Echo-locating Bat. *The Journal of Experimental Biology, 56*, 37-48.
Wallin, N. L. (1991). *Biomusicology - Neurophysiological, Neuropsychological, and Evolutionary Perspectives on the Origins and Purpose of Music*. Pendragon Press.
Wong, J. G., & Waters, D. A. (2001). The Synchonisation of signal emission with wingbeat during the approach phase in Soprano Pipistrelles (*Pipistrellus Pygmaeus*). *The Journal of Experimental Biology*(204), 575-583.
Yartsev, M. M., & Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. *Science, 340*, 367-372. https://doi.org/10.1126/science.1235338

Supplementary Information

This supplementary information contains:

Supplementary Audio Files A1-A3

Validation Analyses – Tempo Changes

Supplementary Figures S1-S4

Supplementary Audio Files

https://royalsocietypublishing.org/doi/suppl/10.1098/rsos.181076

Audio File A1: rsos181076supp1.wav Isolation call of *S. bilineata* pup overlaid with an isochronous rhythm of 12.3Hz. The recording is slowed down to 10% of its original speed.

Audio File A2: rsos181076supp2.wav Territorial song of *S. bilineata* male overlaid with an isochronous rhythm of 15Hz. The recording is slowed down to 10% of its original speed.

Audio File A3: rsos181076supp3.wav Echolocation call sequence of adult *S. bilineata* overlaid with an isochronous rhythm of 8.8 Hz. The recording is slowed down to 20% of its original speed.

Validation analysis - Tempo changes

To assess tempo changes within vocalizations, we calculated linear regressions for the Inter-Onset-Intervals (IOI) sequences to test whether these were significantly different from zero (using an F-test), which would indicate a significant change in tempo. We conducted this analysis for a random subset of each vocalization type (two sequences per individual for isolation calls and territorial songs, all data for echolocation call sequences). To corroborate results from the tempo analysis, individual syllable deviations of the first, middle and last syllable were compared per vocalization type by means of a

Friedman test; this was done to test whether deviations changed throughout a syllable sequence. This analysis was conducted on a subset of the data, chosen in the same way as for the tempo analysis.

The majority of isolation calls (74%) had a stable tempo, 22% of calls showed a decrease in tempo and 4% of calls an increase. On the contrary, the majority of territorial songs (79%) decreased in tempo, especially in the last fifth of songs (Supplementary Figure S3). However, inter-onset intervals did not increase continuously but rather abruptly, often doubling and quadrupling. These multiples of inter-onset intervals make it unlikely that the observed change in tempo had a negative effect on $rhythm^S$ in our study. Furthermore, results were confirmed by the FFT analysis, which is stable against tempo changes.

To corroborate that changes in tempo did not affect $rhythm^S$, we calculated individual element deviations to the nearest single pulse. Element deviations did not change throughout vocalizations, since a best fitting rhythm was found by an optimization task regarding all elements of a sequence. Nevertheless, individual element deviations of vocalizations with tempo changes (territorial songs) did not differ from vocalizations without tempo changes (isolation calls) (Kruskal-Wallis, $p=0.78$, $F=2.47$, $df=6$, Supplementary Figure S4), suggesting that changes in tempo played a negligible role in our study. Another argument for this interpretation is the results from FFT analysis. Since the same pattern was found with a method in which tempo changes cannot affect the outcome, it is reasonable to say that tempo changes did not influence the results from GAT analysis in a crucial way.

Supplementary Figures

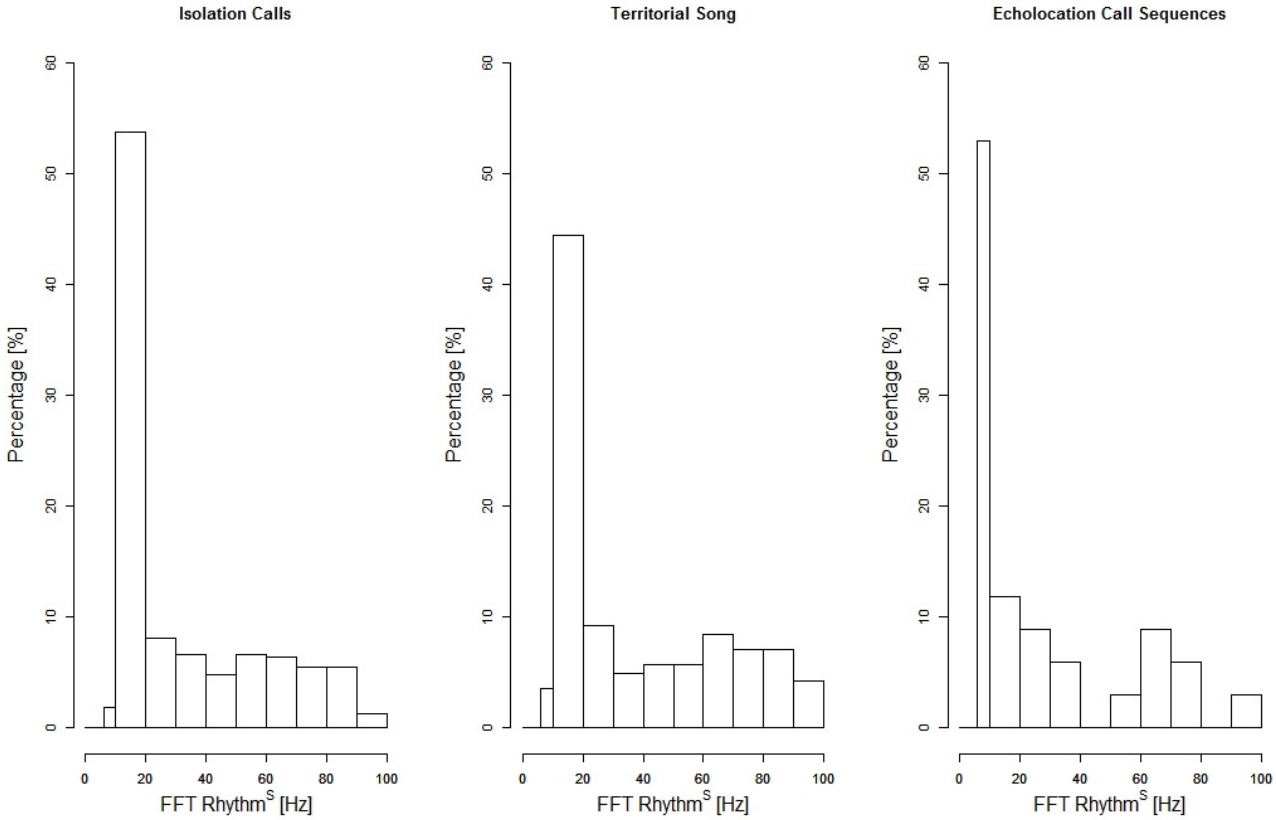


Figure 7: Fourier Analysis
Regular rhythm^S in bat vocalizations. The relative majority of calls/songs occur in rhythm frequencies below 20 Hz for all vocalization types.

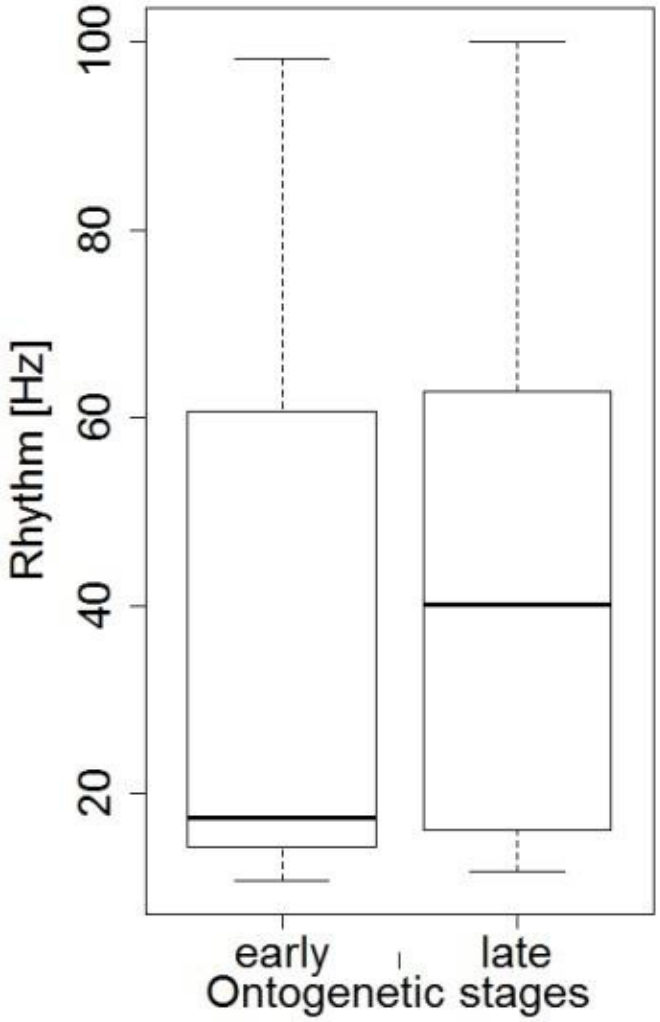


Figure 8: Effect of Ontogenetic Stage on Rhythm in Pup Isolation Calls.
Early ontogeny did not differ from late ontogeny. Medians, interquartile range (25-75%) and whiskers (0-100%) are shown.

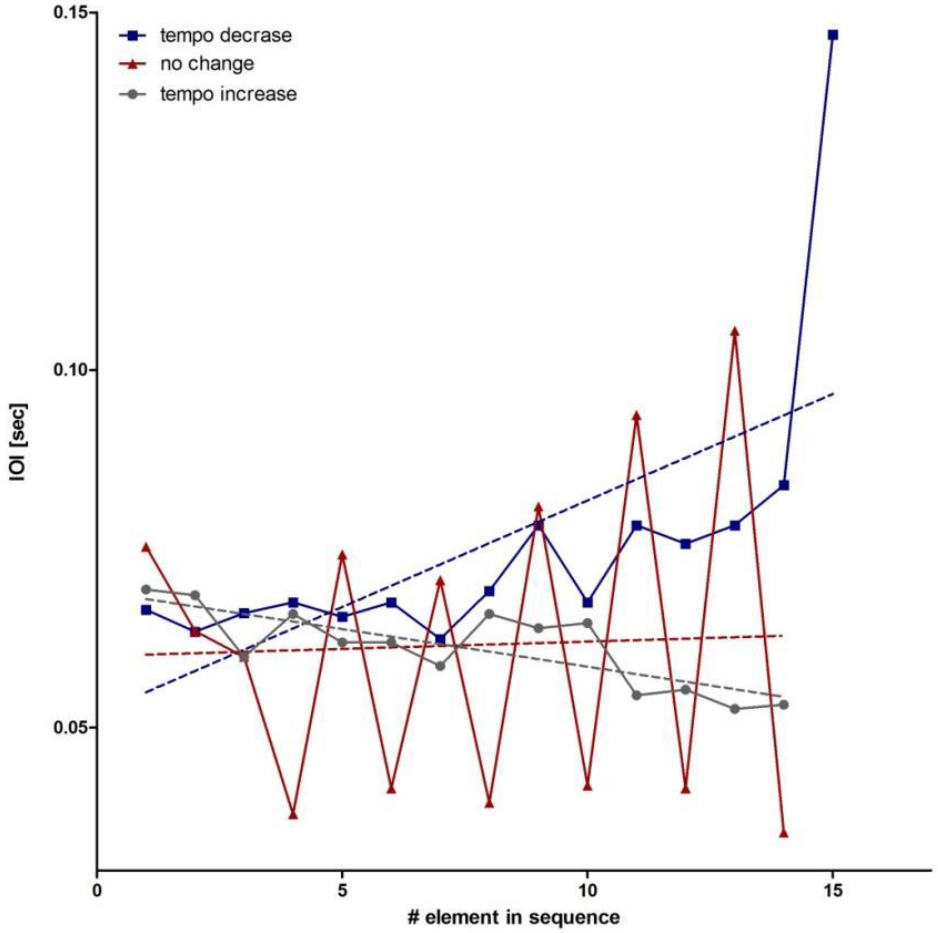


Figure 9: Tempo Changes in Sequences.
Three IOI sequences are shown as solid lines; dashed lines show corresponding linear regressions. Slopes of regression lines were tested against zero. Significant difference from zero was interpreted as tempo change. In red (triangle) an isolation call with no tempo change is shown, in grey (circle) an isolation call increasing in tempo and in blue (square) a territorial song decreasing in tempo rather abruptly are shown.

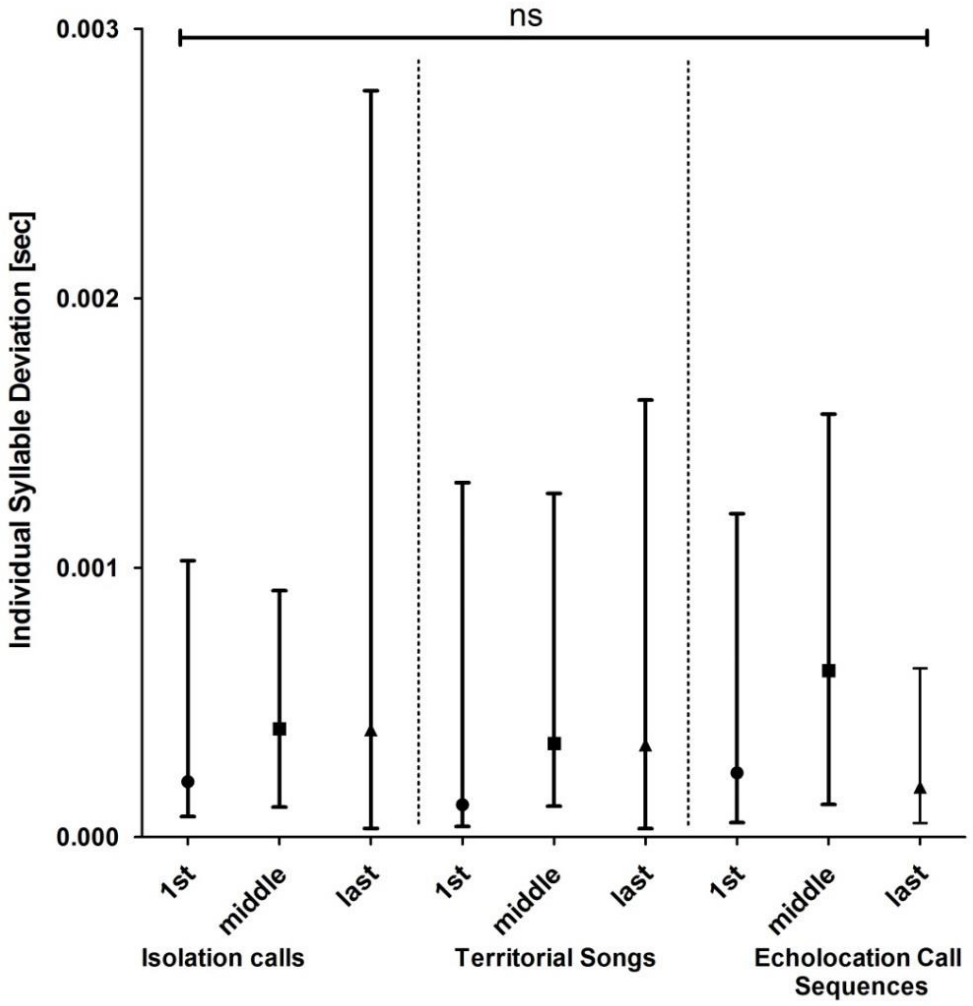


Figure 10: Syllable Deviation of Individual Syllables.
Individual deviations from rhythm^{S} of first, middle and last syllable of calls/songs were compared. Median and interquartile range are shown. No significant differences were found.

Chapter II

Comparison of Methods for Rhythm Analysis of Complex Animals' Acoustic Signals

Adapted from: Burchardt, L.S.; Knörnschild, M. (2020): Comparison of methods for rhythm analysis of complex animals' acoustic signals. In: PLoS Computational Biology 16(4): e1007755.

https://doi.org/10.1371/journal.pcbi.1007755

Abstract

Analysing the rhythm of animals' acoustic signals is of interest to a growing number of researchers: evolutionary biologists want to disentangle how these structures evolved and what patterns can be found, and ecologists and conservation biologists aim to discriminate cryptic species on the basis of parameters of acoustic signals such as temporal structures. Temporal structures are also relevant for research on vocal production learning, a part of which is for the animal to learn a temporal structure. These structures, in other words, these rhythms, are the topic of this paper. How can they be investigated in a meaningful, comparable and universal way? Several approaches exist. Here we used five methods to compare their suitability and interpretability for different questions and datasets and test how they support the reproducibility of results and bypass biases. Three very different datasets with regards to recording situation, length and context were analysed: two social vocalizations of Neotropical bats (multisyllabic, medium long isolation calls of *Saccopteryx bilineata*, and monosyllabic, very short isolation calls of *Carollia perspicillata)* and click trains of sperm whales, *Physeter macrocephalus*. Techniques to be compared included Fourier analysis with a newly developed goodness-of-fit value, a generate-and-test approach where data was overlaid with varying artificial beats, and the analysis of inter-onset-intervals and calculations of a normalized Pairwise Variability Index (nPVI). We discuss the the advantages and disadvantages of the methods and we also show suggestions on how to best visualize rhythm analysis results. Furthermore, we developed a decision tree that will enable researchers to select a suitable and comparable method on the basis of their data.

Author summary

In the analysis of animal communication more and more interest is shown in rhythm of animal communication and what information this might convey. In this paper, we establish a workflow to analyse the temporal structure – namely the rhythm – of any particular animals' acoustic signal with methods that are applicable for a wide range of signals and results that are easily comparable and interpretable. This workflow will enhance the understanding of rhythmicality in animals' acoustic signals as well as facilitate comparison between species. Methods we conducted ranged from simple distributional and visual analysis to higher mathematics such as Fourier analysis. All analyses rely on

Inter-Onset-Intervals, the duration between the beginning of one element and the next. We used different datasets from two neotropical bat species as well as from the sperm whale. With this selection, we cover very short sequences with only few elements up to sequences of around 200 elements, multisyllabic and monosyllabic sequences and social communication as well as sounds used for orientation and foraging.

Introduction

Rhythms can be found anywhere in the world: our hearts have rhythms, circadian rhythms are all around, music across all cultures shares certain components such as rhythm, public transportation (should) follow a certain schedule which in fact is nothing but rhythm. We learn more and more about how important a certain temporal structure is in human language, in their production as well and probably even more so in their perception; stuttering, for example, is most likely connected to a misfunction of rhythm perception (Wieland et al., 2015). This raises the question of whether rhythms, or temporal structures to use a more precise terminology, play an equally important role in animal communication and sound production. Can we learn something about rhythm in animals that will help us understand their communication better and also find underpinnings of the abundance of rhythm in human biology and culture?

Rhythm has a very narrow definition in musicality studies that does not necessarily fit the focus of this paper. We are describing temporal structures and are searching for periodicity. To prevent confusion and since terms might be used in different contexts depending on the research area, we define some key terms in a glossary (Table 2). Nevertheless, we still use the term ‘rhythm’ as a concept that will be understood by a broad audience, as most people have an intuitive understanding of ‘rhythm’, independent of whether this study analyses ‘rhythm’ in the musicological sense of the term.

The rhythmicality of animals’ acoustic signals has an impact on a vast field of related questions. For instance, the evolution of music is investigated in the field of biomusicology, a research area that studies musicality in animals – where musicality is used as a term for different traits that occur spontaneously and are based on and constrained by biology and cognition in an animals’ acoustic signals, such as harmony, timbre or rhythm (Ravignani et al., 2018; Wallin, 1991). Moreover, knowledge about temporal structures is necessary to find coupled biological processes, such as the correlation between beat frequencies in bat’s acoustic signals (also called vocalizations) with their wingbeat frequencies (i.e.

wingbeats per second), independent from whether a bat might actually be flying in a vocalizing context or hanging in a roost (Burchardt et al., 2019). Rhythmicality might also influence mate choice and individual recognition (Mathevon et al., 2017; Norton & Scharff, 2016). Furthermore, neural correlates might play a role so that careful rhythm analysis can give insights into internal clocks or the importance of certain brain waves on different behavioural aspects such as the production of acoustic signals(Norton & Scharff, 2016; Yartsev & Ulanovsky, 2013). Rhythm analysis can also be used to disentangle cryptic species (distinct species that are combined under one species name, because they cannot be distinguished morphologically) that produce sounds in different rhythms (David et al., 2003) or is informative in the context of vocal production learning, a part of which is for the animal to learn the correct temporal structure of a signal (Wirthlin et al., 2019). A growing body of research is addressing questions on rhythm in animal vocalizations and animal sounds (in contrast to vocalizations, sounds are produced by something other than vocal cords, e.g. sperm whale clicks; both are combined under the term acoustic signals). But before we can elaborate on this, it is important to again note different connotations of rhythm in this context. Where we speak of rhythms in animals' acoustic signals a musicologist might only talk about different beats and tempi. What we mean in this paper with rhythm and the connotation of rhythm used in other studies on the subject (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017) describes a temporal structure that might have varying complexity but is mostly based on an isochronous beat (i.e. sounds produced by a metronome). These isochronous beats might be produced in different tempi by different species and individuals. Therefore, one could also say, we search for periodicity in animals' acoustic signals. The definition for periodicity we use here is the following: we regard a sequence as periodic, when there is an underlying isochronous pattern describing it. An isochronous rhythm is a metronome like rhythm with the same beat and the same gap length' (although beat and gap length are not necessarily similar). Not every beat of that isochronous sequence needs to be corresponding with an element in the sequence that is analysed. A beat here is every element of the isochronous pattern. It is also the actual 'beat frequency' of the isochronous rhythm. We refrain from using the word 'pulse', to prevent confusion with the use of the word 'pulse' in echolocation research. Keeping these definitions in mind, we are still using the term "rhythm" as a summary of these concepts in the text for reasons of readability and understanding.

Exemplary studies on the rhythmic production of acoustic signals come from male zebra finches (*Taeniopygia guttata*) (Norton & Scharff, 2016), the bat *Saccopteryx bilineata* (Burchardt et al., 2019) or the humpback whale (*Megaptera novaeangliae*) (Schneider & Mercado, 2018a). While male zebra finches sing with different rhythms depending on the individual, *S. bilineata* vocalizations share a common temporal structure, likely coupled to wingbeat frequencies (Burchardt et al., 2019; Norton & Scharff, 2016). Yet another pattern was found in the song of humpback whales (*Megaptera novaeangliae*), where individuals can produce very stable temporal structures or sound sequences that vary rapidly in tempo and rhythm (Schneider & Mercado, 2018a).

Other forms of rhythm production were found in the palm cockatoo (*Probosciger aterrimus*). The males of this species drum quasi-isochronous patterns, using tools, in a consistent manner (Heinsohn et al., 2017b). Chimpanzees use individual rhythm signatures - likely in a fashion to help recognize unseen companions – when cracking baobab fruits (Merguerditchian et al., 2018).

Table 2: Glossary with Important Terms and Concepts

Glossary	
Animals' acoustic	All acoustic signals that animals produce on purpose
Animal	The entirety of sounds and vocalizations animals produce willingly to communicate with each other
Vocalization	A sound produced on purpose; sound origin: vocal cords; a species can have various vocalization types
Social vocalization	A vocalization uttered in a social context, e.g. isolation calls
Isolation call	Uttered by pups to solicit maternal/paternal care
Animal Sounds	Willingly produced sounds by animals, with another origin than vocal cords, e.g. whales produce their sounds not with vocal cords; the term vocalizations could be misleading in this context
Musicality	different traits that occur spontaneously and are based on and constrained by biology and cognition in an animals' acoustic signals, such as harmony, timbre or rhythm
Rhythm	e.g. an ordered and recurrent alternation of different elements in a sequence of sound and silence in speech, music or animals' communication
Periodicity	underlying reoccurring pattern describing as sequence as periodic, e.g. an isochronous pattern
Isochrony	A stereotyped pattern with same beat and same gap length (gaps and beats do not have to be the same length as well, though), a metronome like acoustic pattern/beat
Heterochrony	A pattern with more than one underlying beat
Beat	The unit to describe an isochronous pattern, given in Hertz (beats per second); a beat frequency of 5 Hz would describe a sequence with an underlying pattern of 5 beats per second, i.e. 5 vocalizations per second or a temporal structure where elements are distributed regularly in a way that you could fit a maximum of 5 element into one second
Inter-Onset-Interval	In a sequence of acoustic signals, the time span between the start of an element and the next element, comprising the element duration and the following gap duration; in other contexts, also called Inter-Pulse-Interval, Inter-Click-Interval or Inter-Call-Interval

Element	The smallest subunit of a sequence of acoustic signals, i.e. a distinct syllable, call, click, pulse etc. surrounded by silence
Exact beat frequency	The beat frequency we calculated to describe a specific sequence best (e.g. 5 Hz as in 5 beats per second)

Studies on the perception of rhythms or periodicity deal for example with the ability of animals to discriminate rhythms, e.g. in rats and European starlings (Celma-Miralles & Toro, 2018a; Hulse et al., 1984a). Moreover, the first instance for a biologically relevant rhythm in non-human mammalian acoustic signals was found in the northern elephant seal, where males can discriminate between familiar and unfamiliar male opponents using the temporal structure of vocalizations. Rhythms apparently differ between individuals in a way that facilitates the discrimination of individuals (Mathevon et al., 2017).

With the growing body of studies and its implications for other research questions, it is important to present methods in a reproducible way and find methods that are applicable to a vast majority of datasets in which temporal structures can be analysed. Reproducibility, interpretation biases, p-hacking (the distortion or manipulation of results through data mining) and apophenia (the tendency to see a pattern in random data) are key issues in all research fields. Defining clear methodologies with open access to code and data is one way of tackling those issues (Munafó et al., 2017). Results must be clearly structured and comparable between species and contexts. A number of papers address these issues and describe suitable methods by means of artificial data, with a decision tree depending on the respective question (Ravignani & Norton, 2017). Nevertheless, a comparison of different methods on different original datasets and of the influence of differences in datasets on the decision for a method is missing, even though this would help researchers to choose which methods to use for their data depending on the question at hand.

Acoustic recordings can differ enormously in their features. Depending on the recording situation and signals to record, one faces very different sampling rates and recording lengths. Moreover, the number of elements (i.e. a distinct syllable, call or click in a given sequence, surrounded by silence) in a recording differs greatly as well as element durations, noise level or amplitudes. Also, the recording situation differs a lot between a zebra finch recorded in a controlled recording box, a whale tagged with a recording device in the Pacific Ocean or a bat vocalizing in its roost. Nevertheless, all these acoustic signals are suitable and interesting to check for periodicity (or rhythmicality). It is crucial that a

comparable method can be applied to all these different recordings. Methods that have been used for rhythm analysis include Fourier analysis (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017; Saar & Mitra, 2008) or calculation of nPVIs, the normalized Pairwise Variability Index, which was originally developed to assess temporal variability in human speech rhythm (Grabe & Low, 2002b; Lin & Wang, 2007; Ravignani & Norton, 2017; Toussaint, 2012, 2013) . The nPVI is a measure of variability between Inter-Onset-Intervals. It will be zero for a perfectly isochronous sequence with all Inter-Onset-Intervals being equal. Furthermore different variations of the analysis of Inter-Onset-Interval – the duration between two adjacent elements (IOI (Jung, Kalko, & von Helversen, 2007; Ratcliffe et al., 2011; Schneider & Mercado, 2018a); also called Inter-Pulse-Interval, IPI (Širović et al., 2017) or Inter-Click-Interval, ICI (Ladegaard et al., 2017; Ladegaard et al., 2015b; Sorensen et al., 2018)) - and a so called generate-and-test or GAT approach (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017) were used in rhythm analysis so far. All these methods search for isochronous patterns, therefore, again, we are rather searching for periodicity and isochronous beats underlying a sequence.

This paper aims to help research decide on a method for the analysis of the temporal structure of their biological data. Five methods were used on three different datasets to assess 1) what kind of rhythm an acoustic signal might have (e.g. isochronous vs. heterochronous) and 2) which exact beat frequencies describe a given sequence best. Rhythm analysis can be done on different levels (Fig 1). Depending on the question at hand and the detail of the analysis, different methods can be used. At first, one has to establish whether a given acoustic signal sequence is rhythmic (periodic) at all. The general hypothesis is that a signal is periodic. This can be assessed by a detailed analysis of Inter-Onset-Intervals (IOIs) and by visual assessment of the data. The next step is to decide whether a signal shows an isochronous – that is a metronome-like – rhythm or a heterochronous rhythm. This again can be inferred from IOI analysis and nPVI calculations. If an isochronous rhythm is to be detected and one wants to know the exact beat frequencies of a signal, a Generate-and-test approach (GAT) (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017) or a fast Fourier transformations (FFT (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017; Saar & Mitra, 2008)) can be used; which one to use depends on the data. We developed a goodness-of-fit value for exact beat frequencies

calculated with FFT and by IOI analyis, as these were missing so far. This makes it now possible to not only infer exact beat frequencies but how good a beat frequency actually fits a dataset and how good a 'beat producer' an animal is. To find an underlying pattern within or between individuals a cluster analysis can be run. If a heterochronous beat is to be expected, recurrence plots are a good way to

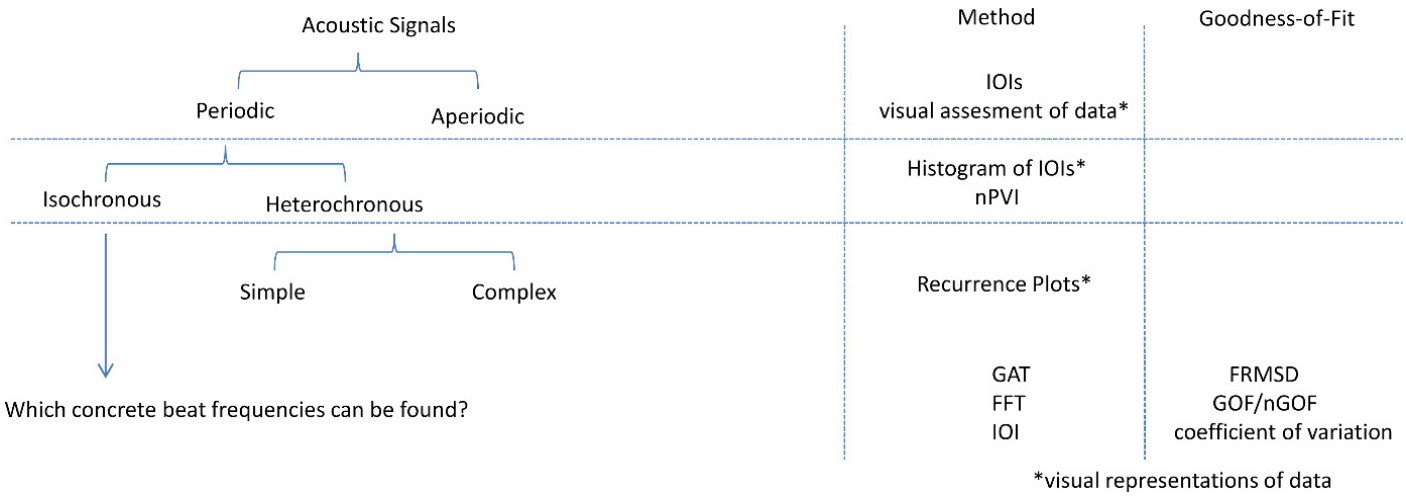


Figure 11: Which Methods to Use Depending on the Level of Analysis.
A first evaluation of whether a signal is periodic or aperiodic relies on IOI and visual assessment of the data. Whether an acoustic signal sequence might be isochronous or heterochronous can be inferred from IOIs and nPVI calculations. To find exact beat frequencies a GAT approach, FFTs or again an assessment of IOIs can be used, and the detection of simple or complex heterochronous patterns is guided visually by recurrence plots. Exact beat frequencies are only interpretable if accompanied by a goodness-of-fit value. The figure was adjusted after (Ravignani et al., 2014).

visualize the data, to find underlying structures and to be able to decide how to proceed in the analysis. Visualizing underlying or sub-structures can also be relevant in the context of nested signals, where a small part of a sequence might have a very different tempo than the rest. In that case it might be worthwhile to rerun parts of the analysis on that specific part. We also introduce recurrence plots on isochronous data in this paper. All of the above-mentioned methods were used on three datasets to compare results and to show the advantages and disadvantages of the different methods as well as their interpretation.

Methods

Labelling of elements and datasets

We chose three different datasets for the analysis with very different properties: 1) monosyllabic (i.e. only one element type in a sequence), short isolation calls of the neotropical bat *Carollia perspicillata* (Figure 12A), 2) multisyllabic, medium long isolation calls of the neotropical bat *Saccopteryx bilineata* (Figure 12B) – both social vocalizations – and 3) monosyllabic, very long echolocation click trains of the sperm whale *Physeter macrocephalus* used for orientation and foraging (Figure 12C).

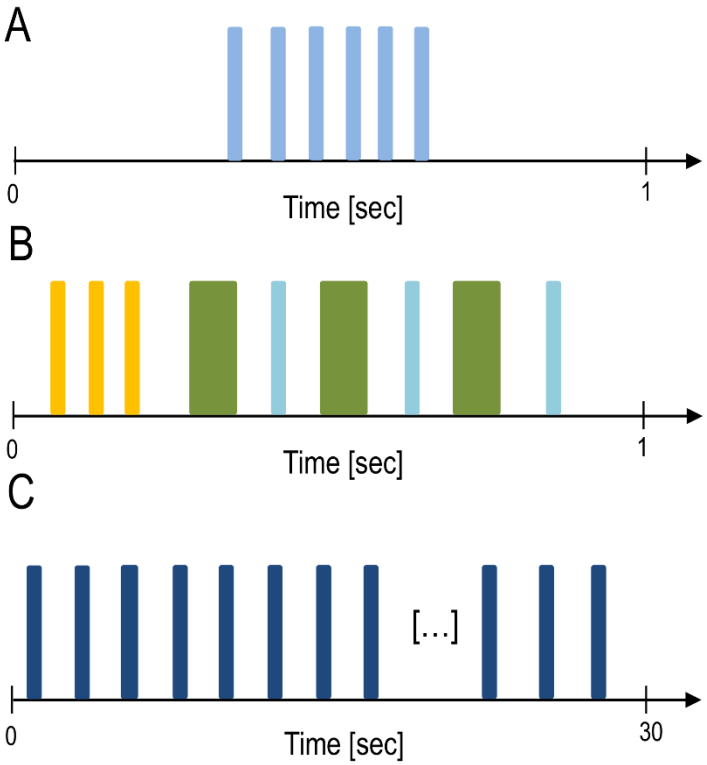


Figure 12: Visual Representation of the Different Sequences.
Different colours indicate different element types. (A) An exemplary sequence of C. perspicillata isolation calls. (B) An exemplary sequence of S. bilineata isolation calls. (C) An exemplary sequence of P. macrocephalus echolocation clicks as used for orientation and foraging. Click trains can be up to 200 elements long.

With this, we cover a broad range of possible acoustic signal sequence structures and can infer the applicability of the methods for a broad range of acoustic signals.

The basis for all analyses were element onsets. An element is a distinct syllable, call or click in a given sequence that is surrounded by silence. It is necessary that elements and their onsets are clearly recognizable. For each acoustic signal sequence, the on- and offset of its elements were determined for subsequent analyses. For multisyllabic isolation calls of *S. bilineata*, we manually determined element

on- and offsets based on oscillograms (see (Knörnschild, Nagy, et al., 2012) for details). For sperm whale echolocation click sequences and isolation call bouts of *C. perspicillata*, we used an automatized procedure in Avisoft SASLab Pro (based on amplitude detection threshold; - 20 dB relative to the element's peak frequency for bats; adjusted manually to not include buzzes for sperm whales) to determine element on- and offsets.

We analysed multisyllabic isolation calls from 5 pups of *S. bilineata* (see (Knörnschild, Nagy, et al., 2012) for details on study site and sound recordings). Each isolation call contained 5 – 26 elements, i.e. syllables (14 ± 3.5, mean ± SD) and was composed of 2 – 4 different element types (mean: 3 element types), but this distinction was not relevant for further analyses. Furthermore, isolation call bouts of 5 *C. perspicillata* pups were analysed (see (Knörnschild et al., 2013) for details on study site and sound recordings). Each bout contained 3-11 elements (mean: 3 elements) and was composed of a single element type. We assessed a total of 47 bouts (Pup 1: 11 bouts, Pup 2: 8 bouts, Pup 3 and 5: 9 bouts, Pup 4: 10 bouts). Furthermore, we analysed 60 sequences of echolocation clicks from a single deep dive of the female sperm whale Sophocles, recorded by the Dominican sperm whale project on 24. April 2014 (for details on study site and recordings see (Bøttcher et al., 2018; Tønnesen et al., 2018)). We extracted trains manually with the software CoolEdit 2000. Single trains were distinguished visually by a clear silent gap of at least 3 seconds (in most cases at least 5 seconds). The elements were afterwards labelled with the software Avisoft SASLab Pro; only the search phase was labelled and feeding buzzes – if at all present – ignored. Feeding buzzes can occur at the very end of a click train when an animal is hunting; they are characterized by a higher repetition rate and less energy (Teloni et al., 2008). Trains contained 13 to 248 elements, i.e. clicks (115 ± 48, mean ± SD).

Rhythm analyses

The different methods used are IOI analyses, including the calculations of coefficients of variation and three methods using the IOIs as input, namely nPVI calculations, Fourier analyses, and a generate-and-test approach. IOIs can be used to visualize the data in histograms or recurrence plots. When one wants to find the exact beat frequencies which best describe an acoustic signal sequence Fourier analysis, IOI analysis, and the GAT approach can be used. To assess how good any of those exact beats describe a given sequence, goodness-of-fit values are crucial. Different values serve as a proxy for the goodness-

of-fit of the best fitting beat in the three different methods and play an important part in the interpretability and comparability of results between species and studies.

IOI. The Inter-Onset-Intervals (IOI) were assessed, and the mean IOI of each sequence converted into the corresponding exact beat frequency by dividing it by 1 [as Hertz is 1/second]. The coefficient of variation was calculated as an indicator of variability. It is estimated as the ratio of the standard deviation to the mean of the sample ((Everitt & Skrondal, 1998), equation 1). The formula for an unbiased estimator ((R. Sokal & Rohlf, 2011), equation 2) was used.

$$\widehat{C_V} = \frac{s}{\bar{x}} \qquad (1)$$

$$\widehat{C_V}^* = \left(1 + \frac{1}{4n}\right)\widehat{C_V} \qquad (2)$$

$\widehat{C_V} = Coefficient\ of\ Variation$ $\widehat{C_V}^* = unbiased\ Coefficient\ of\ Variation$

$s = standard\ deviation$ $\overline{x} = sample\ mean$ $n = sample\ size$

nPVI. Two adjacent IOIs were compared: their difference was calculated and divided by their average; the nPVI gives the average of all these ratios in a sequence multiplied by 100. The obtained values have little explanatory power, beyond being able to assess whether a sequence is isochronous or not (Grabe & Low, 2002; Ravignani & Norton, 2017; Toussaint, 2012, 2013) . We calculated nPVI for all sequences of a dataset separately (named 'sequence' in results) and for all IOIs of a dataset combined (named 'overall' in the results).

$$nPVI = \sum_{k=1}^{m-1} \left| \frac{IOI_k - IOI_{k+1}}{\frac{IOI_k + IOI_{k+1}}{2}} \right| * \frac{100}{m-1} \qquad (3)$$

$nPVI = normalized\ Pairwise\ Variability\ Index$ $k = index\ number$

$m = total\ number\ of\ indices$ $IOI = Inter - Onset - Interval$

Recurrence Plots. In a recurrence plot higher-order patterns within an acoustic signal sequence can be visualized. It plots the sequence of IOIs as their differences, building a raster showing the differences between every IOI with every n-th IOI. The differences are marked by colour code (for code see (Ravignani & Norton, 2017)). Both axes represent the IOI indices in their sequential order.

Fourier analysis. Timestamps of element onsets were used to form a binary point process. Sequences with a time resolution of 5 ms were created, in which only events (i.e., element onsets) were represented by '1', everything else in the sequence was represented by '0'. Each sequence started and ended with an event, represented as a '1'. A fast Fourier transformation was calculated (FFT). After that, frequencies of maximum power were selected as 'best fitting beat' (Burchardt et al., 2019; Saar & Mitra, 2008), which are the exact beat frequencies we subsequently described a sequence with.

A normalized goodness-of-fit value based on the zero-bin component (DC Offset) of the FFT signal was established. In a normal oscillating signal, the zero-bin-component – the amplitude of the signal at 0 Hz – is zero. In a binary sequence, the zero-bin component is not 0 but, instead, same as the mean of the signal in the time domain (adjusted after (Cooley & Tukey, 1965); equations 4 & 5); therefore it is dependent on the total number of elements and the number of samples. It thus functions as an internal reference (Figure 13).

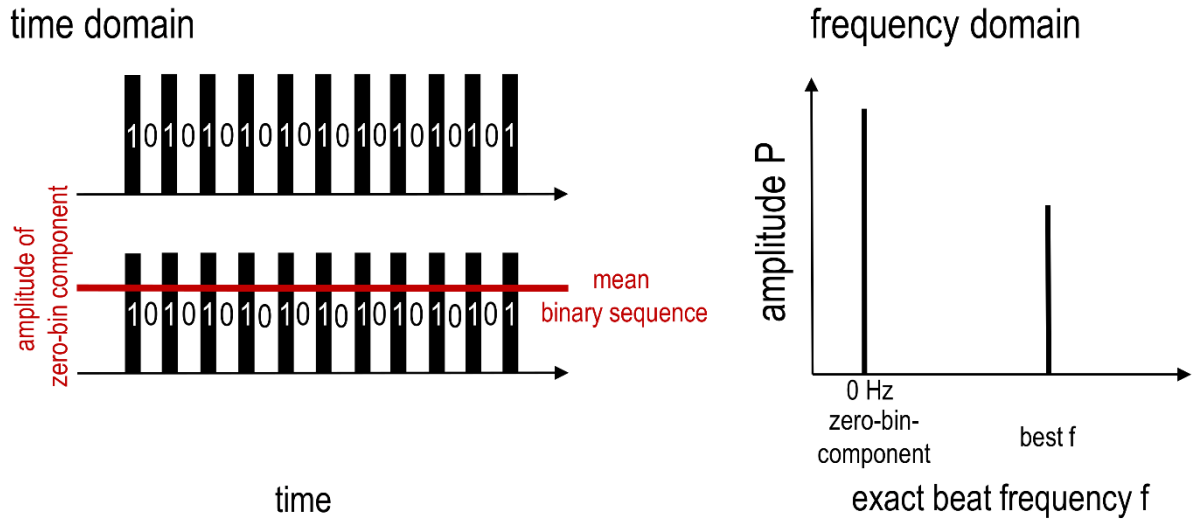


Figure 13: Visual Explanation of the Internal Reference.
The mean of the binary sequence that serves as input for the Fourier analysis determines the amplitude of the zero-bin-component (DC-term). This amplitude will always be the highest in this kind of analysis serving as an internal reference for the second highest peak that determines the best fitting exact beat frequency.

$$X(f) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)e^{-j2\pi\frac{f}{N}n} \qquad (4)$$

$$X(0) = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \qquad (5)$$

$X(f) = Signal\ frequency\ domain \quad N = sample\ size \quad j = j\ function \quad e = Euler\ number$

$n = index\ number \quad x(n) = signal\ in\ time\ domain \quad f = frequency$

The nGOF value is calculated by dividing the amplitude P of the best fitting beat frequency (P_{best}) by the amplitude P of the zero-bin-component (P_0) multiplied with the sampling length (L) (equation 6).

$$nGOF = \frac{|P_{best}|}{L * |P_0|} \qquad (6)$$

$nGOF = normalized\ Goodness\ of\ Fit\ value \quad P_{best} = highest\ Amplitude$

$L = sampling\ length \quad P_0 = Amplitude\ at\ 0\ Hz, zero - bin - component$

GAT. In the generate-and-test approach (developed by (Norton & Scharff, 2016)) the original sequence of element onsets gets tested against computed perfectly isochronous onset sequences of a predefined frequency window (i.e. 5-100 Hz as in beats per second). Sequences were computed in a frequency window from 2 – 100 Hz in 0.01 Hz increments. For each beat frequency, the root-mean-square deviation (RMSD) of all elements in a sequence from their nearest single beat was calculated. The parameter was then normalized for frequency (by dividing it by the frequency), resulting in the frequency normalized root-mean-square deviation – FRMSD.

Artificial data. To further the understanding of the analysis principles we ran all methods on three artificial datasets: 1) perfectly isochronous sequences with IOIs of 0.1, 0.3 or 0.5 seconds; 2) Ten sequences á 100 elements randomly drawn from a uniform distribution between 0 and 1 and 3) three sub datasets, that were drawn from a Gaussian distribution with means of 1, 0.2 or 0.1 seconds with standard deviations of 0.5, 0.1 and 0.05 respectively. Again, each data set consisted of 10 sequences with 100 elements in each sequence. Negative numbers were permitted and the drawing of a negative number re-run until a positive number was drawn. This was done manually (dataset 1) and in Matlab with the 'rand' (dataset 2) and 'normrand' function (dataset 3).

Cluster Analysis

An agglomerative, hierarchical clustering algorithm that used the group average of frequency distances as the basis for finding clusters was applied. Dissimilarities were given by Euclidean distances; the dissimilarity threshold to find clusters was set to 0.05 for all data sets. Cluster analyses were performed for all three methods yielding exact beat frequencies in Matlab.

Software and Code

We used Matlab (Version 2017b & 2016b) and R (Version 3.5.3) for the analyses. CoolEdit 2000 (Syntrillium, Phoenix, USA) was used to extract single echolocation click trains from the dive of a sperm whale. Furthermore, we used Avisoft SASLab Pro Version 5.2.10 (Berlin, Germany) to visualize recordings and to determine element onsets automatically (for isolation call bouts of *C. perspicillata* and click trains of *P. macrocephalus*) and manually (for multisyllabic isolation calls of *S. bilineata*).

The code for the GAT approach was published elsewhere (see (Ravignani & Norton, 2017)) and the code to run the FFT as well as exemplary data is provided here: https://github.com/LSBurchardt/FFT-Method.

Results and Methods discussion

IOIs and nPVIs

We show key data for all datasets in Table 3: the mean of the IOIs in seconds, the standard deviation of IOIs, as well as the coefficient of variation over all IOIs of a dataset (C_V overall) and the average coefficient of variation between sequences (C_V sequences) of a dataset. In contrast to the commonly used parameters variance and standard deviation, the coefficient of variation is neither sample size nor mean dependent. Therefore, it yields comparable results independent of the dataset. To ensure comparability we used the formula for an unbiased estimator ((R. Sokal & Rohlf, 2011), equation 2) since especially for smaller sample sizes the normal coefficient of variation (equation 1) tends to underestimate the variation.

Furthermore, we give information on the range of IOIs in seconds and the number of IOIs comprising the datasets. The range of calculated nPVIs as well as their mean is given together with the information

on the number of sequences underlying the nPVI analysis and subsequent analysis of exact beat frequencies per sequence via GAT, Fourier analysis and IOI calculation.

Table 3: Summary of IOI Results

	Mean IOI [sec]	**SD** (σ)	**Coefficient of Variation (overall)**	**Coefficient of Variation (mean of sequences)**	**Range[sec]**	**n IOIs**	**nPVI**	**n sequences**
C. perspicillata	0.043	0.013	0.31	0.23	0.01 – 0.1	195	2.3 to 110.7 **mean** 35.9	47
S. bilineata	0.078	0.022	0.29	0.19	0.028 – 0.28	646	6.4 to 99.4 **mean** 22.8	50
P. macrocephalus	0.46	0.1	0.22	0.14	0.03 – 3.1	6913	0.4 to 13.6 **mean** 5.2	60

A visual inspection of IOIs is the first step in determining the temporal structure of any given dataset. A unimodal distribution of IOIs is a strong indicator for isochrony because all IOIs spread around the one most prominent duration category. The steeper the distribution, the more consistent an isochronous pattern should be. We find unimodal distributions for all three datasets (Figure 14, first column).

The smaller the C_V (sequences), the less variation we find in IOIs of a dataset, indicating a more consistent structure and possibly isochrony. Smaller nPVIs suggest a similar interpretation. A small nPVI value does not only show a consistent structure but an isochronous structure. When interpreting nPVI values we must consider that even though a very small nPVI indicates isochrony, a middle (20-40) or even high (60-100) nPVI does not necessarily disagree with isochrony and definitely not with rhythmicity. In a computer simulated element sequence with a stress pattern, namely a pattern with an isochronous occurrence of stressed elements, an nPVI value of 94.54 was calculated (see (Ravignani & Norton, 2017)). An indicator of variation between sequences and possibly between individuals is the difference between the C_V (sequences) and C_V (overall). The C_V (sequences) should always be smaller than the C_V (overall), the bigger the difference between the two, the higher the variation between sequences and possibly individuals (see Supporting Information for examples on artificial data).

Looking at the results for our datasets, we can infer isochrony for all three datasets, with *P. macrocephalus* showing the strongest patterning and likely a very strict isochrony and only a few variations between sequences. *S. bilineata* and *C. perspicillata* show values that hint at an underlying isochronous structure with small (*S. bilineata*) and medium (*C. perspicillata*) differences between sequences and individuals.

Exact beat frequencies

After the overall analysis of the pattern, it is interesting to analyse exact beat frequencies, which would describe individual sequences best. Depending on the results of the overall patterns (isochrony or not, high variability vs. low variability), different methods are appropriate to analyse these exact beat frequencies. For example, if results indicate a higher probability of differences between sequences and individuals, an IOI analysis would oversimplify results and we do not consider it fitting. In that case

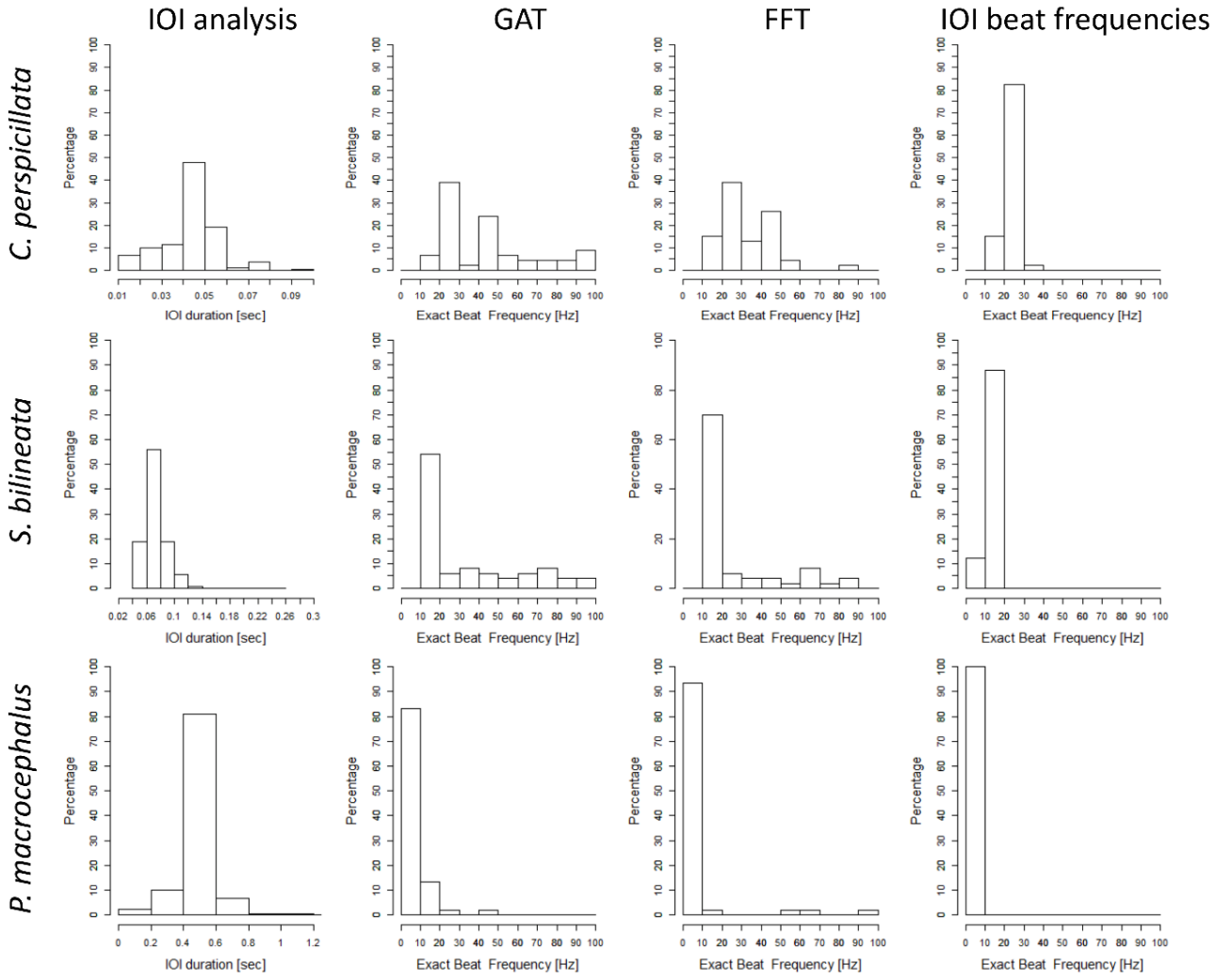


Figure 14: Analysis of the Datasets per Method.
The first column shows the distribution of IOIs for all datasets. The second to fourth column depict exact beat frequency distributions for the three datasets (1. *C. perspicillata*, 2. *S. bilineata*, 3. *P. macrocephalus*) and different methods.

GAT analysis is useful. Nevertheless, if the overall pattern suggests a very strong rhythm, the computationally more intensive analysis of the GAT approach can be spared, because it would most probably not add substantially to the results of the IOI analysis or Fourier analysis. Fourier analysis is a very strong tool to analyse rhythm, but also needs some consideration, for example when deciding which time resolution to choose for the binary sequence. The rather coarse time resolution of 5 ms used in our analysis was chosen for different reasons. A time resolution of 5 ms results in a sampling rate of 200 Hz. Since in an Fourier analysis, signals up to half of the sampling rate can be deconstructed, a sampling rate of 200 Hz will result in frequencies between 0 and 100 Hz being analysed. In other studies, 100 Hz as the upper boundary for the investigation proved suitable for bird song as well as the much faster echolocation pulses of neotropical bats (Burchardt et al., 2019; Norton & Scharff, 2016); therefore, we also used this frequency window for the analysis here. Another very important point to be kept in mind: the chosen time resolution directly influences the frequency resolution of the Fourier signal; the higher the time resolution, the lower the frequency resolution will be and vice versa (equation 7). This problem diminishes with long sampling length but especially in short signals of under and around 1 second, it is a considerable issue. Our chosen time resolution gives suitable frequency resolutions even with short sampling length.

$$\frac{temporal\ resolution}{sample\ length} = frequency\ resolution \qquad (7)$$

Keeping advantages and disadvantages in mind, one should always run more than one analysis method to get a better picture of the data at hand. Our results for the exact beat frequencies are presented in Table 4. For each method for calculating exact beat frequencies (GAT, FFT, IOI) the range of detected beat frequencies is given for the three datasets. The results cluster around certain values. We divided the frequency window we looked at (0-100 Hz) in 10 Hz categories; one category will encompass most of the found sequences (i.e., the category 20 – 30 Hz). This most prominent category is given alongside the percentage of sequences showing beat frequencies in that category. In addition, the results are

visualized in Figure 14, with the different methods in columns and the datasets as rows. The most prominent categories are clearly visible in the histograms for all methods and datasets.

Table 4: Overview of Exact Beat Frequencies Found for Three Datasets with Three Methods.

	GAT			**FFT**			**IOI**		
	Min[Hz]	Max[Hz]	Prominent category [Hz] and %	Min[Hz]	Max[Hz]	Prominent category [Hz] and %	Min[Hz]	Max[Hz]	Prominent category [Hz] and %
C. perspicillata	17.9	100	20-30 39.1%	11.8	83.3	20-30 39.1%	12.4	30.6	20-30 82.6%
S. bilineata	8	100	10-20 54%	11.4	86.6	10-20 70%	8.1	17.1	10-20 88%
P. macrocephalus	2	40.9	0-10 83%	1.7	93.7	0-10 93.3%	1.9	2.4	0-10 100%

Goodness-of-Fit

Finding such strong categories as we can see in the histograms (Figure 14) hints at an underlying isochronous pattern, and we can be sure that the exact beat frequencies we found describe the sequences well. It is very unlikely that we find random exact beat frequencies by chance that show such a pattern of up to 100 % of beats found falling into the same bin category. But what if such an overall pattern is uniformly distributed? How can we be sure that we did not find random beats and how can we compare species and contexts with regards to how well a single beat describes a sequence? For that, we used and developed different goodness-of-fit values which quantify how well a beat describes a sequence.

There are different ways to assess the goodness-of-fit of a beat. By and large, it represents how close the original sequence of elements is described by one certain beat. Since we are searching for the best fitting beat, it describes how well this beat describes the sequence. The goodness-of-fit values for the different methods are correlated to different measures like the number of elements and length of the sequence, and sometimes to a certain extent to beat frequencies; they fall on very different scales and therefore need careful consideration (see Supporting Information for examples on artificial data).

For the GAT approach, the FRMSD (Frequency-normalized Root Mean Square Deviation) depicts the goodness-of-fit. It is positively correlated to the number of elements in a sequence in a non-linear way, but superior to the RMSD which is in addition highly frequency dependent. Using the FRMSD results in finding the slowest beat, coinciding best with element onsets (Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017). It describes the average temporal deviation as a fraction of a full cycle and therefore has no unit (Burchardt et al., 2019; Norton & Scharff, 2016). For the most part, FRMSD values for *C. perspicillata* pups overlap with FRMSD values in *S. bilineata* pups. Nevertheless, the minimum value we find in *S. bilineata* is much higher, while the highest value is lower than in *C. perspicillata* pups. Goodness-of-fit values for the GAT approach show a much broader range for *C. perspicillata.* Element numbers in *S. bilineata* pups are 2- to 9-fold higher, therefore values for *S. bilineata* are considered to show a better fit than the ones for *C. perspicillata*. Due to the FRMSDs positive correlation to element numbers, it is not surprising that we find higher values in the very long sequences of *P. macrocephalus*. Exact values for all three species are shown in Table 5.

Table 5: Comparison of Goodness-of-Fit Values for All Datasets and Methods.

	GAT	FFT		IOI	
Dataset/Method	FRMSD	GOF	nGOF	C_V (overall)	C_V (sequence)
C. perspicillata	0.007-0.214	0.57- 0.98	0.012-0.064	0.31	0.23
S. bilineata	0.059-0.183	0.5 – 0.92	5.5e-4 – 0.014	0.29	0.19
P. macrocephalus	0.07-0.26	0.23- 0.87	1.2e-5 – 0.0032	0.22	0.14

For the Fourier analysis the basis for the goodness-of-fit is the amplitude of the Fourier signal. The amplitude P of the Fourier signal, which is used to determine the best fitting beat, is also indicative of how good the beat actually fits: the higher the amplitude, the better the fit. Nevertheless, the amplitude is strongly correlated to sample length and number of events in the sequence. Therefore, amplitudes could so far only be compared within one dataset and with good knowledge about the correlations. The nGOF on the other hand shows a much smaller correlation with sample length and number of events (Supporting Information) and is therefore more appropriate to use as a goodness-of-fit value. The nGOF was validated by correlating it to the already established goodness-of-fit value of the Generate-and-test approach, the FRMSD value (Supporting Information). The nGOF values range from 8e-6 to 1.3e-3

with a median of 2.5e-5 for *P. macrocephalus*. The measure only set into relation with the internal references – but not normalized for the length of the signal – lie between 0.22 and 0.67 (GOF) which can be thought of as the percentage this one particular beat frequency has on describing the original sequence. This value is easier to interpret, but – again – the signal length has a strong impact, which gets clear when comparing the values of the very long sperm whale click trains with the way shorter values for isolation call bouts of *C. perspicillata,* that show much higher and therefore actually "better" values. All other results, on the other hand, have to lead to the interpretation, that the sperm whale echolocation click trains are a lot more regular and therefore closer to a "perfect" beat than the bat isolation calls. This also shows in the nGOF values for the FFT of *C. perspicillata* pups. They show values that are more than a thousand-fold larger than in the sperm whale data.

The goodness-of-fit values for FFT analysis of *S. bilineata* isolation calls fall in between sperm whales and *C. perspicillata* pups, being 10-fold smaller than the values from *C. perspicillata* and 100-fold larger than *P. macrocephalus*. Again, exact values are shown in Table 5.

In IOI analysis the sample size independent measure of the coefficient of variation (C_V) can be used as an indicator of the goodness-of-fit; the smaller the C_V, the less spread there is in the IOIs, which means they are more similar to each other, thus corresponding to a more regular beat. Since the IOI analysis bears little sequence information it is just indicative of the overall regularity. All measures are shown in Table 5. The differences between C_V (overall) and C_V (sequence) moreover give insight into the likelihood of finding individual differences. While in the C_V (overall) all IOIs of an acoustic signal sequence are regarded, in C_V (sequence) only one sequence is regarded and the average for all analysed sequences calculated. Therefore, we might have individually very isochronous sequences leading to small values for C_V (sequence) but very different sequences, leading to a high value for C_V (overall). Therefore, the bigger the difference between C_V (overall) and C_V (sequence), the higher the likelihood of finding differences between individuals. The difference between the two is the smallest for *P. macrocephalus* and highest for *C. perspicillata*. This leads to the interpretation that it is most likely to find individual differences in exact beat frequency patterns in *C. perspicillata* and we do not expect them in *P. macrocephalus*.

Cluster Analysis

Visual inspection of the detected exact beat frequencies per individual confirms what the overall pattern and C_V calculations already indicated. We find a pattern within individuals, where beat frequencies cluster around certain values. Depending on the method and the dataset, these clusters are differently strong and fall around different values.

The cluster analysis is a good way of depicting "preferences" of the different individuals for certain beats. In *S. bilineata* pups, clusters do not differ much between individuals and show cluster strengths of between 30 and 70 % for the GAT approach; clustering the results of the FFT analysis leads to clusters containing 40% to 60 % of sequences per individual. In IOI analysis, clusters contain between 60% and 100% of sequences. All strongest clusters fall between 10 and 20 Hz (also see (Burchardt et al., 2019)). The picture for *C. perspicillata* pups looks slightly different though. We find the strongest clusters containing a third up to 100 % of sequences of an individual falling into one cluster with IOI analysis. The difference is that not all clusters lie in the same beat category. We find most of the strongest clusters between 20 Hz and 30 Hz for GAT and FFT analysis as well as between 40 Hz and 50 Hz. Other clusters fall in different categories. For IOI analysis, on the other hand, all clusters fall at least partially between 20 Hz and 30 Hz (Figure 15 and Supporting Information).

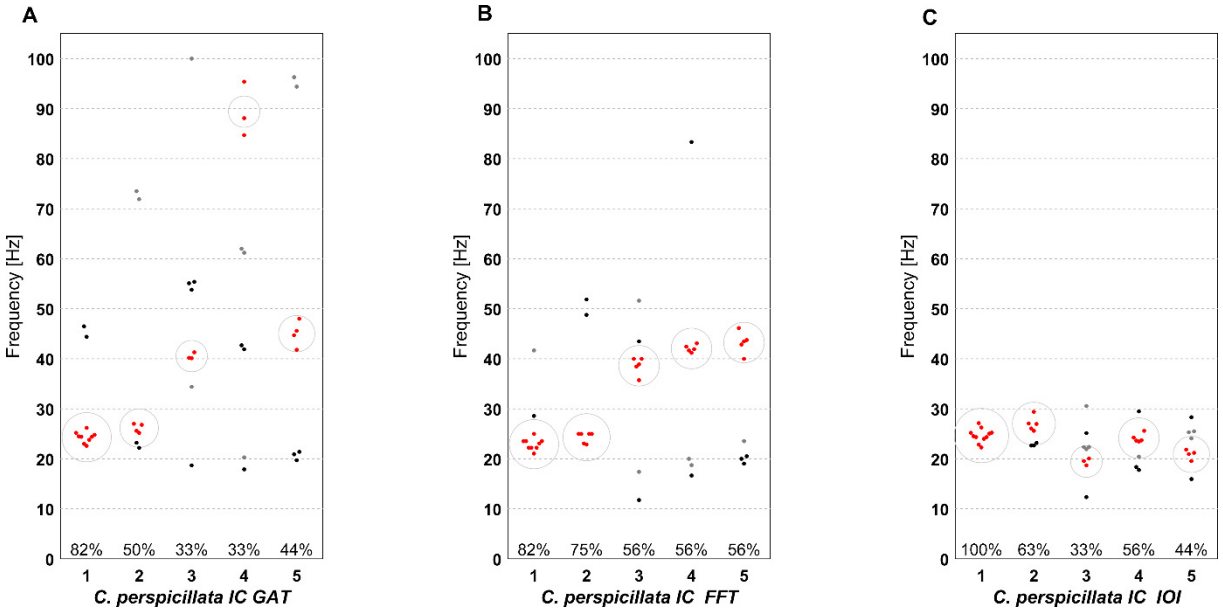


Figure 15: Individual Beat Clusters in C. perspicillata Pups Confirm the Results of Other Methods.
Exact beat frequencies as analysed with the three different methods are shown with clusters in the data. One individual is depicted per column, all exact beat frequencies found are shown as dots. Depicted in red are the sequences falling into the largest cluster of sequences sharing a similar

beat. Percentages at the bottom indicate the percentage of sequences per individual in the largest cluster. (A) Exact beat frequencies and individual clusters as obtained by the GAT approach. (B) Exact beat frequencies and individual clusters as obtained by the FFT method. (C) Exact beat frequencies and individual clusters as obtained by IOI analysis.

Since we analysed echolocation click trains of a single individual for *P. macrocephalus* such a cluster analysis is not useful here. But the very strong patterning and previous research (Whitehead, 2003) let us assume that there are no significant individual differences.

Recurrence Plots

In the following section, we describe two exemplary recurrence plots, one showing a multisyllabic isolation call of *S. bilineata* and the other one showing an echolocation sequence of *P. macrocephalus*. Recurrence plots offer a visual representation of the temporal pattern of a sequence. The more uniform the sequence, the more white and light grey colours can be seen in the plot: white stands for no to very little differences between two adjacent IOIs and the darker a comparison, the bigger the difference. The very strict isochronous pattern of the sperm whale echolocation sequences is depicted in an almost white plot (Figure 16A). In contrast, we can even see the structure of the multisyllabic isolation call of *S. bilineata* pups in the corresponding recurrence plot (Figure 16B), where very similar IOIs are followed by slight pairwise changes of IOIs at the end of the sequence, which corresponds to changes between two element types. These plots could be a very valuable addition in the analysis of more complex temporal structures in acoustic signals because higher order structures – for example, different parts of temporal structure within one acoustic signal sequence – can be visualized and used to determine how to proceed. For very short sequences such as for *C. perspicillata* isolation calls, plotting a recurrence plot most often does not offer additional valuable insights. They are to be interpreted carefully, especially when sequences to be compared via a recurrence plot vary widely in IOI length. The same absolute difference between IOIs might be irrelevant for one but important for another species. The same colour might not stand for the same absolute difference in two plots.

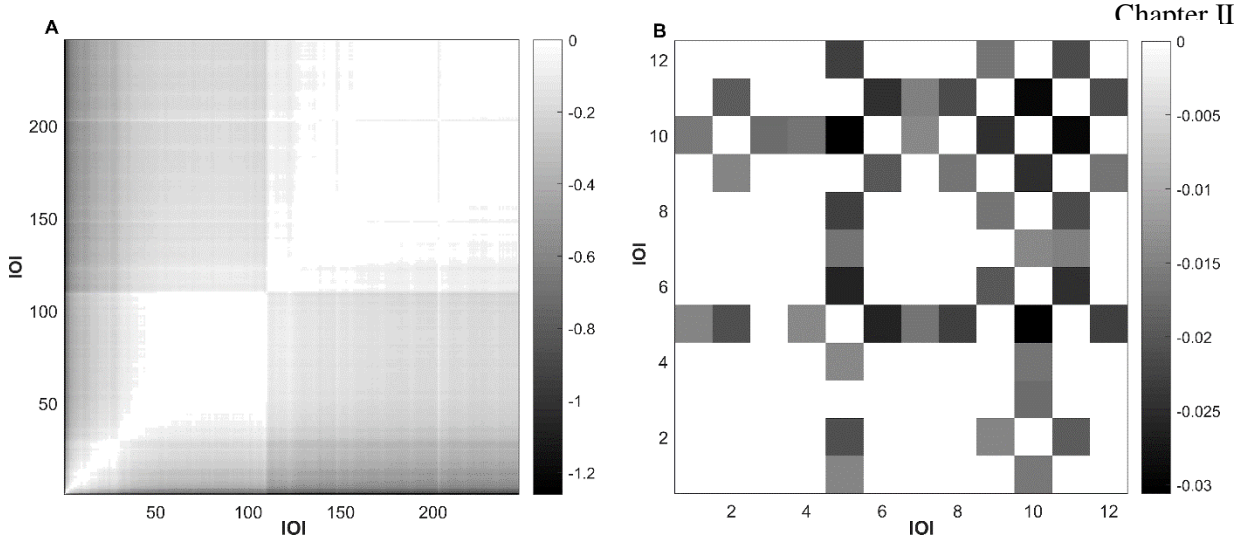


Figure 16: Recurrence Plots of Two Sequences.
No difference is indicated by white; the darker the colour, the bigger the difference. Note that absolute differences are depicted and colours represent different differences in both plots, as shown in the legend. (A) Echolocation click train of *P. macrocephalus:* a very isochronous pattern is visible by only white and light grey colours. (B) The multisyllabic structure of an isolation call of *S. bilineata* is visible in the differences in IOIs: a subsequence of very similar IOIs is followed by an alternating sequence of two more element types.

Decision Tree

Incorporating the different methods into a workflow that includes both the data structure as well as results of early analysis steps leads to a decision tree, describing which methods to use in what case (Figure 17).

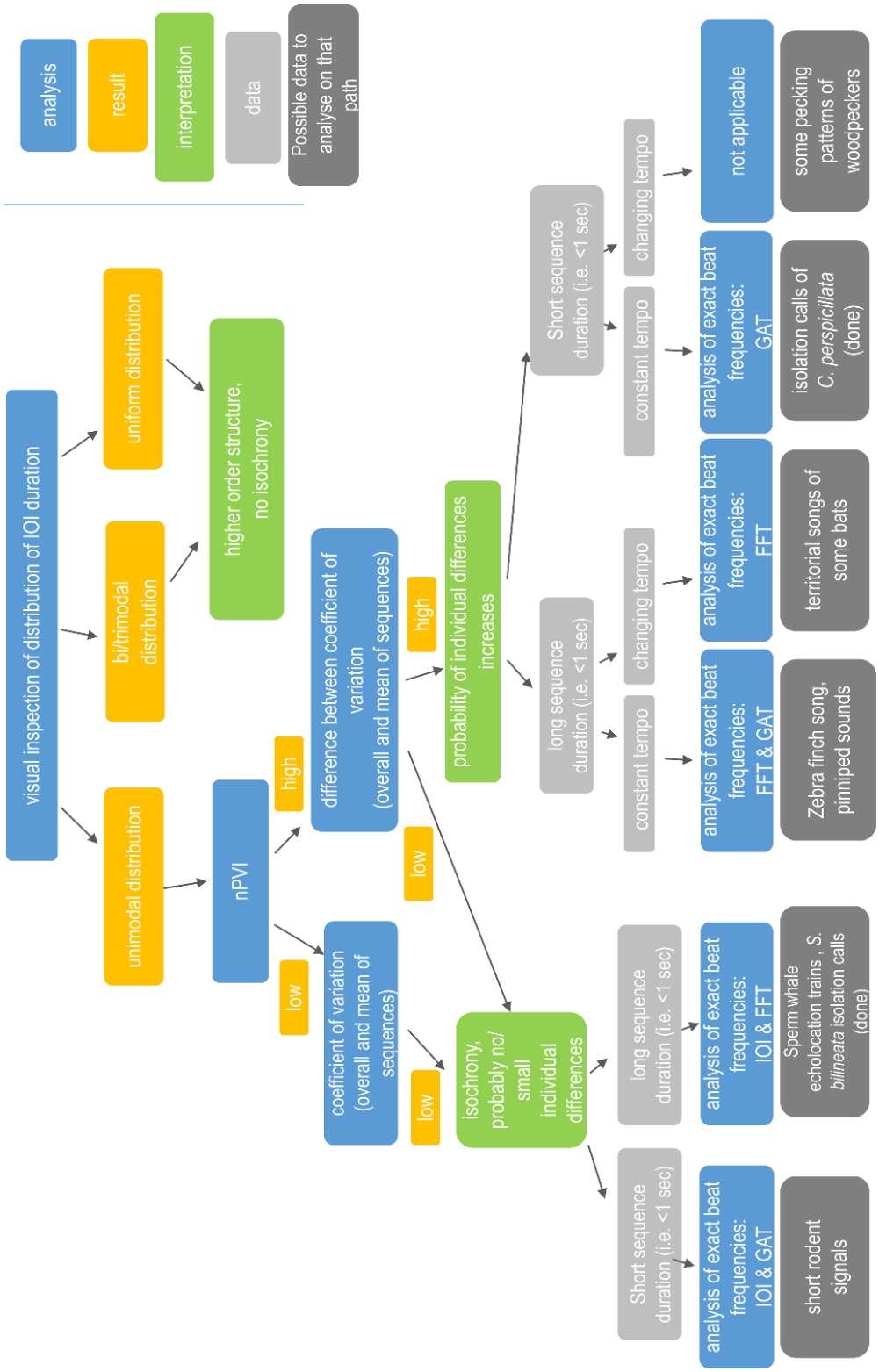


Figure 17: Deciding on a method depending on the dataset and results.
The workflow starts with simple distributional measures such as IOI analysis and nPVI calculations. Questions to be answered in subsequent order are: 1) Is a dataset periodic? 2) If so, can we infer isochrony? 3) Assuming an isochronous pattern, how to analyse exact beat frequencies best, depending on the data at hand? The sequences we analysed fall in three of the decision paths: *S. bilineata* isolation calls would be best to analyse with IOI or FFT; *C. perspicillata* would be best to analyse with the GAT approach, while *P. macrocephalus* echolocation click trains should be analysed with IOI and FFT as well.

Discussion

This study presents a comprehensive overview of the analysis of periodicity and rhythmicality in animal acoustic signals by comparing different methods for three different original datasets and introduces two new goodness-of-fit values for rhythm analysis methods. How to decide on the fitting methods depending on the data is depicted in Figure 17.

Periodicity can be inferred for all three datasets from the results of the IOI analysis and visual assessment of the sequences: multisyllabic isolation calls of *S. bilineata*, isolation call bouts of *C. perspicillata* and echolocation click trains of *P. macrocephalus*. This information might be useful to answer a broad range of questions, but independent of the question at hand are the methods. Those methods enable us to actually infer or exclude periodicity for a given sequence. These methods are the topic of this paper.

The methods (nPVI calculations, C_V, IOI analyses, GAT, Fourier analyses) $_{were}$ adjusted by using three very different kinds of vocalization and sounds for them to be applicable to a broad range of acoustic signals. We used long and short signals in terms of overall duration and element duration, multisyllabic and monosyllabic sequences, and echolocation sequences for navigation as well as social calls. Furthermore, this ensures comparable results and fast and relatively easy implementation of the different analyses, which was the main aim of this study. Nevertheless there might be extreme examples of acoustic signals where the method (i.e. Fourier analysis' time resolution or the frequency window in the GAT approach) could need adjustments; these could include the very slow and long rumbles of elephants (Garstang, 2004; Stoeger et al., 2012) or the extremely fast and short echolocation signals of some bats such as *Kerivoula pellucida*, a small Verspertilionidae bat from Southeast Asia with element lengths of ~1.9 ms and IOIs of around 5 ms (Schmieder et al., 2010) or the even shorter but a little slower calls of *Micronycteris microtis* with an element length of 0.2 ms and IOIs of 14 to 30 ms (Geipel et al., 2013). For the very short elements of some bat species, the sampling rates for creating the binary sequence, serving as input for Fourier analysis, would need to be much higher for two reasons: first, with a time resolution of 5 ms and element lengths of 2 ms or even 0.2 ms, the accuracy of labelling becomes too coarse. Second, the range of frequencies a sequence is described with in a Fourier analysis is dependent on the sampling rate; with the used sampling rate, frequencies of up to 100 Hz can be decomposed but

this is not enough for faster signals (decomposition into frequencies up to half the sampling rate). Changing the sampling rate would on the other hand have implications on the frequency resolution. Duration of samples would need to be at least ~1 second for sampling frequencies up to 1000 Hz; if sequences are shorter, Fourier analysis is not suitable (equation 7). On the other end of extremes, very slow signal sequences should not generate these kinds of problems. The frequency range to test for exact beat frequencies with the GAT approach would need careful consideration in this case though, because 2 Hz, which was the lower boundary in this case, might not be slow enough.

Looking at the different possible analysis paths we describe in the flowchart (Figure 17), we used data fitting into three different paths, leading to two different end categories. All three datasets show a unimodal distribution when looking at the IOI distribution. Results of nPVI and C_V calculations differ. While echolocation click sequences of the sperm whale show low nPVI values and a small difference between C_V (sequence) and C_V (overall), both bat vocalizations do not fall into that category because nPVI values are higher. Nevertheless, since the difference between C_V (overall) and C_V (sequence) are also small in *S. bilineata*, we proceed in the flowchart to the interpretation that, comparable to the sperm whale trains, isolation calls of *S. bilineata* are isochronous with probably no or only small individual differences. For *C. perspicillata* isolation calls, however, we conclude that even though isochronous, the probability for individual differences is increased, and therefore we proceed on a different path in the analysis.

Sequences of *S. bilineata* and *P. macrocephalus* are adequate in length (i.e mostly more than 1 second), therefore the frequency resolution in Fourier analysis is no problem and IOI analysis and Fourier analysis are most suitable for exact beat analysis. *C. perspicillata* sequences are shorter than 1 second and show a constant tempo, which would make the GAT approach the most suitable one. To give possible acoustic signal types for other paths, depending on the data might be from left to right in Figure 17: short call sequences of rodents, for example ultrasonic pulses of *Typhlomus chapensis* (Volodin et al., 2018); for sequences with a higher probability for individual differences that are above 1 second in duration and show a constant tempo one could think of male zebra finch song (Norton & Scharff, 2016) or the vocalization sequences of pinnipeds such as the Northern elephant seal (Mathevon et al., 2017). For a sequence with a changing tempo, one might think of a territorial song of some bat species that

escalate and increase the tempo in the end (Behr et al., 2006; Knörnschild et al., 2017; Voigt et al., 2008). Sequences, where none of the methods would be applicable, could, for example, be short, accelerating pecking patterns of woodpeckers (Miles et al., 2018).

The analysis of echolocation click trains of *P. macrocephalus* shows some interesting discrepancies between methods. Beat frequencies as known from the literature – often termed click rates or repetition rates in the respective literature – lie around 0.7 – 4 Hz (Douglas et al., 2005; Madsen et al., 2002). Using the IOI analysis, we get results fitting perfectly into that frame, which makes sense, as the same methodology is used. The other analyses also show way faster beat frequencies, even though not very prominently. The important message is that Fourier analysis and the GAT approach reproduce the overall pattern that most echolocation trains show beats as previously described in the literature. Nevertheless, it also shows the possibility of the oversimplification of IOI analysis; this needs more analyses, but it might be possible that especially in more variable contexts than whale echolocation, IOI analysis is missing a lot of information, e.g., small differences that might be pronounced between individuals for discrimination purposes. It was already suggested that echolocation click beats of sperm whales may include this information (André & Kamminga, 2000).

There are a few general take home messages regarding methods to analyse the rhythm. Starting with the data, since all analyses rely on IOIs, elements need to be clearly separable, and recordings need to have a good signal-to-noise ratio. Furthermore, the duration of a single sequence (i.e., duration between the first and the last element) should not be too short and a sequence should contain at the very least 3 elements for all methods to be applicable. In general, as many methods as possible should be applied to get a full picture of the data. Different methods have different flaws; by using various methods and comparing the results, artefacts or inconsistencies are easier to detect. Methods to calculate exact beat frequencies do have very different major flaws: Fourier analysis is not well applicable for very short sequences, because of the trade-of between time resolution in the original signal and frequency resolution in the Fourier signal (equation 7). The GAT approach has issues with sequences changing in tempo since the optimization task is carried out for all elements within a sequence, such that one outlier can influence the results strongly. IOI analysis tends to oversimplify structures since it depends only on the mean of IOIs in a sequence, which is not depicting the variation in a sequence at all (Figure 15).

To enable reproducible rhythm analysis, one needs to provide at least the original IOI sequences of the data or even the raw acoustic signals with labels. Information on the generation of the binary sequence for the Fourier analysis is essential; this mainly refers to the time resolution used. If cluster analyses are run to detect individual patterns, reporting the used distance measures, as well as clustering algorithm and distance threshold, are necessary to make results comparable between studies.

Considering all these things, rhythm analysis can be used to tackle many questions. Not only can we further investigate couplings of biological processes such as motor rhythms (Moss et al., 2006; Suthers & Fattu, 1973), but it can be used to find possible guiding neural processes (Norton & Scharff, 2016; Yartsev & Ulanovsky, 2013) and can give valuable information for studies on the perception of temporal structures (García-Rosales, Martin, et al., 2018). Especially in echolocating animals such as whales and bats, rhythm analysis yields a good background for studies on rhythm perception. Furthermore, rhythm analysis might prove to be a valuable tool for the analysis of vocal production learning, as was already suggested for example for the vocal learning in zebra finches, where very stereotyped elements are learned, with a difference only in the temporal structure (Hyland Bruno & Tchernichovski, 2017). In other species, one aspect of vocal production learning is for the animal to learn the temporal structure of an acoustic signal. Without knowing the beats produced by animal tutors and tutees, this is difficult to achieve (Wirthlin et al., 2019).

Looking at a broad range of animal acoustic signals and uncovering broader patterns between animal taxa can, in the end, inform us about the origins and importance of periodicity and rhythmicity.

Acknowledgments

We thank the Dominica Sperm Whale Project and The Marine Bioacoustics Lab, Aarhus University, for sharing their sperm whale audio recording and helpful discussions, as well as the whole Knörnschild Lab for fruitful discussions. This work was supported by an Elsa-Neumann Stipend to Lara S. Burchardt and a Heisenberg Fellowship (DFG KN935 3-1) from the German Research Foundation Mirjam Knörnschild.

References

André, M., & Kamminga, C. (2000). Rhythmic dimension in the echolocation click trains of sperm whales: a possible function of identification and communication. *Journal of the Marine Biological Association of the United Kingdom, 80*(1), 163-169. https://doi.org/10.1017/S002531549900168X

Behr, O., Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., & Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behavioural Ecology*(17), 810-817. https://doi.org/10.1093/beheco/arl013

Bøttcher, A., Gero, S., Beedholm, K., Whitehead, H., & Madsen, P. T. (2018). Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification. *The Journal of the Acoustical Society of America, 144*(1), 365-374. https://doi.org/10.1121/1.5047657

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science, 6*. http://doi.org/10.1098/rsos.181076

Celma-Miralles, A., & Toro, J. M. (2018). Beat perception in a non-vocal learner: rats can identify isochronous beats. The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)

Cooley, J. W., & Tukey, J. W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation, 19*(90), 297-301. https://doi.org/10.2307/2003354

David, J. A. d. O., Zefa, E., & Fontanetti, C. S. (2003). Cryptic species of *Gryllus* in the light of bioacoustic (Orthoptera: Gryllidae). *Neotropical Entomology, 32*(1), 75-80. https://doi.org/10.1590/S1519-566X2003000100010

Douglas, L. A., Dawson, S. M., & Jaquet, N. (2005). Click rates and silences of sperm whales at Kaikoura, New Zealand. *The Journal of the Acoustical Society of America, 118*(1), 523-529. https://doi.org/10.1121/1.1937283

Everitt, B., & Skrondal, A. (1998). *The Cambridge Dictionary of Statistics* (4 ed.). Cambridge University Press.

García-Rosales, F., Martin, L. M., Beetz, M. J., Cabral-Calderin, Y., Kössl, M., & Hechavarria, J. C. (2018). Low-Frequency Spike-Field Coherence Is a Fingerprint of Periodicity Coding in the Auditory Cortex. *iScience, 9*, 47-62. https://doi.org/10.1016/j.isci.2018.10.009

Garstang, M. (2004). Long-distance, low-frequency elephant communication. *J Comp Physiol A 190*, 791-805. https://doi.org/10.1007/s00359-004-0553-0

Geipel, I., Jung, K., & Kalko, E. K. V. (2013). Perception of silent and motionless prey on vegetation by echolocation in the gleaning bat Micronycteris microtis. *Proc Biol Sci, 280*(1754), 20122830-20122830. https://doi.org/10.1098/rspb.2012.2830

Grabe, E., & Low, E. (2002). Durational variability in speech and the rhythm class hypothesis. In *Laboratory Phonology VII* (Vol. 7, pp. 515-546). Mouton de Gruyter. https://doi.org/10.1515/9783110197105.515

Heinsohn, R., Zdenek, C. N., Cunningham, R. B., Endler, J. A., & Langmore, N. E. (2017, Jun). Tool-assisted rhythmic drumming in palm cockatoos shares key elements of human instrumental music. *Sci Adv, 3*(6), e1602399. https://doi.org/10.1126/sciadv.1602399

Hulse, S. H., Humpal, J., & Cynx, J. (1984). Discrimination and Generalization of Rhythmic and Arrhythmic Sound Patterns by European Starlings (*Sturnus vulgaris*). *Music Perception: An Interdisciplinary Journal, 1*(4), 442. https://doi.org/10.2307/40285272

Hyland Bruno, J., & Tchernichovski, O. (2017). Regularities in zebra finch song beyond the repeated motif. *Behav Processes*. https://doi.org/10.1016/j.beproc.2017.11.001

Jung, K., Kalko, E. K. V., & von Helversen, O. (2007). Echolocation calls in Central American emballonurid bats: signal design and call frequency alternation. *Journal of Zoology*(272), 125-137.

Knörnschild, M., Bluml, S., Steidl, P., Eckenweber, M., & Nagy, M. (2017). Bat songs as acoustic beacons - male territorial songs attract dispersing females. *Scientific Reports, 7*(1), 13918. https://doi.org/10.1038/s41598-017-14434-5

Knörnschild, M., Feifel, M., & Kalko, E. K. V. (2013). Mother-offspring recognition in the bat *Carollia perspicillata*. *Animal Behaviour, 86*, 941-948. https://doi.org/10.1016/j.anbehav.2013.08.011

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Animal Behaviour, 84*, 761-769. http://dx.doi.org/10.1016/j.anbehav.2012.06.029

Ladegaard, M., Jensen, F. H., Beedholm, K., da Silva, V. M. F., & Madsen, P. T. (2017). Amazon river dolphins (*Inia geoffrensis*) modify biosonar output level and directivity during prey interception in the wild. *The Journal of Experimental Biology, 220*, 2654-2665. https://doi.org/10.1242/jeb.159913

Ladegaard, M., Jensen, F. H., de Freitas, M., Ferreira da Silva, V. M., & Madsen, P. T. (2015, Oct). Amazon river dolphins (*Inia geoffrensis*) use a high-frequency short-range biosonar. *The Journal of Experimental Biology, 218*, 3091-3101. https://doi.org/10.1242/jeb.120501

Lin, H., & Wang, Q. (2007). Mandarin rhythm: An acoustic study. *Journal of Chinese Language and Computing, 17*(3), 127-140.

Madsen, P. T., Wahlberg, M., & Møhl, B. (2002). Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology, 53*(1), 31-41. https://doi.org/10.1007/s00265-002-0548-1

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals Voices. *Current Biology, 27*(15), 2352-2356.e2352. https://doi.org/10.1016/j.cub.2017.06.035

Merguerditchian, A., Vuillemin, A., & Pruetz, J. D. (2018). *Identifying the ape beat in the wild: rhythmic indvidual signatures from sounds of manual fruit cracking in Fongoli Chimpanzees* The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12),

Miles, M. C., Schuppe, E. R., Ligon, R. M. t., & Fuxjager, M. J. (2018). Macroevolutionary patterning of woodpecker drums reveals how sexual selection elaborates signals under constraint. *Proc R Soc B 285*(1873), 20172628. https://doi.org/10.1098/rspb.2017.2628

Moss, C. F., Bohn, K., Gilkenson, H., & Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *PLOS Biology, 4*(4, e79). https://doi.org/10.1371/journal.pbio.0040079

Munafó, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *nature human behaviour, 1*. https://doi.org/10.1038/s41562-016-0021 |

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

R. Sokal, R., & Rohlf, F. (2011). *Biometry* (4 ed.). W. H. Freeman

Ratcliffe, J. M., Jakobsen, L., Kalko, E. K. V., & Surlykke, A. (2011). Frequency alternation and an offbeat rhythm indicate foraging behavior in the echolocating bat, *Saccopteryx bilineata*. *Journal of Comparative Physiology A*. https://doi.org/10.1007/s00359-011-0630-0

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Front Psychol, 5*, 1118. https://doi.org/10.3389/fpsyg.2014.01118

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speec, movement and animal vocalizations. *Journal of Language Evolution, 2*(1), 4-19. https://doi.org/10.1093/jole/lzx002

Ravignani, A., Thompson, B., & Filippi, P. (2018). The Evolution of Musicality: What Can Be Learned from Language Evolution Research? *Frontiers in Neuroscience, 12*(20). https://doi.org/10.3389/fnins.2018.00020

Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One, 3*(1), e1461. https://doi.org/10.1371/journal.pone.0001461

Schmieder, D. A., Kingston, T., Hashim, R., & Siemers, B. M. (2010). Breaking the trade-off: rainforest bats maximize bandwidth and repetition rate of echolocation calls as they approach prey. *Biology Letters, 6*(5), 604-609. https://doi.org/10.1098/rsbl.2010.0114

Schneider, J. N., & Mercado, E. (2018). Characterizing the rhythm and tempo of sound production by singing whales. *Bioacoustics, 28*(3), 239-256. https://doi.org/10.1080/09524622.2018.1428827

Širović, A., Oleson, E. M., Buccowich, J., Rice, A., & Bayless, A. R. (2017). Fin whale song variability in southern California and the Gulf of California. *Scientific Reports, 7*. https://doi.org/10.1038/s41598-017-09979-4

Sorensen, P. M., Wisniewska, D. M., Jensen, F. H., Johnson, M., Teilmann, J., & Madsen, P. T. (2018, Jun 26). Click communication in wild harbour porpoises (*Phocoena phocoena*). *Scientific Reports, 8*. https://doi.org/10.1038/s41598-018-28022-8

Stoeger, A. S., Heilmann, G., Zeppelzauer, M., Ganswindt, A., Hensman, S., & Charlton, B. D. (2012). Visualizing sound emission of elephant vocalizations: evidence for two rumble production types. *PLoS One, 7*(11), e48907-e48907. https://doi.org/10.1371/journal.pone.0048907

Suthers, R. A., & Fattu, J. M. (1973). Mechanisms of Sound Production by Echolocating Bats. *American Zoologist*(13), 1215-1226.

Teloni, V., Mark, J. P., Patrick, M. J. O., & Peter, M. T. (2008). Shallow food for deep divers: Dynamic foraging behavior of male sperm whales in a high latitude habitat. *Journal of Experimental Marine Biology and Ecology, 354*(1), 119-131. https://doi.org/10.1016/j.jembe.2007.10.010

Tønnesen, P., Gero, S., Ladegaard, M., Johnson, M., & Madsen, P. T. (2018). First-year sperm whale calves echolocate and perform long, deep dives. *Behavioral Ecology and Sociobiology, 72*(10), 165. https://doi.org/10.1007/s00265-018-2570-y

Toussaint, G. T. (2012, 23.07-28.07). The Pairwise Variability Index as a Tool in Musical Rhythm Analysis. In E. Cambouroploulos, C. Tsougras, P. Mavromatis, & K. Pastiadis, Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Science of Music, Thessaloniki, Greece.

Toussaint, G. T. (2013). *The geometry of musical rhythm: what makes a "good" rhythm good?* CRC Press.

Voigt, C. C., Behr, O., Caspers, B., Helversen, O., & Knörnschild, M. (2008). Songs, Scents, and senses: Sexual Selection in the greater sac-winged bat, Saccopteryx Bilineata. *Journal of Mammalogy, 89*(6), 1401-1410.

Volodin, I., A. Panyutina, A., Abramov, A., Ilchenko, O., & Volodina, E. (2018). Ultrasonic bouts of a blind climbing rodent (*Typhlomys chapensis*): acoustic analysis. https://doi.org/10.1080/09524622.2018.1509374

Wallin, N. L. (1991). *Biomusicology - Neurophysiological, Neuropsychological, and Evolutionary Perspectives on the Origins and Purpose of Music*. Pendragon Press.

Whitehead, H. (2003). *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press.

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain Lang, 144*, 26-34. https://doi.org/10.1016/j.bandl.2015.03.008

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., & Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron, 104*(1), 87-99. https://doi.org/10.1016/j.neuron.2019.09.036

Yartsev, M. M., & Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. *Science, 340*(6130), 367-372. https://doi.org/10.1126/science.1235338

Supplementary Information

This document includes supporting information on artificial data and the validation and explanation of the goodness-of-fit value nGOF.

Validation normalized Goodness-of-Fit value, Fourier analysis

To validate the use of the normalized goodness-of-fit value (nGOF) in Fourier analysis, we correlate it with the sample length, to show that the nGOF shows the smallest correlation coefficient as compared to the goodness-of-fit value that was not normalized (GOF) or the amplitude in the frequency spectrum. These correlations only show when including all analysed data, ranging from very short to very long sequences. To ensure comparability between exactly these very different sequences, it is important to use the value least correlated to the sample length.

Furthermore, we validated the nGOF by comparing and correlating it to the already published goodness-of-fit value for the Generate-and-Test approach, i.e., the frequency-normalized root-mean-square-deviation (FRMSD). The two values strongly correlate, which ensures us it is appropriate to use it as a goodness-of-fit value. All correlation coefficients were calculated in R (version 3.5.3) with the function 'cor', which by default is calculating a Pearson correlation. We show the results in Table 6.

Table 6: Correlation of Goodness-of-Fit and Related Values

	nGOF	GOF	P	sample length	FRMSD	# elements
nGOF						
GOF	0.73					
P	0.94	0.83				
sample length	-0.48	-0.78	-0.66			
FRMSD	-0.77	-0.84	-0.86	0.71		
# elements	-0.51	-0.79	-0.69	0.99	0.73	

Artificial data

To make it easier to set results into perspective, we analysed three artificial data sets, using the same workflow as for the original biological data.

Results were calculated for perfect isochronous sequences with Inter-Onset-Intervals (IOIs) of 0.1, 0.3 and 0.5 seconds. In a second dataset numbers were drawn randomly from a uniform distribution ranging

from 0 to 1. A third dataset consisted of numbers randomly drawn from a Gaussian distribution with three different parameter combinations. 100 elements per sequence, 10 sequences each were drawn for the following parameter combinations: 1) mean: 0.1 seconds, standard deviation: 0.05 seconds; 2) mean: 0.2 seconds, standard deviation: 0.1 seconds and 3) mean: 1 second, standard deviation: 0.5 seconds. That way, we expected results to be around 10 Hz, 5 Hz and 1 Hz respectively for Inter-Onset-Interval analysis since this simply depends on the mean. The coefficient of variation (C_v) should be very similar between the three groups, since the relation between mean and standard deviation was chosen to be the same in all three cases. That way, the difference between the mean coefficient of variation for all sequences and the coefficient of variation overall over different subsets of the data illustrates the relation between the difference between the two and the possibility of different underlying beats very nicely. Different beats, analysed separately, can have the same C_v mean for all sequences, since the variation is similar in all the sequences (as the modelled relation between mean and standard deviation is similar). As we calculate the results over the different subsets combined, the variation increases, because Inter-Onset-Intervals from the different modelled distributions form the basis for the calculation and therefore the variation increases. If we transfer that to our original data sets this means that the higher the difference between the two values, the higher the probability that different sub distributions underlie the different sequences. That indicates differences between sequences and/or individuals. All results are shown in Table 7, where the first given value represents the expected value, whereas the second value is the calculated value. They fit very nicely in all cases.

Table 7: IOI Analysis of Artificial Data, Expected Values and Calculated Results

Dataset	Mean[sec]	Std [sec]	C_v mean of sequences	C_v overall	Beat [Hz]
1 Hz	1/1.05	0.5/0.48	0.45/0.45	0.45/0.46	1 / 0.96
5 Hz	0.2/0.21	0.1/0.1	0.45/0.46	0.45/0.46	5 / 4.78
10 Hz	0.1/0.11	0.05/0.05	0.45/0.46	0.45/0.46	10 / 9.57

1 Hz + 5 Hz	0.6 /0.63	0.5 / 0.54	0.45/0.46	-- / 0.86	1.7 / 1.6
1 Hz + 10 Hz	0.55/0.57	0.5 / 0.58	0.45/46	-- / 1.01	1.8 / 1.74
5 Hz + 10 Hz	0.15/0.16	0.1 / 0.1	0.45/46	-- / 0.58	6.7 / 6.4
1 Hz + 5 Hz + 10 Hz	0.44 /0.45	0.5 / 0.5	0.45/46	-- / 1.18	2.3/ 2.21

Rhythm analysis results of artificial data

Here we show the resulting exact beats for the three different methods Inter-Onset-Interval analysis, Fourier analysis and the Generate-and-test approach. We show the detailed numbers for the Fourier analysis, to give insight into what the newly established goodness-of-fit values looks like for different datasets (Table 8-10). Furthermore, we show results for all methods in a histogram depiction for the datasets drawn from normally distributed numbers (Figure 18). Furthermore, we show the results of the cluster analysis on dataset 3, drawn from three Gaussian distributions (Figure 19). We can see very clear clusters for the IOI analysis, as these are based on the modelled means, but for the other two methods we do not see clear clusters at all. This is especially interesting as in our original datasets we see in parts similar coefficients of variation, the resulting clusters are still much stronger (20-30% in one cluster in artificial data versus 33-82% in biological data for the GAT approach for example). Also, the expected beat frequencies, based on the modelled distributions would be 1 Hz for sub-dataset 1, 5 Hz for sub-dataset 2 and 10 Hz for sub-dataset 3. This is not at all supported by the results, which again shows that Inter-Onset-Interval analysis oversimplifies results considerably.

Table 8: Results of Fourier Analysis for Perfectly Isochronous Sequences of Different IOIs

Modelled IOI	Amplitude P	Resulting beat [Hz]	Sample length	GOF	nGOF
0.1 sec	0.0503	9.995	1981	0.9958	5.0268e-04
0.3 sec	0.0168	3.333	5941	0.9995	1.6824e-04
0.5 sec	0.0101	1.9998	9901	0.9998	1.0098e-04

Table 9: Results of Fourier Analysis for Sequences Drawn from a Uniform Distribution Between 0 & 1

Sequence #	Amplitude P	Resulting beat [Hz]	Sample length	GOF	nGOF
1	0.0031	54.49	10398	0.3180	3.0578e-05
2	0.0031	91.67	10045	0.3105	3.0912e-05
3	0.0032	82.06	9318	0.2937	3.1516e-05
4	0.0032	24.62	9961	0.3200	3.2129e-05
5	0.0029	46.45	9555	0.2753	3.8814e-05
6	0.0029	88.89	9774	0.2856	3.9225e-05
7	0.0031	32.91	9906	0.3119	3.1488e-05
8	0.0030	47.43	10196	0.3098	3.0384e-05
9	0.0031	81.36	8987	0.2793	3.1074e-05
10	0.0038	15.96	8483	0.3244	3.8238e-05

Table 10: Results of Fourier Analysis for Sequences Drawn from Three Different Gaussian Distributions

Sequence #	Amplitude P	Resulting beat [Hz]	Sample length	GOF	nGOF
mean_1_seq01	0.0014718	3.40623546	21490	0.31948573	1.4867E-05
mean_1_seq02	0.00136936	26.9280919	20707	0.28355351	1.3694E-05
mean_1_seq03	0.00129352	37.3855158	22383	0.28952889	1.2935E-05
mean_1_seq04	0.00144668	60.5123294	20885	0.30213879	1.4467E-05
mean_1_seq05	0.00158602	20.1604867	19067	0.30240621	1.586E-05
mean_1_seq06	0.00150281	19.8842456	20388	0.30639232	1.5028E-05
mean_1_seq07	0.00147356	4.13854704	20152	0.29695108	1.4736E-05
mean_1_seq08	0.00145706	21.9145803	20370	0.29680217	1.4571E-05
mean_1_seq09	0.00153698	74.6243407	20098	0.30890217	1.537E-05
Mean_1_seq10	0.00142628	71.777362	21380	0.3049378	1.4263E-05
mean_0.2_seq01	0.00649635	17.0134073	4326	0.28103226	6.4964E-05
mean_0.2_seq02	0.00602181	97.4408498	4142	0.24942354	6.0218E-05
mean_0.2_seq03	0.00656363	34.7476552	4478	0.29391954	6.5636E-05
mean_0.2_seq04	0.00678646	7.03685974	4178	0.28353816	6.7865E-05
mean_0.2_seq05	0.00783899	47.5616151	3814	0.29897908	7.839E-05
mean_0.2_seq06	0.00711769	14.9092692	4078	0.29025924	7.1177E-05
mean_0.2_seq07	0.00716861	42.2227735	4031	0.28896665	7.1686E-05
mean_0.2_seq08	0.00641353	74.3067485	4075	0.26135143	6.4135E-05
mean_0.2_seq09	0.00704148	6.41631435	4021	0.28313806	7.0415E-05
mean_0.2_seq10	0.00668986	35.9214219	4276	0.28605838	6.6899E-05
mean_0.1_seq01	0.01192391	19.4085028	2164	0.25803339	0.00011924
mean_0.1_seq02	0.01399666	72.4903475	2072	0.29001089	0.00013997
mean_0.1_seq03	0.01315409	69.4953104	2239	0.29452001	0.00013154
mean_0.1_seq04	0.01288567	14.0737195	2089	0.26918168	0.00012886
mean_0.1_seq05	0.01582618	35.8678553	1907	0.30485385	0.00015986
mean_0.1_seq06	0.01279744	29.8039216	2040	0.26106775	0.00012797
mean_0.1_seq07	0.01363039	41.3690476	2016	0.27478874	0.0001363
mean_0.1_seq08	0.01276541	21.5897939	2038	0.260159	0.00012765
mean_0.1_seq09	0.01502785	55.2238806	2010	0.30205979	0.00015028
mean_0.1_seq10	0.01458886	91.8186068	2139	0.31205579	0.00014589

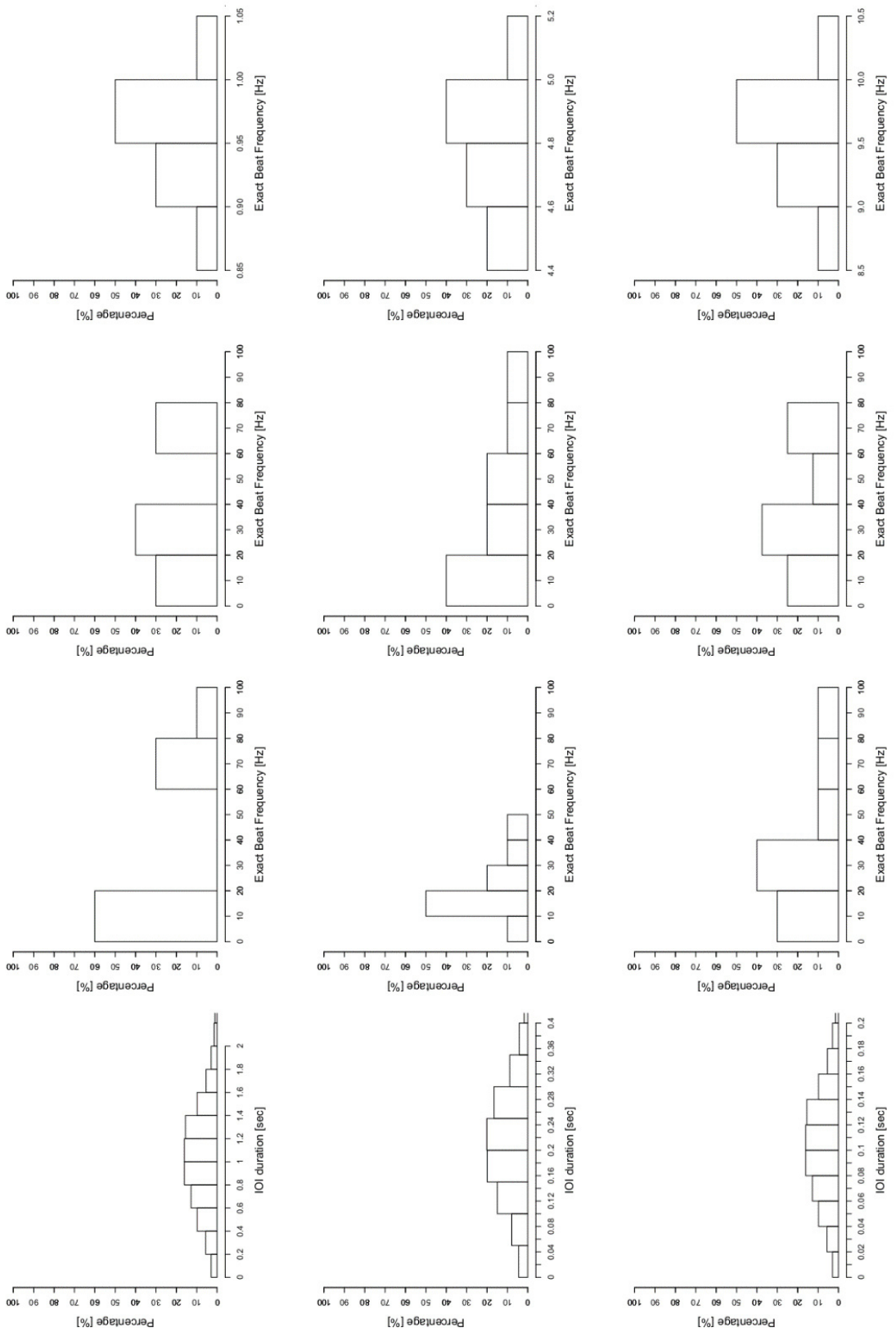


Figure 18: Artificial Data, Results of Rhythm analysis.
Drawn from Gaussian distribution with differing means were three datasets (means of datasets from row 1 to 3: 1 sec, 0.2 sec, 0.1 sec respectively). The first column depicts the distribution of Inter-Onset-Intervals in the 10 sequences (1000 elements) the second to fourth column depict the analysed exact beat frequencies found with: a Generate and test approach, Fourier Analysis and Inter-Onset-Interval analysis.

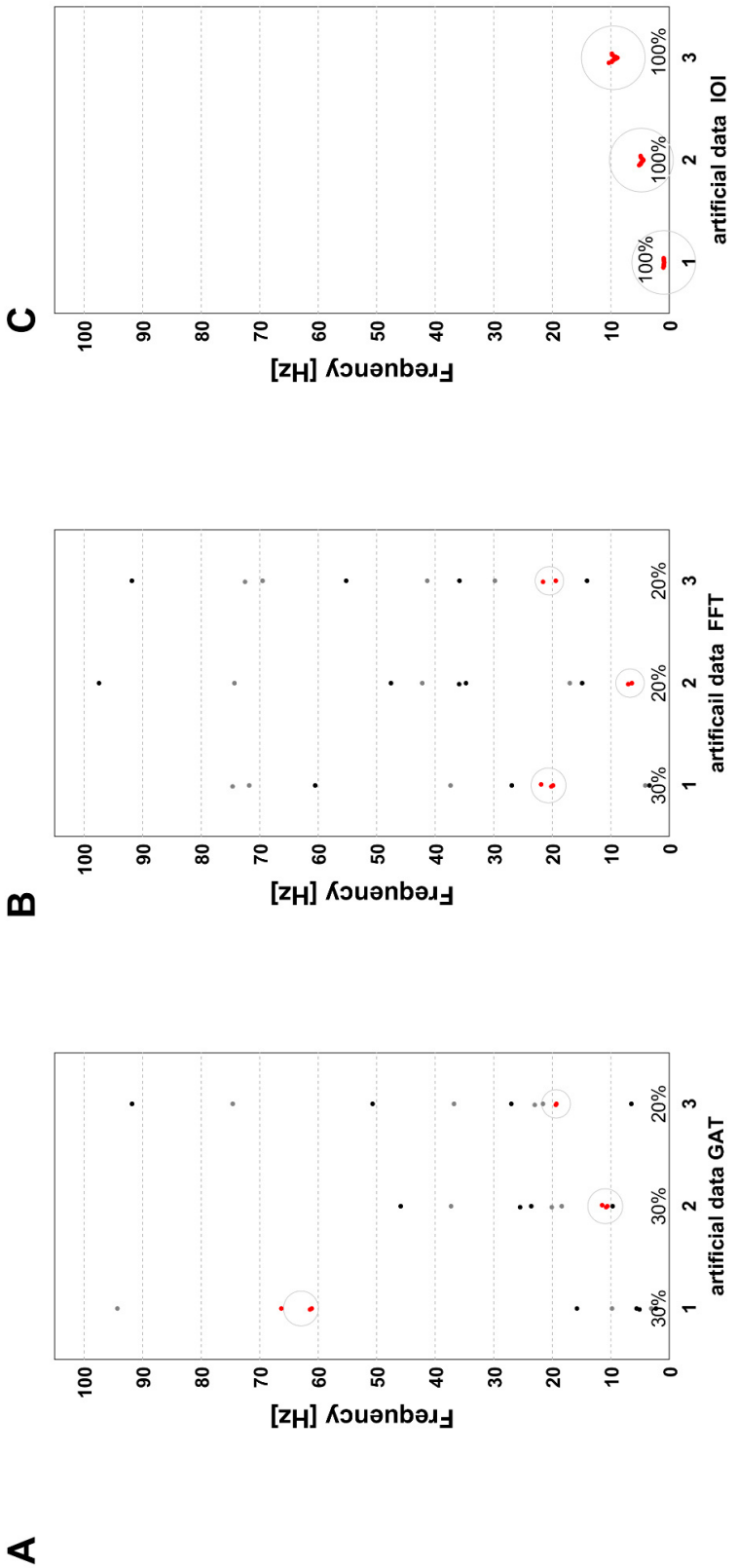


Figure 19: Cluster Analysis of Artificial Data.
(A) results for the GAT approach, (B) results for the Fourier analysis, (C) results for the Inter-Onset-Interval analysis. The numbers on the bottom indicate the sub dataset. '1' stands for the dataset drawn from a distribution with a mean of 1 with a SD of 0.5, '2' is the dataset with a mean of 0.2 seconds and SD of 0.1 and '3' is the dataset with a mean of 0.1 seconds and SD of 0.05.

Chapter III

Adjusted Auditory Brainstem Response Procedure to Measure Rhythm Perception in Small Mammals.

Abstract

The production of rhythmic signals by animals has received increasing scientific attention in recent years but knowledge about the perception of different temporal patterns is still scarce. We developed a method to quantify rhythm perception in small mammals. In an auditory brainstem response (ABR) experiment the direct response of the auditory nuclei to an auditory stimulus was measured. By varying the presented stimulus rates, we adjusted the classical ABR paradigm to measure rhythm perception. We recorded ABRs from 78 wild living individuals of 12 bat species from tropical America in response to pure-tone stimuli (in 17 rates ranging from 6 to 100 Hz) and from 20 individuals of captive bred *Carollia perspicillata* presented with natural stimuli (i.e., isolation calls of *C. perspicillata* at rates of 6, 25, and 44 Hz). A general decline of response strength towards higher stimulus presentation rates could be found for all species tested with pure tones. No clear differences were found between the three presentation rates for natural stimuli tested in *C. perspicillata*. Natural stimuli on average elicited higher reactions than artificial stimuli. Furthermore, the measured response strengths for different perception rhythms were compared to vocal production rhythms of the respective species' echolocation calls and a limited overlap between production and perception rhythms was found.

We confirmed the applicability of the adjusted ABR procedure to measure rhythm perception in small mammals. Using bats from tropical America as an example, we found differences in perception strength dependent on the stimulus presentation rate for artificial and natural stimuli in untrained wild and captive bats. Our results are important for future considerations of stimulus choice in ABR experiments, as they suggest preferring natural stimuli and slower stimulus presentation rates.

Introduction

Audiometry, the research of acoustic perception, focuses on studying the processes of acoustic perception, auditory processing, the involvement of specific brain areas in these processes, and ultimately investigates acoustic information encoding (e.g., (García-Rosales, Beetz, et al., 2018; Land et al., 2016; Linnenschmidt & Wiegrebe, 2019; Portfors & Wenstrup, 1999; Wetekam et al., 2020)). One non-invasive way to study a part of auditory perception are auditory brainstem response (ABR)

measurements that were developed in the 1970s by Jewett and Williston. An ABR visualizes and measures the neuronal activity in the brainstem part of the auditory pathway as field potentials (Jewett & Williston, 1971). The auditory pathway includes five brainstem nuclei: the cochlear nucleus, the superior olivary complex, the nucleus of the lateral lemniscus, the inferior colliculus and the medial geniculate (Claesdotter-Hybbinette et al., 2016; Purves et al., 2012). ABRs are a widely used standard technique to assess hearing in humans and other animals (Hecox & Galambos, 1974; Juselius Baghdassarian et al., 2018; Land et al., 2016; Obrist & Wenstrup, 1998; Szymanski et al., 1999).

In this study, we repurposed the classical model of ABRs to investigate the perception of isochronous rhythms (i.e., metronome like rhythms with the same interval between all elements) in different tempi, using the integrated response strength measured in the ABRs as a proxy for how well a stimulus is perceived in the anaesthetized animal. Animals can produce sounds in a variety of different temporal structures. A growing body of research is looking into temporal patterns in animals' acoustic signals focusing on mammals and birds including courtship signals, offspring-parent interactions, contact calls, echolocation signals of bats and whales and other types of social communication (Burchardt & Knörnschild, 2020; Burchardt et al., 2019; Heinsohn et al., 2017a; Mathevon et al., 2017; Merguerditchian et al., 2018; Norton & Scharff, 2016; Ravignani, in press; Ravignani et al., 2019; Ravignani & Norton, 2017). As important as knowledge about production rhythms is the question of how animals can perceive these rhythms. For instance, we know that male northern elephant seals use the temporal structure (amongst others) to discriminate between male opponents and must therefore be able to perceive differences in beat frequencies (i.e., a concrete rhythm of a sequence described in Hertz as in beats per second) (Mathevon et al., 2017). In woodpeckers, different species use different pecking patterns ranging from steady short pecking bouts to bouts increasing or decreasing in tempo. The woodpeckers can discriminate between general patterns to differentiate between con- and heterospecifics (Garcia et al., 2020).

To date, it is unclear whether different tempi are perceived equally well or why this would be reasonable to assume. Bats are an interesting candidate taxon to answer this question. In various bat species, the production of echolocation calls is coupled to wing beat frequencies in situations where the sensory needs allow it, that is during search flights (Kalko, 1994; Moss et al., 2006; Schnitzler, 1971; Suthers et

al., 1972; Wong & Waters, 2001). On the other hand, production rhythms of echolocation calls need to be very flexible, and depending on the sensory needs can increase substantially compared to search flight rhythms for example in environments with a lot of clutter or in the feeding buzz immediately before insect capture (Moss et al., 2006; Moss & Surlykke, 2001; Ratcliffe et al., 2013). In a previous study on the greater sac-winged bat *Saccopteryx bilineata* (Emballonuridae), we found that acoustic signal rhythms match wing beat frequencies also while bats are perched, for example in isolation calls uttered by pups and territorial song uttered by males. Independent of context (search flight echolocation or social communication), age, and individual, this species uses rhythms of around 6 to 20 Hz (corresponding to Inter-Onset-Intervals of ~166 – 50 ms (Burchardt et al., 2019)). Even taking the variability of echolocation call rhythms in cluttered environments or a feeding situation into account, bats of the species *S. bilineata* are exposed to this range of acoustic beats or rhythms throughout their entire lifetime in various situations. We therefore hypothesize that their sensory apparatus is specifically tuned to these rhythms. In species where different social communications and echolocation calls are uttered in very different rhythms, such sensory tuning is less likely, and we would expect their rhythm perception to be similar over a range of different rhythms. We further hypothesize that we can show such differences in rhythm perception with the adjusted ABR paradigm. This hypothesis is strengthened by previous findings showing that not only pitch perception but also tempo perception is taking place in bat auditory nuclei (Portfors & Wenstrup, 1999). In the moustached bat, *Pteronotus parnellii* (Mormoopidae), neurons in the inferior colliculus, one of the auditory nuclei in the brainstem, facilitate response strength by echo delay-tuning (Portfors & Wenstrup, 1999). Furthermore, the inferior colliculus also shows a high temporal precision in electrophysiological recordings of Seba's short-tailed bat *Carollia perspicillata* (Phyllostomidae), a common model organism (Macias et al., 2016). The summed electrical potentials generated by these auditory nuclei can be captured with ABR measurements, which work very well in bats (Burkard & Moss, 1994; Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019; Obrist & Wenstrup, 1998; Wetekam et al., 2020).

In this study, we investigated the hypothesis of specialized rhythm perception in bats with an adjusted ABR procedure. While in a classical ABR experiment different frequencies are presented at a number of different sound levels, we here focused on presenting a single frequency at varying presentation rates,

as was done in a similar way in other species to study adaptation processes in the auditory pathway (Burkard et al., 1994; Burkard et al., 1997; Burkard et al., 2017; Burkard et al., 1996a; Burkard et al., 1996b; Burkard & Voigt, 1989; Ridgway et al., 1981). We conducted these rhythm ABR experiments in 12 wild tropical American bat species, from six different families using artificial stimuli, i.e., pure tones at the species' echolocation call peak frequency at 17 different presentation rates (between 6 and 100 Hz). Furthermore, in one of the species, *C. perspicillata*, we conducted a second rhythm ABR experiment with captive bred individuals to investigate whether this procedure can also be applied using natural stimuli (i.e., isolation calls uttered by pups to solicit maternal care (Knörnschild et al., 2013)). Two different durations of natural stimuli were used: 1) exceptionally short isolation calls matching the duration of the artificial stimuli and 2) isolation calls of average duration, which were 4-fold longer. This approach allowed us to test if natural stimuli that retained their behavioural relevance elicit similar reactions as artificial stimuli. These natural stimuli were presented at three different rates (i.e., 6, 25, and 44 Hz). Overall, we established, for the first time in bats, an adjusted ABR procedure allowing the direct and non-invasive measurement of rhythm perception using the integrated response strength as a proxy for perception.

Methods

Animals and experimental approval

Wild caught animals

ABRs were measured from a total of 78 bats from 12 different species (6 families, Supplementary Table 1). All animals were adult and wild caught around Gamboa, Panama, in March 2019. The research was conducted in accordance with the Panamanian government (MiAmbiente permit 117SE/A-5-19) and the regulations of the Smithsonian Tropical Research Institute (STRI ACUC protocol 1182019-0302-2022).

Animals from captive breeding

ABRs were measured for a total of 20 adult, captive bred bats of the species *C. perspicillata*. We measured 10 males and 10 females. Females were neither pregnant nor lactating. All measurements were conducted under the permission of the responsible administrations of Schleswig-Holstein, Germany (permit number G10-1/19).

Anesthetics application

Experiment 1: Wild caught bats were anesthetized with a combination drug of Medetomidine, Midazolam, and Fentanyl at least 10 minutes before the experiment. The details of the anaesthesia application and dosage, and animal welfare details are published elsewhere (Lattenkamp et al., 2021).

Experiment 2: Bats in captive breeding were anesthetized with a combination drug of Ketamin and Xylazin (7.5 mg/kg Ketamin-hydrochlorid and 16.5 mg/kg Xylazinhydrochlorid). Captive animals weighed between 16 and 24 g, and injection volumina lay between 0.02 and 0.03 ml, lasting for 1 ½ to 2 hours depending on the individual. The drug was injected subcutaneously between the shoulder blades (needle: Sterican® brown 0.45 x 12 mm, B. Braun, Melsungen AG, Melsungen, Germany). During anaesthesia, eye cream (Bepanthen®, 5% Dexpanthenol, Bayer AG, Leverkusen, Germany) was applied to prevent the eyes from drying out. The experiment lasted approximately one hour. After that bats were transferred to a small, padded chamber and given medical oxygen (Compact Mini, 5 L, 200 bar, Air Liquide Healthcare, Düsseldorf, Germany) for 10 to 15 minutes. Bats were transferred back to the colony and kept in single chambers of a temporary holding cage and hand released as soon as they initiated flight, when freely placed on the palm of the experimenter's hand. Water and fruit juice were given as needed, to further facilitate the wake-up process.

ABR setup

The ABRs were measured in a small sound-attenuating box (PELI 1450 case, peli products, Torrance, CA, USA; inside measurements: 37.1 x 25.8 x 15.2 cm^3). Details of the custom-made setup are published elsewhere (Lattenkamp et al., 2021). The box was lined with a copper mesh (for the reduction of electrical interferences), and then covered with sound-attenuation foam (to reduce acoustic reverberation). The bats were positioned with their outer ear opening at a distance of 4 cm and on one level with the horizontal centre of the loudspeaker (R2004/602000, ScanSpeak, Videbæk, Denmark). The loudspeaker was connected to an amplifier (M032N, Kemo® Electronic, Germany). Both stimulus presentation and ABR recording were done by an audio interface (ADI-2 PRO FS, RME, Haimhausen, Germany), running at a sampling rate of 384 kHz. The sound system of the setup was calibrated to 120 dB for Experiment 1 with a 1/8" measuring microphone (B&K4138 without protective grid, Bruel & Kjaer, Bremen, Germany) connected to a measuring amplifier (B&K Measuring Amplifier Type 2636,

Bruel & Kjaer, Bremen, Germany). The calibration was confirmed for 90 dB with a 1 kHz test tone and a calibration microphone mounted on a level meter (TEAC Df-1, Germany) for Experiment 2.

Stimuli

Wild caught animals – Experiment 1

Recordings of ABRs were done in response to pure tone pip stimuli. The tone pips were sinusoids of 2.5 ms duration (Hanning windowed) with different carrier frequencies according to the species-specific echolocation call peak frequency (Table 11). The tone pips were presented 256 times with 17 different presentation rates (6-100 Hz) at an amplitude of 90 dB peak-equivalent sound pressure level (peSPL). Specifically, the used presentation rates were 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 40, 60, 80, and 100 Hz. As we hypothesized perception to be best around often heard rhythms, and with knowledge that these rhythms lie between 6 and 20 Hz in bats (Burchardt & Knörnschild, 2020; Burchardt et al., 2019), we chose a high resolution for that range and up to 30 Hz, to account for variability. As production rhythms were analysed up to 100 Hz before, we also analysed perception rhythms up to 100 Hz (Burchardt et al., 2019; Norton & Scharff, 2016). The order of the presentation rates was randomized for each measurement and individual. Every other presented tone pip was phase-inverted to cancel out electrical stimulus artefacts picked up by the ABR electrodes after averaging in the time-domain. Stimuli were generated at a sampling rate of 384 kHz and a digital word length of 24 bit. A custom written MATLAB script (Matlab, R2018b, MathWorks, Natick, NA, USA) was used to generate the stimuli and coordinate their presentation via the above-mentioned audio interface.

Animals in captive breeding – Experiment 2

Recordings of auditory brainstem responses were done in response to natural stimuli, namely isolation calls of the tested species *C. perspicillata* (for details on dataset and recording see (Knörnschild et al., 2013)). The mean duration of isolation calls in the dataset was 7.3 ms and the minimum duration was 1.7 ms. We selected 10 short isolation calls with a duration of 2 ms from 10 different individuals. Furthermore, 10 average long stimuli of around 8 ms durations were selected from the same set of recordings and individuals, so that we always had one short and one long isolation call of the same individual serving as stimuli in the experiments (Figure 24). In total we presented each individual with 20 stimuli, i.e., 10 short and 10 long natural isolation calls. The presentation rates were 6, 25, and 44

Hz, and were again presented in a randomized order. Presentation rates were chosen so that the known production rhythms of isolation calls were presented (~25 Hz (Burchardt & Knörnschild, 2020)). We chose to then test one significantly slower and one significantly faster presentation rate. 6 Hz was chosen to have the same lower limit as Experiment 1. This results in a difference of 19 Hz between 6 and 25 Hz, the faster presentation rate therefore needed to be 19 Hz faster than 25 Hz resulting in the third presentation rate of 44 Hz. As a total of 20 stimuli were presented to each individual, only three presentation rates could be tested in order to keep the total duration of the experiment within the range of one hour. Natural stimuli had a sampling rate of 384 kHz, to mimic artificial stimuli. Natural stimuli were presented via the same audio-interface and setup that was used in Experiment 1.

ABR recording

Two subdermal electrodes (clipped needles, Sterican® brown 0.45 - 12 mm, B. Braun, Melsungen AG, Melsungen, Germany) were placed at the caudal midline of the head, close to the brainstem (recording electrode) and at the dorsal midline of the head between the ears (reference electrode) (following(Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019). The ground electrode was either placed on the base of the left ear of the animal or on the wing or tail membrane. The electrodes were connected to a bio amplifier (BMA-200, CWE Inc., USA), which bandpass filtered the brainstem responses (between 100 Hz and 3 kHz) and initially amplified the signal by 60 dB. The signal was converted to digital by the above-mentioned audio interface. The ABR signals were down sampled by a factor of 20 to 19.2 kHz before each of the 256 recordings was saved. ABR signals averaged in the time-domain for each combination were displayed to the experimenter for quality monitoring during the measurements.

ABR data analysis

First, the 256 measurements for each presentation rate and stimulus were averaged, multiplied with a calibration dependent scale-factor, and converted to µV. We obtain a timeseries of voltages for each individual, stimulus, and stimulus presentation rate. To minimize noise and restrain values to be positive, a moving-minimum-subtraction-method (Källstrand et al., 2014) was performed. Then the root-mean square (RMS) and standard deviation (STD) value was calculated for a time window of 10 ms after stimulus onset for the transformed curve. Afterwards a trapezoidal integration over the first 10 ms after

stimulus onset was performed in Matlab (Matlab, R2017b, MathWorks, Natick, NA, USA). The result of this integration is an area under the normalized curve in µV. We calculated that area for each presentation rate and individual.

A resampling approach was chosen to confirm the presence of a signal (n = 500; 95% confidence). 500 sequences were simulated by drawing randomly 192 samples (equivalent to 10 ms, with replacement) from the whole averaged and converted original data of the respective individual and stimulus presentation rate. The 500 simulated sequences consisted of 256 measurements each, just like the original data. Those 256 measurements were averaged, and the moving-minimum procedure applied. For each of the 500 simulated sequences, we then calculated the RMS and STD. We calculated the percentage of RMS and STD values being lower than the original values. A signal was regarded as being present if that percentage was above 95. This was done following the procedure of other studies to statistically verify the presence of an ABR signal (Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019; Lv et al., 2007; Wetekam et al., 2020).

For results of Experiment 1 with artificial stimuli, responses were scaled to the measurements of the stimulus at 6 Hz presentation rate. To scale the result, we subtracted the integral value of the 6 Hz rate from all other presentation rates. For measurements of natural stimuli, this scaling was not necessary. Results of 6 Hz, 40 Hz, and 100 Hz from Experiment 1 were compared with a Friedman test and post hoc Dunn's Comparison, while results for different presentation rates and sexes in Experiment 2 were tested via a Kruskal-Wallis test followed by a Dunn's Multiple Comparisons test (Prism 5, GraphPad, San Diego, USA).

Rhythm analysis

We obtained recordings of echolocation calls in free flight for all 12 bat species tested in Experiment 1 (see Supplementary Information, Table 12). We determined the production rhythms of their echolocation calls using their respective inter-onset-interval (IOI). A beat frequency was calculated by taking the inverse average IOI of a sequence. IOI calculations are described in more detail elsewhere (Burchardt & Knörnschild, 2020). Detailed information on the datasets (i.e. study location, recording situation, devices) can be found in the Supplementary Information (Table 12) or in already published literature (*S. bilineata*: (Knörnschild, Jung, et al., 2012b), *Lonchorhina aurita*: (Gessinger et al., 2019),

Rhynchonycteris naso: (Jung, Kalko, & von Helversen, 2007)). Exemplary results for five of the tested families are shown in the main manuscript. Results for all species are shown in the Supplements in Table 14 and Figures 25-36.

Results

We were able to measure rhythm perception with an adjusted ABR procedure with a portable setup, in wild and captive untrained and anesthetized animals with artificial and natural stimuli. We obtained rhythm ABR results for 12 wild Central American bat species (artificial stimuli), including *C. perspicillata.* Furthermore, ABRs were obtained for captive bred individuals of the species *C. perspicillata* (natural stimuli). The results are organized in three parts: (1) an overview about the perception of all 12 species and a detailed illustration of the perception in *C. perspicillata* (Experiment 1), (2) a comparison between the different natural stimulus types (Experiment 2) and between the relevant natural and artificial stimulus types, and (3) a correlation between production rhythms and perception rhythms for five of the analysed species (details on the other species are provided in the Supplementary Information).

Experiment 1 – Artificial Stimuli

Results were significant for all 12 species, i.e., all measured brain responses differed significantly from noise level, as was detected by a resampling approach. 78 wild individuals of 12 Central American bat species from six different families were measured and tested with artificial stimuli for 17 different stimulus presentation rates.

Overview of Rhythm Perception in 12 Bat Species

All measured species showed a decline in response strength from slow to fast rhythms (Figure 20). We depicted the averaged response per species, scaled to the response at 6 Hz. Only few stimulus presentation rates resulted in higher responses than at 6 Hz. Only for one species, *Rhynchonycteris naso*, a substantial number of responses (six presentation rates: 10, 12, 20, 30, 40, and 60 Hz) was stronger than the response at a presentation rate of 6 Hz (Supplements, Figure 25).

Nevertheless, all species showed a significant decline when comparing 6 Hz with a presentation rate of 100 Hz (Figure 20C, $p < 0.0001$***, Friedman test with Dunn's Multiple Comparison Test, groups = 3, Friedman statistic = 20.67, difference in rank sum = 22). The decline was significant between 40 Hz and

100 Hz, but not between 6 Hz and 40 Hz (Figure 20C, 6 Hz vs 40 Hz: difference in rank sum = 8, ns; 40 Hz via 100 Hz: difference in rank sum = 14, $p < 0.05$*). The integrated response strength declined between 22.8 and 256.5 µV with a median decline of 32.6 µV from 6 Hz to 100 Hz (Supplements, Table 13).

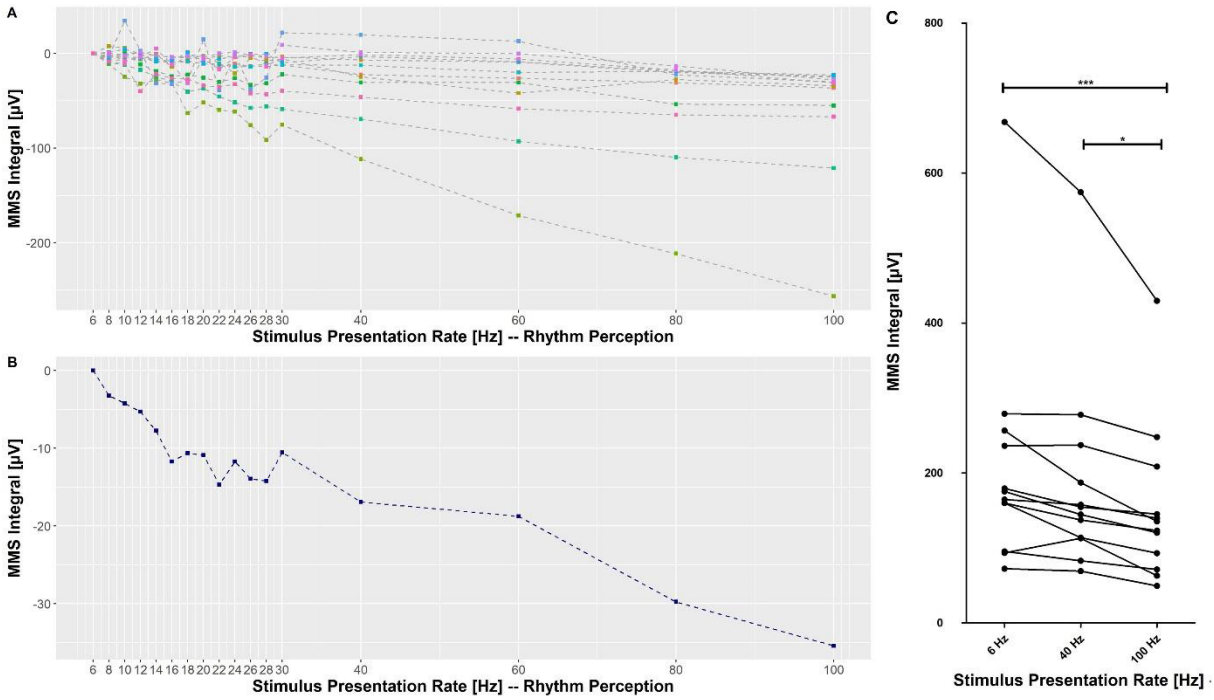


Figure 20: ABR Response Strength of 12 Central American Bat Species Depicted in µV.
(A) ABR response strength scaled to the 6 Hz response strength. All 12 species showed a decline in response strength from 6 Hz to 100 Hz. (B) Averaged ABR response strength of 12 bat species from 6 Hz to 100 Hz scaled to 6 Hz. (C) Comparison between the ABR response strength at 6 Hz vs. 100 Hz for all 12 species; the response strength at 6 Hz was significantly higher than at 100 Hz (Wilcoxon signed rank test, one-tailed, p = 0.0013, positive rang sum 78, negative rang sum 0.00)**

Detailed Results for *Carollia perspicillata*

We detected individual differences in ABR waveform between different presentation rates, as well as in general ABR response strength between individuals. This can be illustrated by the results of *C. perspicillata* (Figure 21). We found a clear difference between presentation rates of 6 Hz and 100 Hz with regards to general response strength, but also in the distinctiveness of single peaks. These differences are not discussed further, as we only looked at the integrated area under the whole transformed curve for subsequent analysis, which we interpret as the general response strength. Individuals all showed a general decline of response strength from slow to fast rhythms but decline strength as well as individual response peaks differed (Figure 21C). The averaged response showed an

almost linear decline (Figure 21D). Detailed results for the remaining 11 species are shown in the Supplementary Information (Figures 25- 36).

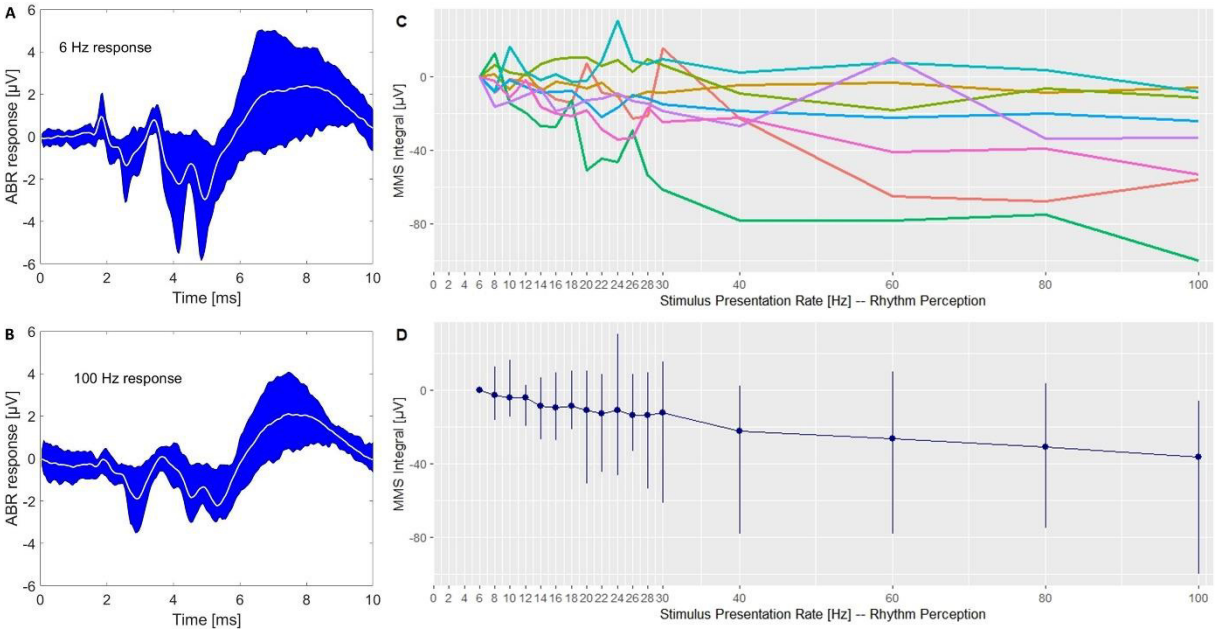


Figure 21: Detailed Results of *C. perspicillata* for Experiment 1 Presenting Artificial Stimuli to Free Living Bats.
(A, B) The averaged raw traces for the first 10 ms of the ABR response for the presentation rates 6 and 100 Hz are displayed. The slopes depict an average as well as minimum and maximum (blue area) of all eight measured individuals. (A) Average ABR slope of eight *C. perspicillata* individuals at a 6 Hz stimulus presentation rate. (B) Average ABR slope of eight measured *C. perspicillata* individuals at 100 Hz stimulus presentation, showing a generally lower response and less distinct peaks. (C) Single measurements for eight *C. perspicillata* individuals, all showing a general decline in response strength towards faster rhythms of up to 100 Hz. Shown is the moving-minimum subtraction (mms) transformed data. (D) The scaled averaged response strength (as calculated as trapezoidal integral on mms transformed data) for eight *C. perspicillata* individuals is depicted. Error bars are the minimum and maximum values measured.

Experiment 2 – Natural Stimuli

In a second experiment, 20 captive individuals (10 males and 10 females) of the species *C. perspicillata* were tested with natural stimuli, namely isolation calls of the respective species. Two different kinds of natural stimuli were presented: 2 ms long isolation calls and 8 ms long isolation calls. We presented 10 calls of each length to every tested specimen in three presentation rates. Therefore, every individual was exposed to 60 combinations. At least 30 of the combinations needed to result in a significant ABR response for us to regard the individual in the reported results. Four males and three females did not reach that criterion and are therefore not shown in the results. From all other individuals, only significant responses were taken into account (average of 95.5% over all remaining 13 individuals). Significant responses were averaged for all 10 stimuli per individual and duration.

Most intriguingly, significant responses to natural stimuli were consistently higher than to artificial stimuli as measured in the same species in Experiment 1 (Figure 22A). For natural stimuli, response strength was similar between males and females for both long and short isolation calls (Kruskal-Wallis test, p = 0.1189, groups = 12, Kruskal-Wallis statistic = 16.64). We could not make this comparison in wild *C. perspicillata* exposed to artificial stimuli because only males were recorded.

When comparing results within females via paired testing, there were differences between stimulus presentation rates but also between short and long stimuli. We showed significant differences only for comparisons with either the same stimulus presentation rate or with the same stimulus duration. Females' responses to short isolation call stimuli at 25 Hz presentation rate differed significantly from long isolation calls, with long isolation calls eliciting lower reactions (Friedman test with Dunn's Multiple Comparison Test, p = **, difference in rank sum = 24, Figure 22B). The same was true for stimulus presentation rates of 44 Hz, again long isolation calls resulted in lower responses (Friedman test with Dunn's Multiple Comparison Test, p = *, difference in rank sum = 22, Figure 22B). No significant differences in the corresponding pairs in males were found, even though long stimuli tended to elicit lower reactions than short stimuli (Figure 22B). There were also no differences between stimulus presentation rates within one stimulus type for males or females.

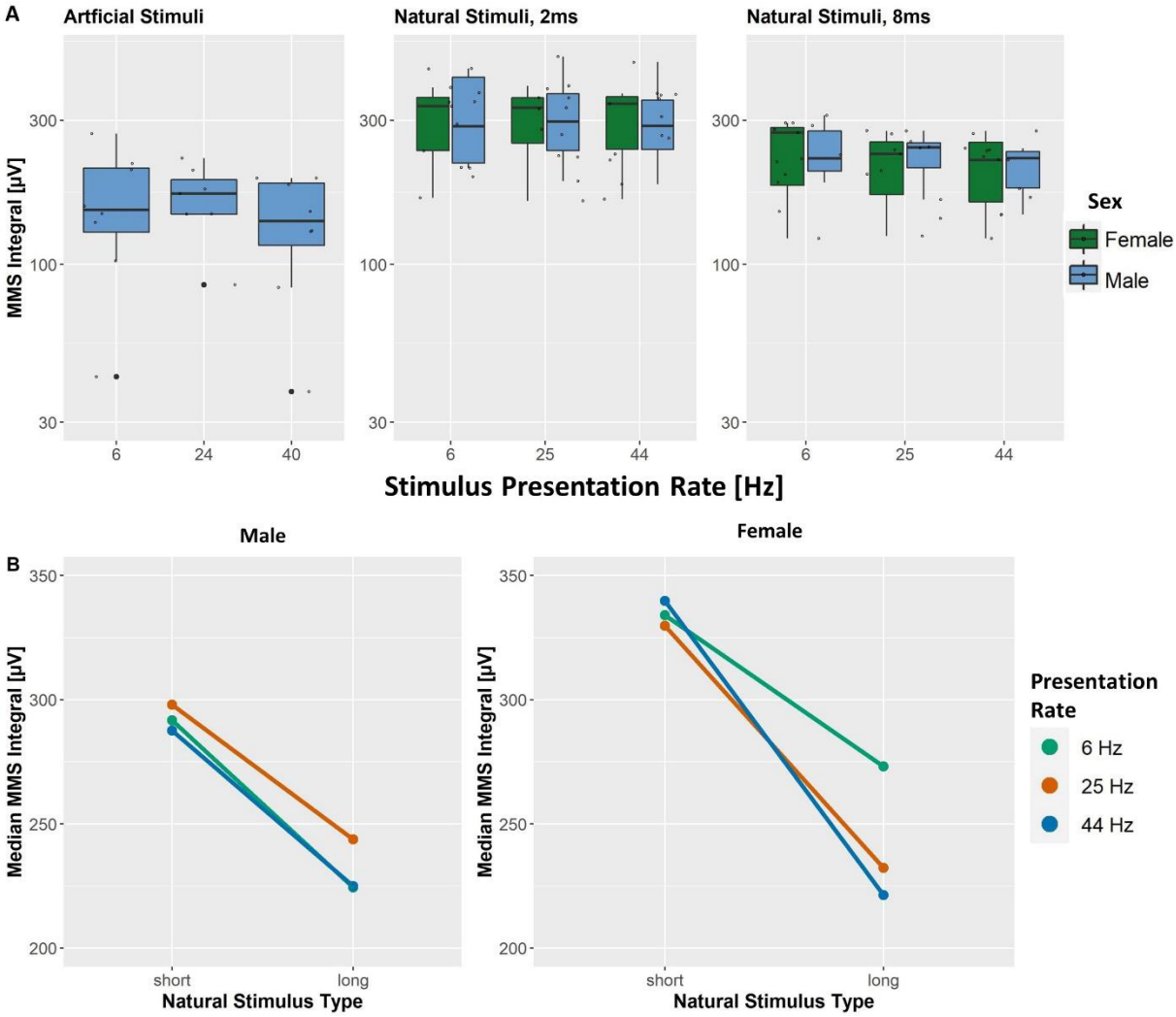


Figure 22: Detailed Results of *C. perspicillata* for Experiment 2.
(A) Comparison between Experiment 1 and 2 for *C. perspicillata*. Natural stimuli elicited higher reactions compared to artificial stimuli. There was a mild decline from 6 Hz to 100 Hz in both experiments, confirming the overall results. Males (n = 6) and females (n = 7) did not respond differently to different presentation rates and different stimulus types. (B) Difference in response strength between short and long natural stimuli. Long stimuli tend to elicit lower reactions, this change is only significant for females for stimulus presentation rates of 25 Hz and 44 Hz.

Comparison between Production and Perception Rhythms

To test if commonly produced rhythms are perceived better, we compared the most prominent rhythms of echolocation call production (during search flight) of five species from five families to their responses to artificial stimuli. This comparison is guided through the visual assessment of results in Figure 23. The results for the other species are summarized in the Supplementary Material.

We focused on the five species *Saccopteryx bilineata* (Emballonuridae, Figure 23A), *Molossus molossus* (Molossidae, Figure 23B), *Pteronotus parnellii* (Mormoopidae, Figure 23C), *Lonchorhina aurita*

(Phyllostomidae, Figure 23D), and *Myotis nigricans* (Vespertilionidae, Figure 23E) which highlight the diverse relation between rhythm production and perception in five bat families.

For *S. bilineata*, we did not observe large differences in ABR signal strength between presentation rates of 6 and 40 Hz, even though most echolocation sequences are produced at 8 to 12 Hz during search flight (Burchardt et al., 2019). Thus, in this species there is a complete overlap of production rhythms and perception rhythms but not a clear correlation between the two (Figure 23A). For *M. molossus* we observed the expected pattern: search flight echolocation sequences were uttered with rhythms corresponding well to the rhythms that elicited higher responses in the ABR experiments. Thus, production rhythms and perception rhythms seemed to be correlated in this species (Figure 23B). For *P. parnellii* production rhythms of search flight echolocation overlapped well with stimulus presentation rates, eliciting the highest reactions in this species as well (Figure 23C). For *L. aurita*, we also observed the expected pattern: search flight echolocation sequences were uttered with rhythms corresponding well to the rhythms that elicited the highest responses in the ABR experiments. Thus, production rhythms and perception rhythms seemed to be correlated in this species (Figure 23D). For *M. nigricans*, we see the same picture as for *M. molossus* and *L. aurita*, production rhythms correlated well with presentation rates that elicited high reactions in the ABR procedure (Figure 23E).

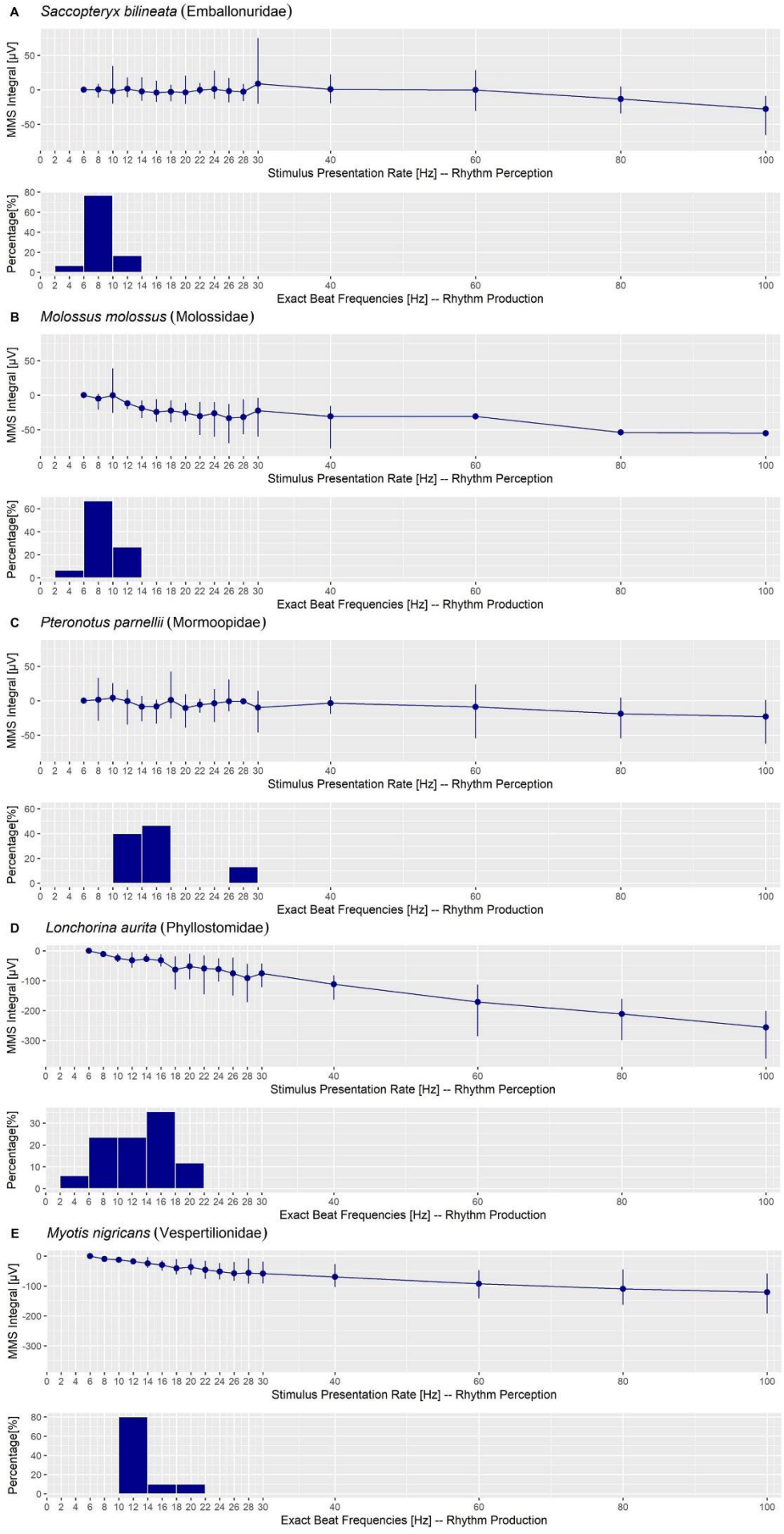

A Saccopteryx bilineata (Emballonuridae)
MMS Integral [µV]
Stimulus Presentation Rate [Hz] -- Rhythm Perception
Percentage[%]
Exact Beat Frequencies [Hz] -- Rhythm Production
B Molossus molossus (Molossidae)
MMS Integral [µV]
Stimulus Presentation Rate [Hz] -- Rhythm Perception
Percentage[%]
Exact Beat Frequencies [Hz] -- Rhythm Production
C Pteronotus parnellii (Mormoopidae)
MMS Integral [µV]
Stimulus Presentation Rate [Hz] -- Rhythm Perception
Percentage[%]
Exact Beat Frequencies [Hz] -- Rhythm Production
D Lonchorina aurita (Phyllostomidae)
MMS Integral [µV]
Stimulus Presentation Rate [Hz] -- Rhythm Perception
Percentage[%]
Exact Beat Frequencies [Hz] -- Rhythm Production
E Myotis nigricans (Vespertilionidae)
MMS Integral [µV]
Stimulus Presentation Rate [Hz] -- Rhythm Perception
Percentage[%]
Exact Beat Frequencies [Hz] -- Rhythm Production

Figure 23: Correlation Between Rhythm Perception and Production in Five Bat Species.
Rhythm perception is shown in the line graph as the integrated response strength scaled to the response at 6 Hz. Rhythm production is shown as a histogram; rhythms were analysed with IOI analysis in search flight echolocation call sequences. (A) In *S. bilineata*, rhythms that were perceived well overlapped with rhythms that were produced often. (B) In *M. molossus* rhythms that were perceived well were produced most often. (C) In *P. parnellii* rhythms that were perceived well overlapped with rhythms that were produced often. (D) In *L. aurita*, rhythms that were perceived well were produced most often. (E) In *M. nigricans*, rhythms that were perceived well were produced most often.

Discussion

We analysed the rhythm perception of 12 species of Central American bats with an adjusted workflow adapting the principles of ABR recordings to measure differences in rhythm perception in small mammals. The method described here was used for the first time in bats and successfully tested on a total of 98 individuals. Differences in perception were found on an individual as well as species level. This opens a new field of research, as this method can be easily adopted for other small mammals, such as rodents, or lagomorphs. Thus, our new method could facilitate future studies on beat perception that contribute to the growing field of research on rhythm in animals' acoustic signals, for instance the role of rhythm in vocal production learning (Hyland Bruno & Tchernichovski, 2017; Wirthlin et al., 2019) or individual recognition (Mathevon et al., 2017).

The ABR method has clear advantages over behavioural tests, as it is faster and more standardized than for example playback experiments or other behavioural experiments involving the time-consuming training of animals (i.e., (Koay et al., 1997; Lattenkamp et al., 2018)). With ABRs it is possible to directly measure the spontaneous reaction of an animal on an individual level and control for most parameters. Of course, ABRs do not replace behavioural experiments as these are likely to produce more sensitive results (Heffner et al., 2008), but they can provide first insights into the hearing capacity of a species and save a lot of time and effort, because potential subsequent behavioural experiments can be set up in a more informed way. It is a minimally invasive technique to measure summed neuronal responses without harming the animal and avoiding unnecessary stress due to long periods of isolation and training.

In this study, we compared ABR strength in response to a wide range of stimulus presentation rates to the echolocation call rate of 12 species. We expected to find that the auditory perception process is tuned

to rhythms perceived very often, as for example species-specific search flight echolocation call rhythms, coupled to wingbeat frequencies(Burchardt et al., 2019; Kalko, 1994; Schnitzler, 1971; Wong & Waters, 2001). However, despite the enormous variation in production rhythms of echolocation call sequences between species we found that all 12 tested species showed the strongest ABR signals in response to slow stimulus presentation rates of 6 Hz to 20 Hz. Production rhythms can lie at around 100 Hz and higher in *Thyroptera tricolour*, while they lie at around 6 to 12 Hz in *S. bilineata* (see Supplements and (Burchardt et al., 2019)). Corresponding results, demonstrating the decrease of reaction strength with increasing stimulus presentation rates, were found in electrophysiological recordings in the cortex of *C. perspicillata.* These recordings revealed that brain processes involved in the perception of echolocation calls (as a very specialized vocalization used for orientation, a system that can only be found in bats and toothed whales) appear to be far less specialized than one could expect. Cortical neurons were only able to track acoustic stimuli at frequencies (rates) of up to, but not faster than 22 Hz (Martin et al., 2017). This might hold true not only for the processing in the cortex, but also in the more primary brainstem, thus matching our results.

Similar electrophysiological recordings in the inferior colliculus (IC) of *P. parnellii* indicated that the IC includes delay-tuned neurons that facilitate response strength through delay. The "best-delay" ranges from 0 to 20 ms (Portfors & Wenstrup, 1999). Delaying every stimulus by i.e., 20 ms, resembles a stimulus presentation rate of 50 Hz (as 1000/ 20 ms equals 50 Hz). Delaying it by less than 20 ms would then fit with even higher stimulus presentation rates. Overall, the described delay-tuning should possibly facilitate the response strength for fast presentation rates of 50 Hz and higher. However, this potential facilitation is not visible in our results when looking at the overall response strength. Nevertheless, it might be worthwhile for future studies to specifically test a possible effect of delay-tuning on the ABR peak corresponding to the response of the IC. One explanation for our finding of improved perception for slow rhythms could be that these patterns represent temporal sensitivity for social calls rather than echolocation call rates. This would be especially interesting to study in contrasting species such as *Saccopteryx bilineata* (i.e., with slow social call and echolocation call rates (Burchardt et al., 2019)) and *T. tricolour* (with very fast echolocation, but slow social call rates (Supplements and (Chaverri & Gillam, 2015)).

While we can exclude habituation effects shaping the results, because stimuli rates were presented in a random order, we cannot exclude the possibility of an adaptation process taking place. The higher stimulus energy of faster rhythms could result in neuronal adaptation and in an overall lower reaction strength. Previous studies studying adaptation processes in the auditory brainstem by manipulating click rates in ABR recordings found a decrease of the peak amplitude towards higher presentation rates in cats, gerbils, mice, chicken, humans, and even in echolocating bottle-nosed and common dolphins (Burkard et al., 1994; Burkard & Sims, 2002; Burkard et al., 1997; Burkard et al., 2017; Burkard et al., 1996a; Burkard et al., 1996b; Burkard & Voigt, 1989; Ridgway et al., 1981). Adaptation is assumed to lead to an almost linear decline in response strength towards higher presentation rates (Wiegrebe & Schmidt, 1996), as it was observed in the bottle-nosed dolphin (Burkard et al., 2017). However, the decline we observed in the present study was not very linear in many species. Results look linear between presentation rates of 40 Hz and 100 Hz, which is at least in parts due to the coarse resolution of measurement points. Between stimulus presentation rates of 6 Hz and 30 Hz – where we measured with a higher resolution of presentation rates – much more variability can be seen, and the curves could not well be described as linearly declining (Figure 20A). This gets even more apparent when looking at individual results (i.e., Figure 21C and Supplements, Figures 37 - 48). Furthermore, even though our results show an increase in ABR strength in response to lower stimulus rates, the reaction strength varies strongly between the measured species. While for example the decrease in response strength between 6 Hz and 100 Hz in *L. aurita* lies at 256.5 µV it only lies at 22.8 µV in *P. parnellii* (Figure 20C, Supplements, Table 13). Adaptation might influence the results, but we cannot determine to which extent.

Aside from a general decline of ABR strength with increasing stimulus rate in all species, we detected differences between individuals which might be caused by the differences in hearing capacity between the individual animals (i.e., different hearing capacities due to age). This potential sampling bias could have affected results especially for species were only few individuals were tested (i.e., *L. aurita* with three tested individuals). Observed differences between individuals might have been caused by slight differences in electrode positioning, age of the individuals, or the amount of mastication muscles. The more muscles, the farther away the electrodes could be placed from the brainstem, thus leading to a

smaller received signal (Linnenschmidt & Wiegrebe, 2019). Furthermore, differences between species could be caused by a general difference in anatomy and mastication muscles, due to differences in diet and feeding style (Herrel et al., 2008; Santana, 2018; Santana et al., 2010).

An important aspect that is being disregarded in this setup is attention. As animals were anesthetized for the ABR measurements, attention driven reactions were suppressed. How attention would affect the perception of different rhythms is unclear and could only be addressed with subsequent behavioural tests. The results of this study can help to set up such behavioural tests in a more informed way.

The results of this study are of further importance for general stimulus selection and presentation in ABR experiments. Stimulus presentation rate has been broadly neglected in the optimization of ABR recordings. Our findings demonstrate that stimulus presentation rate plays a critical role in ABR experiments and needs to be considered during experimental design. Furthermore, it is important to consider the characteristics of the presented stimuli. ABR strength in response to natural stimuli was heightened in comparison to artificial stimuli. Due to the associated stimulus complexity, natural stimuli are rarely used in ABR experiments. However, the use of natural stimuli is gaining more importance in recent years (Lahti & Foster, 2015; Talebi & Baker, 2012). Our findings show increased reaction strength to natural stimuli and are in line with electrophysiological recordings in *C. perspicillata* showing that the temporal tuning in higher brain areas such as the cortex works better with natural stimulus sequences (Beetz et al., 2016). In this study, we decided to focus on tone pips of the species-specific peak-frequency and selected natural stimuli. These two stimulus types were used to ensure that the bats would be able to perceive them well. For future studies in bats, but also in other species, it would be interesting to further study rhythm perception of clicks (broad band noise stimuli) (Burkard et al., 1994; Burkard & Sims, 2002; Burkard et al., 1997; Burkard et al., 2017; Burkard et al., 1996a; Burkard et al., 1996b; Burkard & Voigt, 1989) or tone pips with important frequencies from social calls as well as more natural stimuli.

Taken together, we can confirm the general applicability of the adjusted ABR procedure to measure rhythm perception in small mammals. Using Central American bats as an example, we found differences

in perception strength dependent on the stimulus presentation rate for artificial and natural stimuli in untrained wild and captive bats.

Acknowledgements

We thank Constance Scharff and Arpik Nshedjan for their support on experiment approval and logistics; Sina Girra for her help in experiment preparations; Martina Nagy, Florian Gloza-Rausch, and Jasmin Bulang for their help in data collection and Leonie Baier for her help on the Matlab code. We thank Meike Linnenschmidt for teaching the ABR procedure and Kirsten Jung for providing echolocation data for *Rhynchonycteris naso*. Furthermore, we thank Julio Hechavarría and his lab for very helpful discussions. We are grateful to the late Lutz Wiegrebe, our friend, colleague, and mentor, who originally designed the recording setup and whose insights and help were instrumental to this study.

Funding

LSB was funded by an Elsa-Neuman Scholarship of the city of Berlin. MK was supported by a Heisenberg Fellowship (DFG KN935 5-1) and a research grant (DFG KN935 4-1) from the German Research Foundation. EZL was funded by a short-term stipend of the German Academic Exchange Program (DAAD) (no. 57438025) and a travelling fellowship of the Company of Biologists (JEBTF18113). GG was founded by the Heinrich Böll Foundation and received a Smithsonian Fellowship for conducting fieldwork. IG was funded by a Smithsonian Tropical Research Institute Tupper Postdoctoral Fellowship. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

References

Beetz, M. J., Hechavarria, J. C., & Kössl, M. (2016). Temporal tuning in the bat auditory cortex is sharper when studied with natural echolocation sequences. *Scientific Reports, 6*. https://doi.org/10.1038/srep29102

Burchardt, L. S., & Knörnschild, M. (2020). Comparison of methods for rhythm analysis of complex animals' acoustic signals. *PLOS Computational Biology, 16*(4), e1007755. https://doi.org/10.1371/journal.pcbi.1007755

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science, 6*. http://doi.org/10.1098/rsos.181076

Burkard, R., Jones, S., & Jones, T. (1994). Conventional and cross-correlation brain-stem auditory evoked responses in the white leghorn chick: rate manipulations. *The Journal of the Acoustical Society of America, 95*(4), 2136-2144. https://doi.org/10.1121/1.408675

Burkard, R., & Moss, C. F. (1994). The brain-stem auditory-evoked response in the big brown bat (*Eptesicus fuscus*) to clicks and frequency-modulated sweeps. *The Journal of the Acoustical Society of America, 96*(2).

Burkard, R., & Sims, D. (2002). The Human Auditory Brainstem Response to High Click Rates. *American journal of audiology, 10*(2), 53-61. https://doi.org/10.1044/1059-0889(2001/008)

Burkard, R., Trautwein, P., & Salvi, R. (1997). The effects of click level, click rate, and level of background masking noise on the inferior colliculus potential (ICP) in the normal and carboplatin-treated chinchilla. *The Journal of the Acoustical Society of America, 102*(6), 3620-3627. 10.1121/1.420149

Burkard, R. F., Finneran, J. J., & Mulsow, J. (2017). The effects of click rate on the auditory brainstem response of bottlenose dolphins. *The Journal of the Acoustical Society of America, 141*(5), 3396. https://doi.org/10.1121/1.4983447

Burkard, R. F., McGee, J., & Walsh, E. J. (1996a). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. I. Peak latencies. *The Journal of the Acoustical Society of America, 100*(2), 978-990. https://doi.org/10.1121/1.416209

Burkard, R. F., McGee, J., & Walsh, E. J. (1996b). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. II. Peak amplitudes. *The Journal of the Acoustical Society of America, 100*(2), 991-1002. https://doi.org/10.1121/1.416210

Burkard, R. F., & Voigt, H. F. (1989). Stimulus dependecies of the gerbil brain-stem auditory-evoked response (BAER). I: Effects of click level, rate, and polarity. *The Journal of the Acoustical Society of America, 85*(6), 2514-2525.

Chaverri, G., & Gillam, E. H. (2015, Jan). Repeatability in the contact calling system of Spix's disc-winged bat (Thyroptera tricolour). *R Soc Open Sci, 2*(1), 140197. 10.1098/rsos.140197

Claesdotter-Hybbinette, E., Matti Cervin, M., & Akerlun, S. (2016). Gender specific differences in auditory brain stem response in young patients with ADHD. *Neuropsychiatry, 06*(01). https://doi.org/10.4172/Neuropsychiatry.1000114

García-Rosales, F., Beetz, M. J., Cabral-Calderin, Y., Kössl, M., & Hechavarria, J. C. (2018). Neuronal coding of multiscale temporal features in communication sequences within the bat auditory cortex. *Communications Biology, 1*(1), 200. https://doi.org/10.1038/s42003-018-0205-5

Garcia, M., Theunissen, F., Sebe, F., Clavel, J., Ravignani, A., Marin-Cudraz, T., Fuchs, J., & Mathevon, N. (2020). Evolution of communication signals and information during species radiation. *Nature Communications, 11*(1), 4970. https://doi.org/10.1038/s41467-020-18772-3

Gessinger, G., Gonzalez-Terrazas, T. P., Page, R. A., Jung, K., & Tschapka, M. (2019). Unusual echolocation behaviour of the common sword-nosed bat *Lonchorhina aurita*: an adaptation to aerial insectivory in a phyllostomid bat? *Royal Society Open Science, 6*(7), 182165. https://doi.org/10.1098/rsos.182165

Hecox, K., & Galambos, R. (1974). Brain Stem Auditory Evoked Responses in Human Infants and Adults. *Archives of Otolaryngology, 99*(1), 30-33. https://doi.org/10.1001/archotol.1974.00780030034006

Heffner, H. E., Koay, G., & Heffner, R. S. (2008). Comparison of behavioural and auditory brainstem response measures of threshold shift in rats exposed to loud sound. *The Journal of the Acoustical Society of America, 124*(2), 1093-1104. https://doi.org/10.1121/1.2949518

Heinsohn, R., Zdenek, C. N., Cunningham, R. B., Endler, J. A., & Langmore, N. E. (2017). Tool-assisted rhythmic drumming in palm cockatoos shares key elements of human instrumental music. *Science Advances, 3*(6), e1602399. https://doi.org/10.1126/sciadv.1602399

Herrel, A., De Smet, A., Aguirre, L. F., & Aerts, P. (2008). Morphological and mechanical determinants of bite force in bats: do muscles matter? *The Journal of Experimental Biology 211*(Pt 1), 86-91. https://doi.org/10.1242/jeb.012211

Hyland Bruno, J., & Tchernichovski, O. (2017). Regularities in zebra finch song beyond the repeated motif. *Behavioural Processes*. https://doi.org/10.1016/j.beproc.2017.11.001

Jewett, D. L., & Williston, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain, 94*(4), 681-696. https://doi.org/10.1093/brain/94.4.681

Jung, K., Kalko, E. K. V., & von Helversen, O. (2007). Echolocation calls in Central American emballonurid bats: signal design and call frequency alternation. *Journal of Zoology*(272), 125-137. https://doi.org/10.1111/J.1469-7998.2006.00250.X

Juselius Baghdassarian, E., Nilsson Markhed, M., Lindström, E., Nilsson, B. M., & Lewander, T. (2018). Auditory brainstem response (ABR) profiling tests as diagnostic support for schizophrenia and adult attention-deficit hyperactivity disorder (ADHD). *Acta Neuropsychiatrica, 30*(3), 137-147. https://doi.org/10.1017/neu.2017.24

Kalko, E. K. V. (1994). Coupling of sound emission and wingbeat in naturally foraging european pipistrelle bats (Microchiroptera: Vespertilionidae). *Folia Zoologica*(43(4)), 363-376.

Källstrand, J., Lewander, T., Baghdassarian, E., & Nielzén, S. (2014). A new method for analysing auditory brain-stem response waveforms using a moving-minimum subtraction procedure of digitized analog recordings. *Neuropsychiatric Disease and Treatment, 10*, 1011-1016. https://doi.org/10.2147/NDT.S59178

Knörnschild, M., Feifel, M., & Kalko, E. K. V. (2013). Mother-offspring recognition in the bat *Carollia perspicillata*. *Animal Behaviour, 86*, 941-948. https://doi.org/10.1016/j.anbehav.2013.08.011

Knörnschild, M., Jung, K., Nagy, M., Metz, M., & Kalko, E. K. V. (2012). Bat echolocation calls facilitate social communication. *Proceedings of the Royal Society B, 279*, 4827-4835. https://doi.org/10.1098/rspb.2012.1995

Koay, G., Heffner, H. E., & Heffner, R. S. (1997). Audiogram of the big brown bat (*Eptesicus fuscus*). *Hearing Research, 105*(1), 202-210. https://doi.org/10.1016/S0378-5955(96)00208-0

Lahti, D. C., & Foster, S. (2015). The Limits of Artificial Stimuli in Behavioural Research: The Umwelt Gamble. *Ethology, 121*(6), 529-537. https://doi.org/10.1111/eth.12361

Land, R., Burghard, A., & Kral, A. (2016). The contribution of inferior colliculus activity to the auditory brainstem response (ABR) in mice. *Hearing Research, 341*, 109-118. https://doi.org/10.1016/j.heares.2016.08.008

Lattenkamp, E. Z., Nagy, M., Drexl, M., Vernes, S. C., Wiegrebe, L., & Knornschild, M. (2021). Hearing sensitivity and amplitude coding in bats are differentially shaped by echolocation calls and social calls. *Proceedings of the Royal Society B, 288*(1942), 20202600. https://doi.org/10.1098/rspb.2020.2600

Lattenkamp, E. Z., Vernes, S. C., & Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *Journal of Experimental Biology, 221*(Pt 14). https://doi.org/10.1242/jeb.180729

Linnenschmidt, M., & Wiegrebe, L. (2019). Ontogeny of auditory brainstem responses in the bat, *Phyllostomus discolour*. *Hearing Research, 373*, 85-95. https://doi.org/10.1016/j.heares.2018.12.010

Lv, J., Simpson, D. M., & Bell, S. L. (2007). Objective detection of evoked potentials using a bootstrap technique. *Medical Engineering & Physics, 29*(2), 191-198. https://doi.org/10.1016/j.medengphy.2006.03.001

Macias, S., Hechavarria, J. C., & Kossl, M. (2016). Temporal encoding precision of bat auditory neurons tuned to target distance deteriorates on the way to the cortex. *Journal of Comparative Physiology A, 202*(3), 195-202. https://doi.org/10.1007/s00359-016-1067-2

Martin, L. M., Garcia-Rosales, F., Beetz, M. J., & Hechavarria, J. C. (2017). Processing of temporally patterned sounds in the auditory cortex of Seba's short-tailed bat, *Carollia perspicillata*. *European Journal of Neuroscience, 46*(8), 2365-2379. https://doi.org/10.1111/ejn.13702

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals Voices. *Current Biology, 27*(15), 2352-2356.e2352. https://doi.org/10.1016/j.cub.2017.06.035

Merguerditchian, A., Vuillemin, A., & Pruetz, J. D. (2018). *Identifying the ape beat in the wild: rhythmic indvidual signatures from sounds of manual fruit cracking in Fongoli Chimpanzees* The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12),

Moss, C. F., Bohn, K., Gilkenson, H., & Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *PLOS Biology, 4*(4), e79. https://doi.org/10.1371/journal.pbio.0040079

Moss, C. F., & Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *The Journal of the Acoustical Society of America, 110*(4), 2207-2226. https://doi.org/10.1121/1.1398051

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

Obrist, M. K., & Wenstrup, J. J. (1998). Hearing and hunting in red bats (*Lasiurus borealis*, Vespertilionidae): audiogram and ear properties. *Journal of Experimental Biology, 201*, 143-154.

Portfors, C. V., & Wenstrup, J. J. (1999). Delay-Tuned Neurons in the Inferior Colliculus of the Mustached Bat: Implications for Analyses of Target Distance. *Journal of Neurophysiology, 82*(3), 1326-1338.

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., & White, L. E. (2012). *Neuroscience* (5th Edition ed.). Sinauer Associates Inc.

Ratcliffe, J. M., Elemans, C. P., Jakobsen, L., & Surlykke, A. (2013). How the bat got its buzz. *Biology Letters, 9*(2). https://doi.org/10.1098/rsbl.2012.1031

Ravignani, A. (in press). Isochrony, vocal learning and the acquisition of rhythm and melody. *Behavioural and Brain Science, in press*.

Ravignani, A., Dalla Bella, S., Falk, S., Kello, C. T., Noriega, F., & Kotz, S. A. (2019). Rhythm in speech and animal vocalizations: a cross-species perspective. *Annals of the New York Academy of Sciences, 1453*, 79-98. https://doi.org/10.1111/nyas.14166

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speec, movement and animal vocalizations. *Journal of Language Evolution, 2*(1), 4-19. doi: 10.1093/jole/lzx002

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., & Galambos, R. (1981). Auditory brainstem response in dolphins. *Proceedings of the National Academy of Science, 78*(3). https://doi.org/10.1073/pnas.78.3.1943

Santana, S. E. (2018). Comparative Anatomy of Bat Jaw Musculature via Diffusible Iodine-Based Contrast-Enhanced Computed Tomography. *The Anatomical Record, 301*(2), 267-278. https://doi.org/10.1002/ar.23721

Santana, S. E., Dumont, E. R., & Davis, J. L. (2010). Mechanics of bite force production and its relationship to diet in bats. *Functional Ecology, 24*(4), 776-784. https://doi.org/10.1111/j.1365-2435.2010.01703.x

Schnitzler, H. U. (1971). Fledermäuse im Windkanal. *Zeitschrift für vergleichende Physiologie*(73), 209-221.

Suthers, R. A., Thomas, S. P., & Suthers, B. J. (1972). Respiration, Wing-beat and Ultrasonic Pulse Emission in an Echo-locating Bat. *The Journal of Experimental Biology, 56*, 37-48.

Szymanski, M., Bain, D., Kiehl, K., Pennington, S., Wong, S., & Henry, K. R. (1999). Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioural audiograms. *The Journal of the Acoustical Society of America, 106*(2), 1134-1141.

Talebi, V., & Baker, C. L., Jr. (2012). Natural versus synthetic stimuli for estimating receptive field models: a comparison of predictive robustness. *The Journal of Neuroscience, 32*(5), 1560-1576. https://doi.org/10.1523/JNEUROSCI.4661-12.2012

Wetekam, J., Reissig, C., Hechavarria, J. C., & Kössl, M. (2020). Auditory brainstem responses in the bat *Carollia perspicillata*: threshold calculation and relation to audiograms based on otoacoustic emission measurement. *Journal of Comparative Physiology A, 206*(1), 95-101. https://doi.org/10.1007/s00359-019-01394-6

Wiegrebe, L., & Schmidt, S. (1996). Temporal integration in the echolocating bat, *Megaderma lyra*. *Hearing Research, 102*, 35-42. https://doi.org/10.1016/s0378-5955(96)00139-6

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., & Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioural Trait. *Neuron, 104*(1), 87-99. https://doi.org/10.1016/j.neuron.2019.09.036

Wong, J. G., & Waters, D. A. (2001). The Synchonisation of signal emission with wingbeat during the approach phase in Soprano Pipistrelles (*Pipistrellus Pygmaeus*). *The Journal of Experimental Biology*(204), 575-583.

Supplementary Information

Contents:

1. Animals in the Wild

Table 11: ABR Experiment 1: Overview of Animals, Sample Size and Measurement Parameters

Family	Species	N (female/ male)	Peak frequency [kHz]	Level [dB peSPL]	Presentation Rate	Repetitions
Emballonuridae	*Rhynchonycteris naso*	6 (2/4)	100	90	6-100	256
Emballonuridae	*Saccopteryx bilineata*	9 (4/5)	46	90	6-100	256
Emballonuridae	*Saccopteryx leptura*	4 (2/2)	53*(1x) /55(3x)	90	6-100	256
Molossidae	*Molossus molossus*	6 (4/2)	39	90	6-100	256
Mormoopidae	*Pteronotus parnellii*	7 (1/6)	60	90	6-100	256
Phyllostomidae	*Carollia perspicillata*	8 (0/8)	90	90	6-100	256
Phyllostomidae	*Desmodus rotundus*	6 (3/3)	78	90	6-100	256
Phyllostomidae	*Glossophaga soricina*	6 (3/3)	113	90	6-100	256
Phyllostomidae	*Lonchorhina aurita*	3 (2/1)	47	90	6-100	256
Phyllostomidae	*Phyllostomus hastatus*	6 (1/5)	57	90	6-100	256
Thyropteridae	*Thyroptera tricolour*	9 (5/4)	100	90	6-100	256
Vespertilionidae	*Myotis nigricans*	8 (0/8)	55	90	6-100	256

* Tone pip frequencies were verbally communicated for every individual, this was a case of miscommunication; as both frequencies fall into the best hearing range of *Saccopteryx leptura*, we decided to keep the result.

Table 12: Echolocation Recordings: Overview of Recording Situation for All Species

Species	Environment	Location	Year	Sampling rate, depth	Device	Sequences (individuals if known)	Recorded From
Rhynchonycteris naso	Free flying in foraging habitat	Costa Rica Panama	2001-2002	480kHz, 16-bit	USG Benedict v. Laar; AKG microphone	8 (4)	KJ ***(Jung, Kalko, & Helversen, 2007)***
Saccopteryx bilineata	Free flying after being released	Costa Rica	2010	500 kHz, 16-bit	Petterson D1000X	30 (30)	MK
Saccopteryx leptura	Free flying in foraging habitat	Peru Panama	2012 2016	250 kHz, 16-bit depth	Avisoft USG 116Hm, CM-16 microphone	15 (12)	MK
Molossus molossus	Free flying in foraging habitat	Peru Panama	2012 2016	250 kHz, 16-bit depth	Avisoft USG 116Hm, CM-16 microphone	15 (15)	MK
Pteronotus parnellii	Free flying in foraging habitat	Panama	2018	500 kHz	Avisoft USG 116Hm, CM-16 microphone	15 (4)	GG
Carollia perspicillata	Flight room (3.7m * 2.5m)	Panama	2015	300 kHz	Avisoft USG 416H, CM-16 microphone	10 (4)	GG
Desmodus rotundus	Flight room (3.7m * 2.5m)	Panama	2015	300 kHz	Avisoft USG 416H, CM-16 microphone	13 (7)	GG
Glossophaga soricina	Flight room (3.7m * 2.5m)	Panama	2015	300 kHz	Avisoft USG 416H, CM-16 microphone	8 (3)	GG
Lonchorhina aurita	Flight room (5.0m*5.0m)	Panama	2019	300 kHz	Avisoft USG 416H, CM-16 microphone	17 (9)	GG
Phyllostomus hastatus	Flight room (3.7m * 2.5m)	Panama	2015	300 kHz	Avisoft USG 416H, CM-16 microphone	14 (4)	GG
Thyroptera tricolour	Flight room (1.4m * 1.0m * 0.8 m)	Panama	2010	666.6 kHz, 16-bit;	Avisoft USG 116Hm, CM-16 microphone	19 (6)	IG

Myotis nigricans	Free flying in foraging habitat	Peru Panama	2012 2019	250 kHz, 16-bit depth	Avisoft USG 116Hm, CM-16 microphone	10 (7)	MK

GG: Gloria Gessinger, IG: Inga Geipel; KJ: Kirsten Jung; MK: Mirjam Knörnschild

2. Natural Stimuli

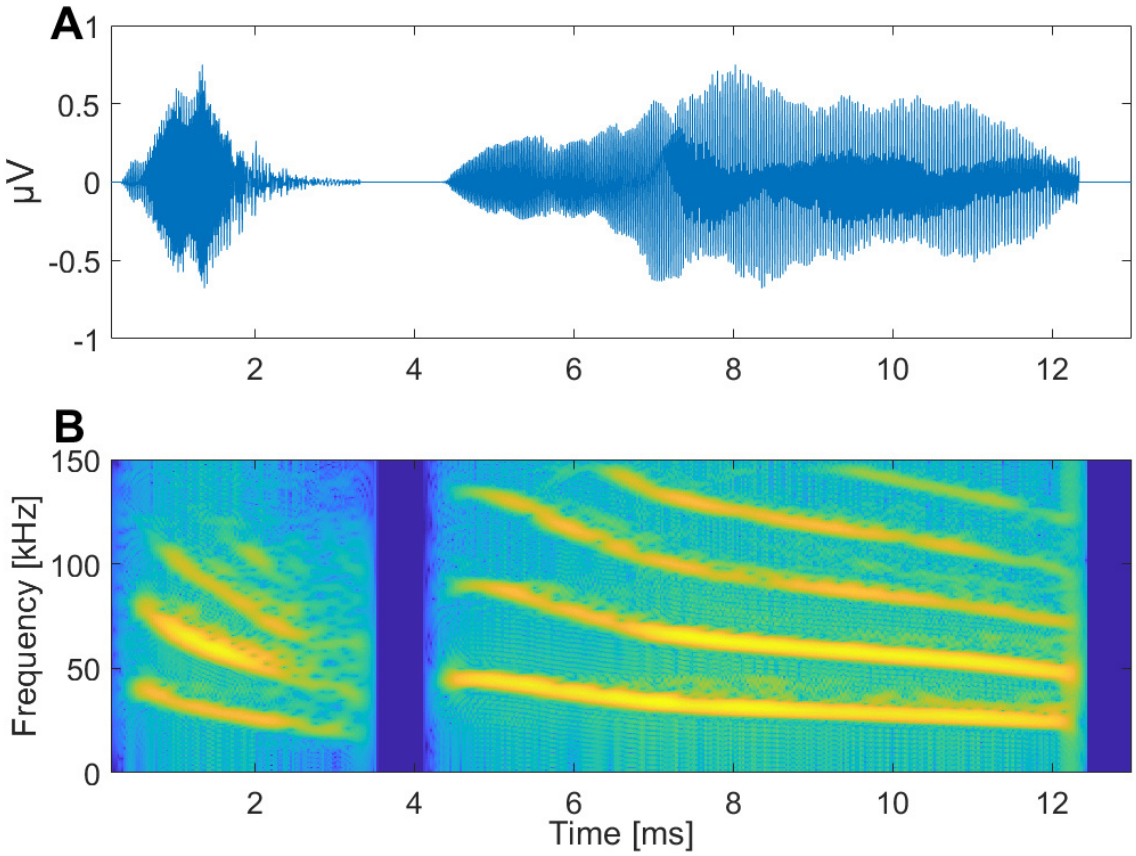


Figure 24: Short and Long Natural Stimuli.
(A) Oscillograms of an exemplary 2 ms long natural stimulus followed by an exemplary 8 ms long natural stimulus as used in Experiment 2. (B) Corresponding spectrograms of the natural stimului in (A): an exemplary 2 ms long natural stimulus followed by an exemplary 8 ms long natural stimulus as used in Experiment 2.

3. Detailed results of Experiment 1

Table 13: MMS Integrated Response Strength for 6, 40 and 100 Hz for All Species.

Species	6 Hz [µV]	40 Hz [µV]	100 Hz [µV]
Rhynchonycteris naso	93.37	112.94	63.05
Saccopteryx bilineata	236.26	237.22	208.50
Saccopteryx leptura	278.91	277.60	247.96
Molossus molossus	175.37	144.70	120.28
Pteronotus parnellii	72.31	69.04	49.43
Carollia perspicillata	159.80	137.28	123.24
Desmodus rotundus	164.70	157.64	139.70
Glossophaga soricina	179.37	154.57	145.17
Lonchorhina aurita	668.10	574.45	429.61
Phyllostomus hastatus	95.41	82.92	71.28
Thyroptera tricolour	159.96	113.79	93.09
Myotis nigricans	256.60	187.16	135.50

4. Detailed results of rhythm analysis

Table 14: Echolocation Recordings, Rhythm Analysis.
Average IOI duration and corresponding IOI beats for all species in analysed search flight call sequences

Species	Mean IOI [ms] ± SD	Coefficient of variation (calculated per sequence, averaged for all sequences)	IOI beat frequency range [Hz]
Rhynchonycteris naso	49.15 ± 11.3	0.17	15.1 – 24.2
Saccopteryx bilineata	119.1 ± 51.0	0.29	3.7 – 13
Saccopteryx leptura	55.8 ± 25.3	0.35	13.4 – 32.4
Molossus molossus	121.5 ± 43.3	0.26	5.2 – 12.5
Pteronotus parnellii	55.2 ± 29.8	0.40	10.4 – 27
Carollia perspicillata	62.5 ± 88.2	0.79	9.8 – 43.2
Desmodus rotundus	83.0 ± 68.0	0.63	7.4 – 30.5
Glossophaga soricina	31.6 ± 18.0	0.56	21.9 – 38.9
Lonchorhina aurita	85.7 ± 86	0.51	4.8 – 19.0
Phyllostomus hastatus	61.7 ± 45.7	0.54	10.4 – 22.5
Thyroptera tricolour	9.3 ± 4.9	0.16	86 – 144.4
Myotis nigricans	82.5 ± 25.5	0.26	10.2 – 18.2

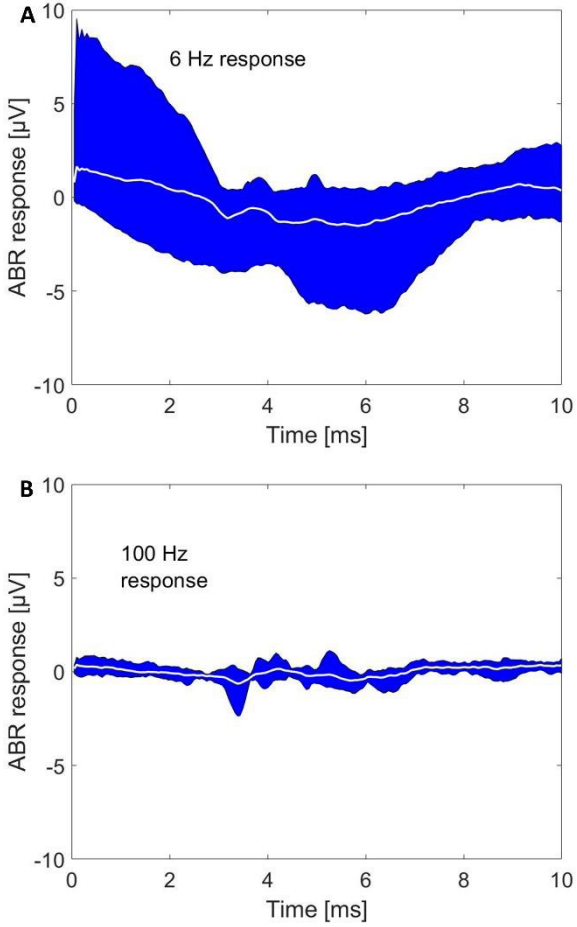
5. Raw ABR Slopes of all analysed species – Experiment 1

Figure Explanation:

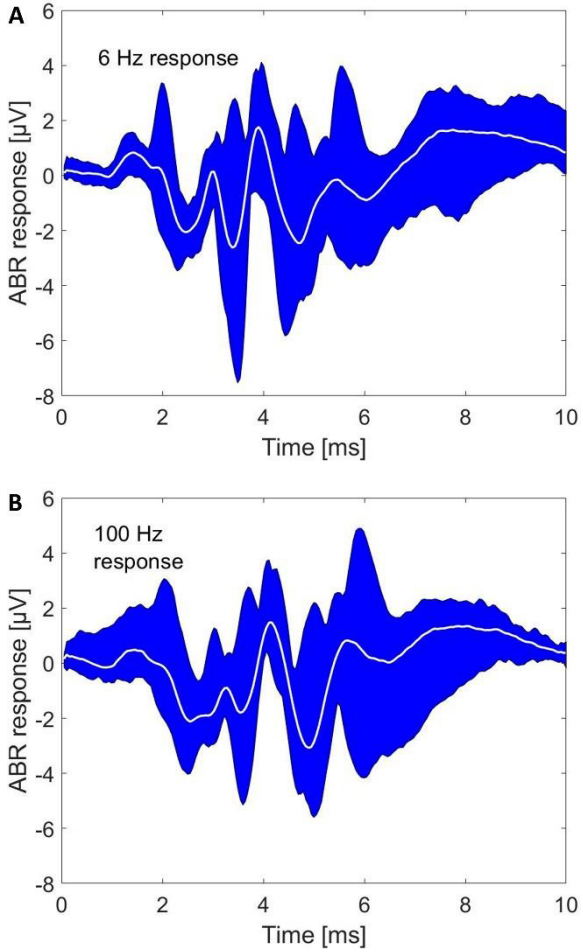
Figure 25-36: Raw ABR Slopes of All Analysed Species – Experiment 1.
(A) ABR trace at a stimulus presentation rate of 6 Hz, shown is the mean of all individuals of that species (white line) and the minimum and maximum of all individuals (blue area).

(B) ABR trace at a stimulus presentation rate of 100 Hz, shown is the mean of all individuals of that species (white line) and the minimum and maximum of all individuals (blue area).

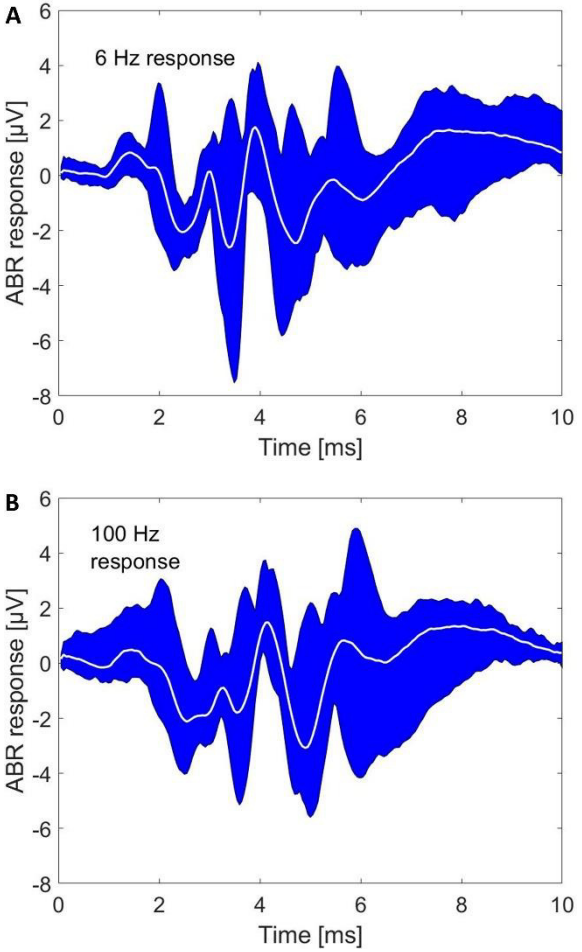
a. Figure 25: ***Rhynchonycteris naso***



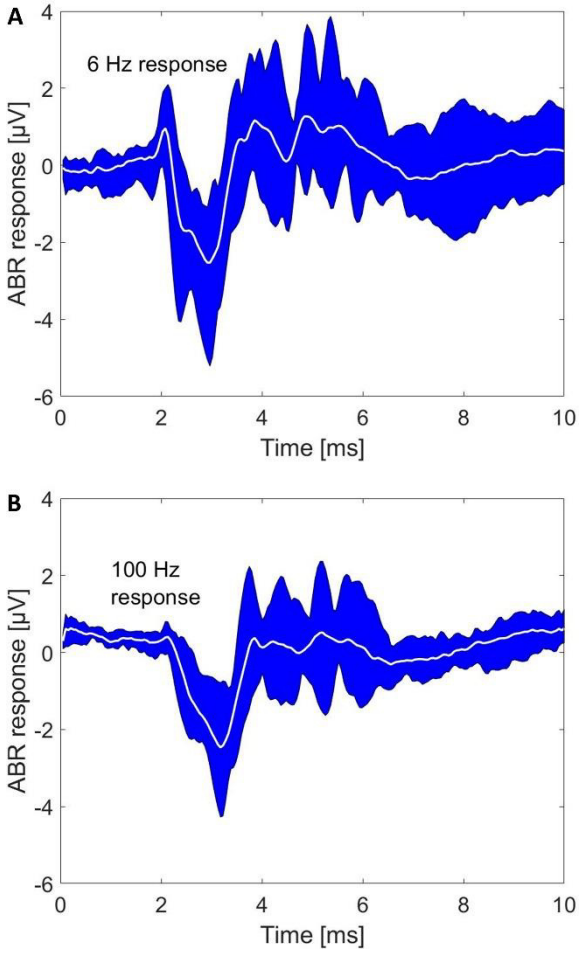
b. Figure 26: ***Saccopteryx bilineata***



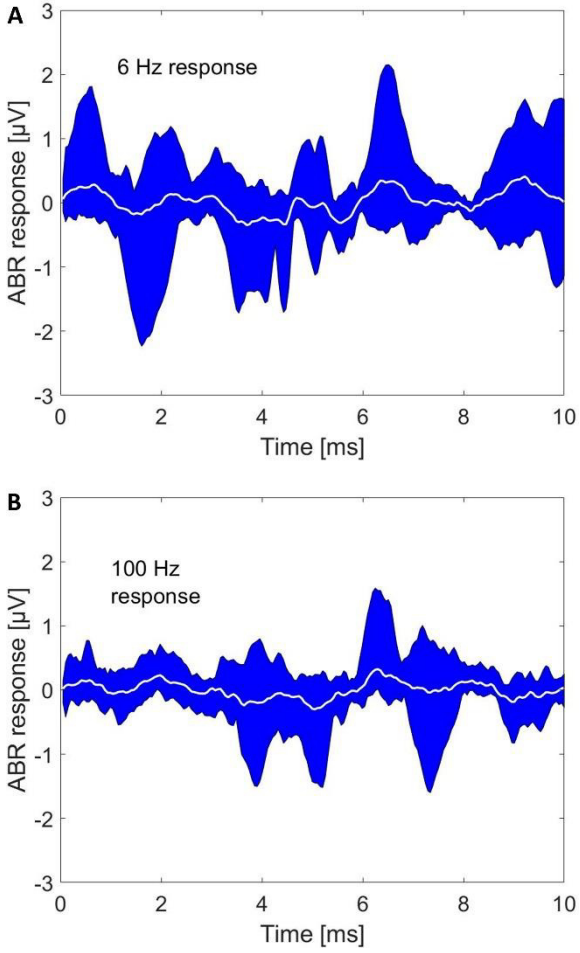
c. Figure 27: ***Saccopteryx leptura***



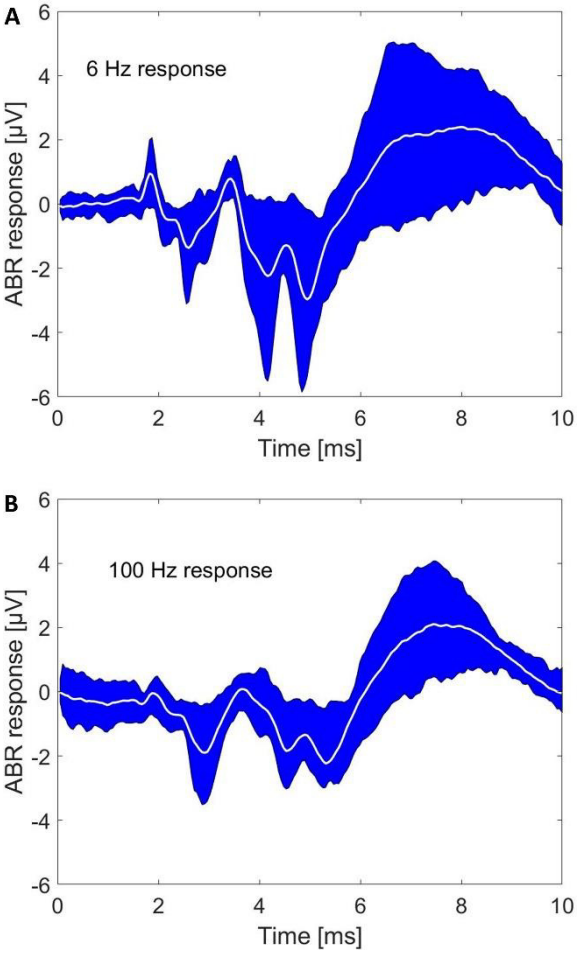
d. Figure 28: ***Molossus molossus***



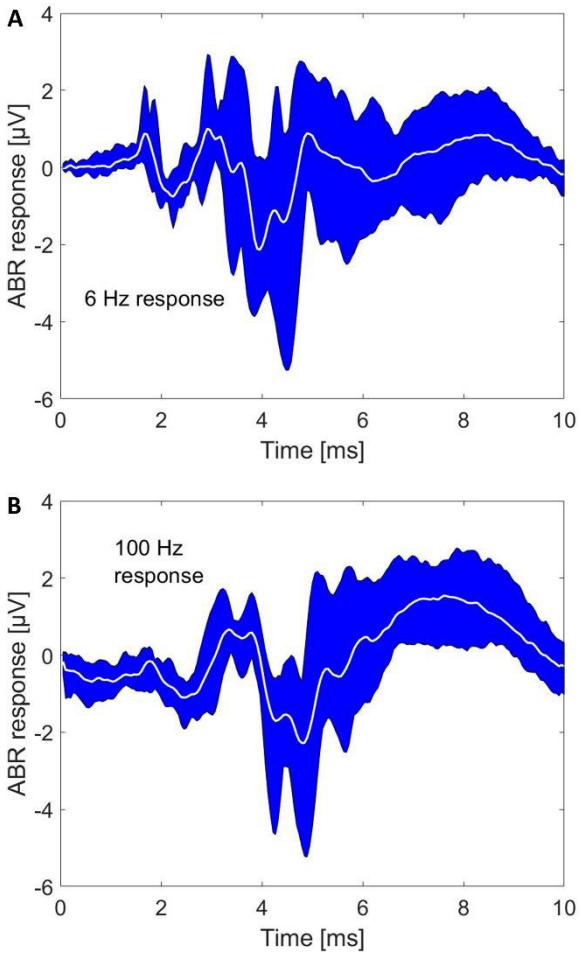
e. Figure 29: ***Pteronotus parnellii***



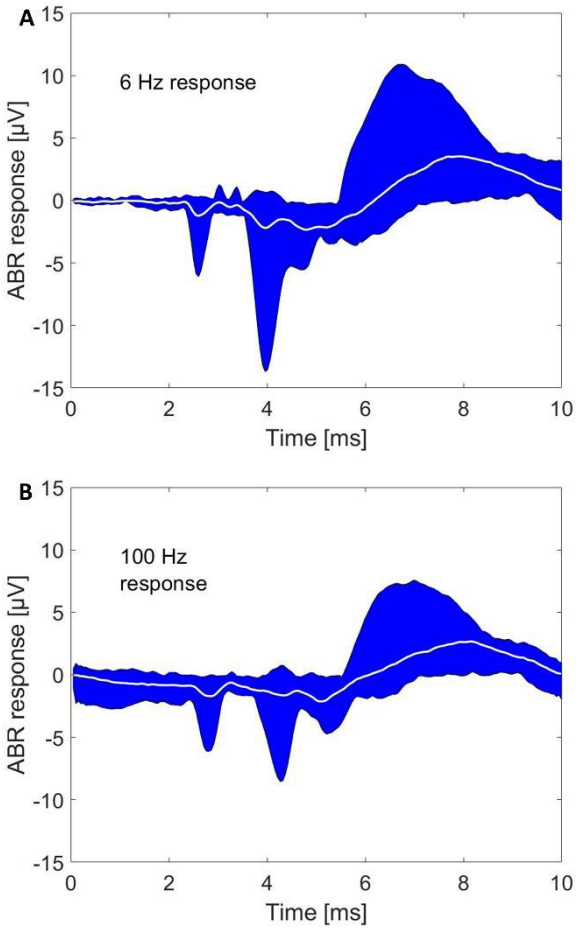
f. Figure 30: ***Carollia perspicillata***



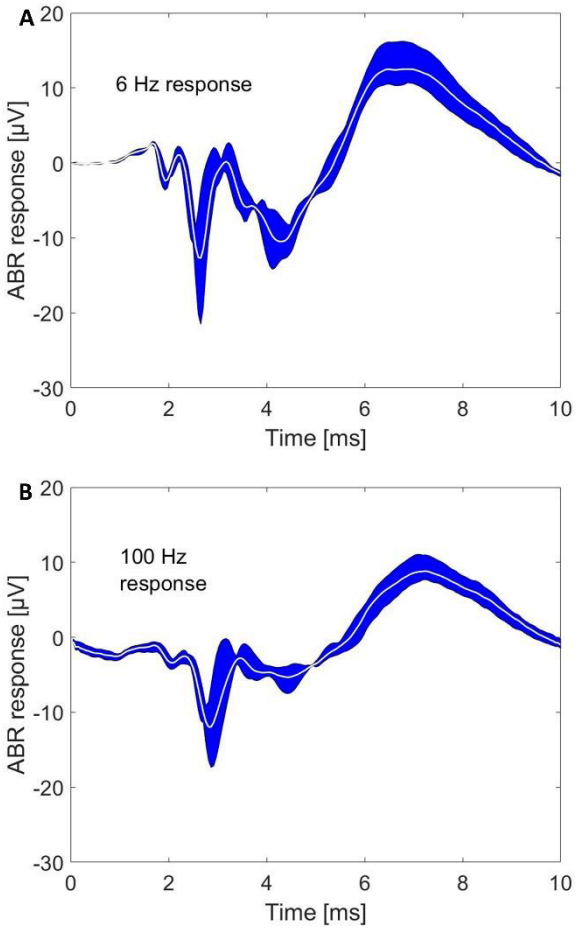
g. *Figure 31:* ***Desmodus rotundus***



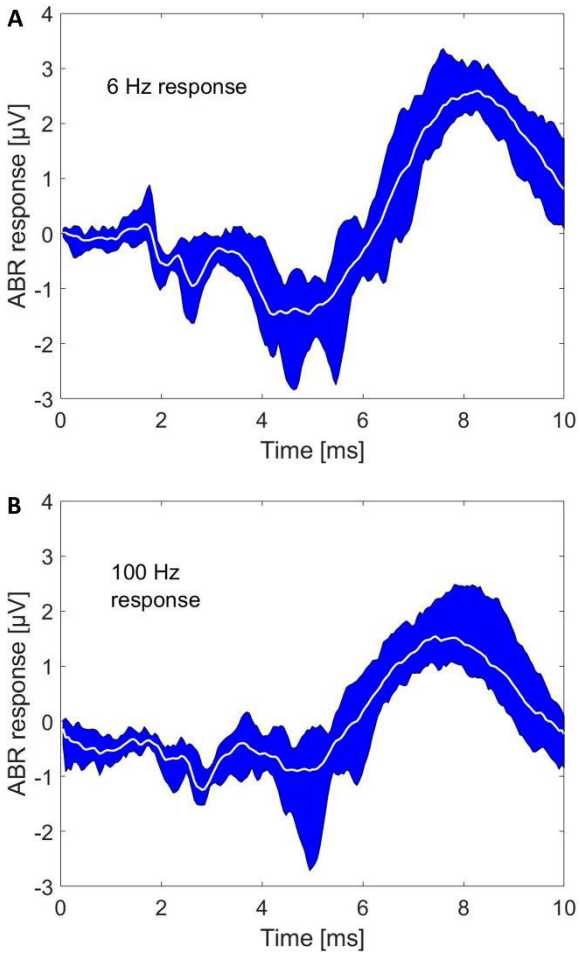
h. *Figure 32:* ***Glossophaga soricina***



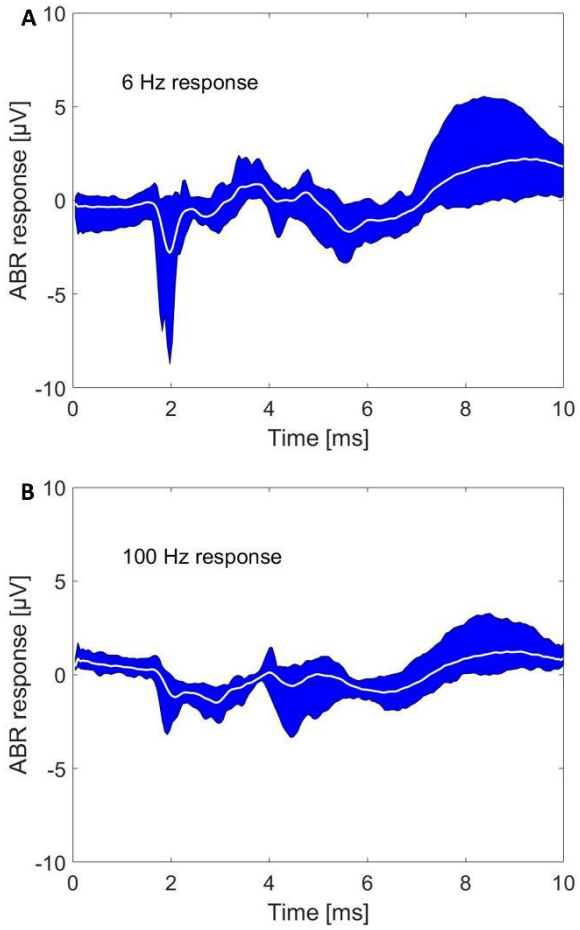
i. *Figure 33:* ***Lonchorhina aurita***



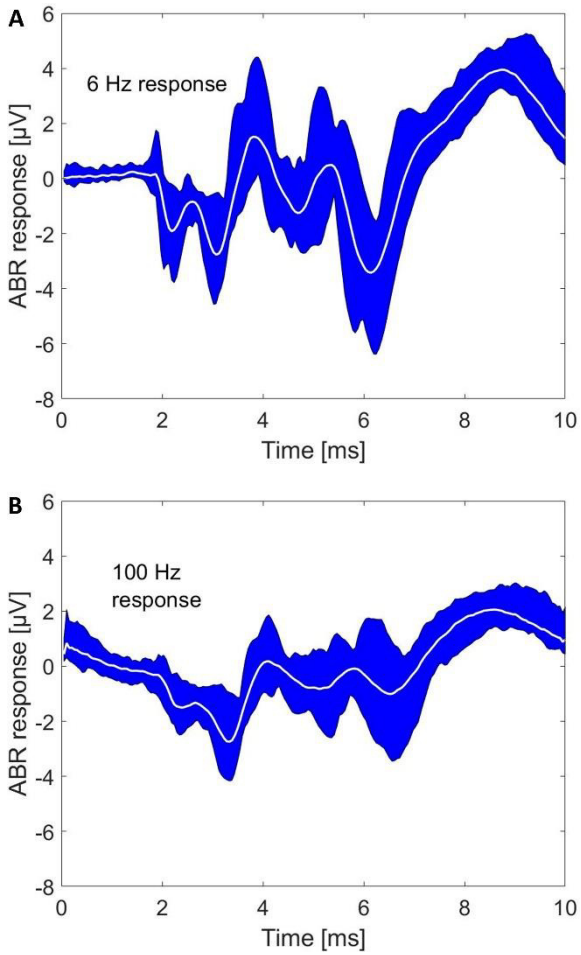
j. *Figure 34:* ***Phyllostomus hastatus***



k. *Figure 35:* ***Thyroptera tricolour***



l. *Figure 36:* ***Myotis nigricans***



6. Detailed Results of all analysed species – Experiment 1

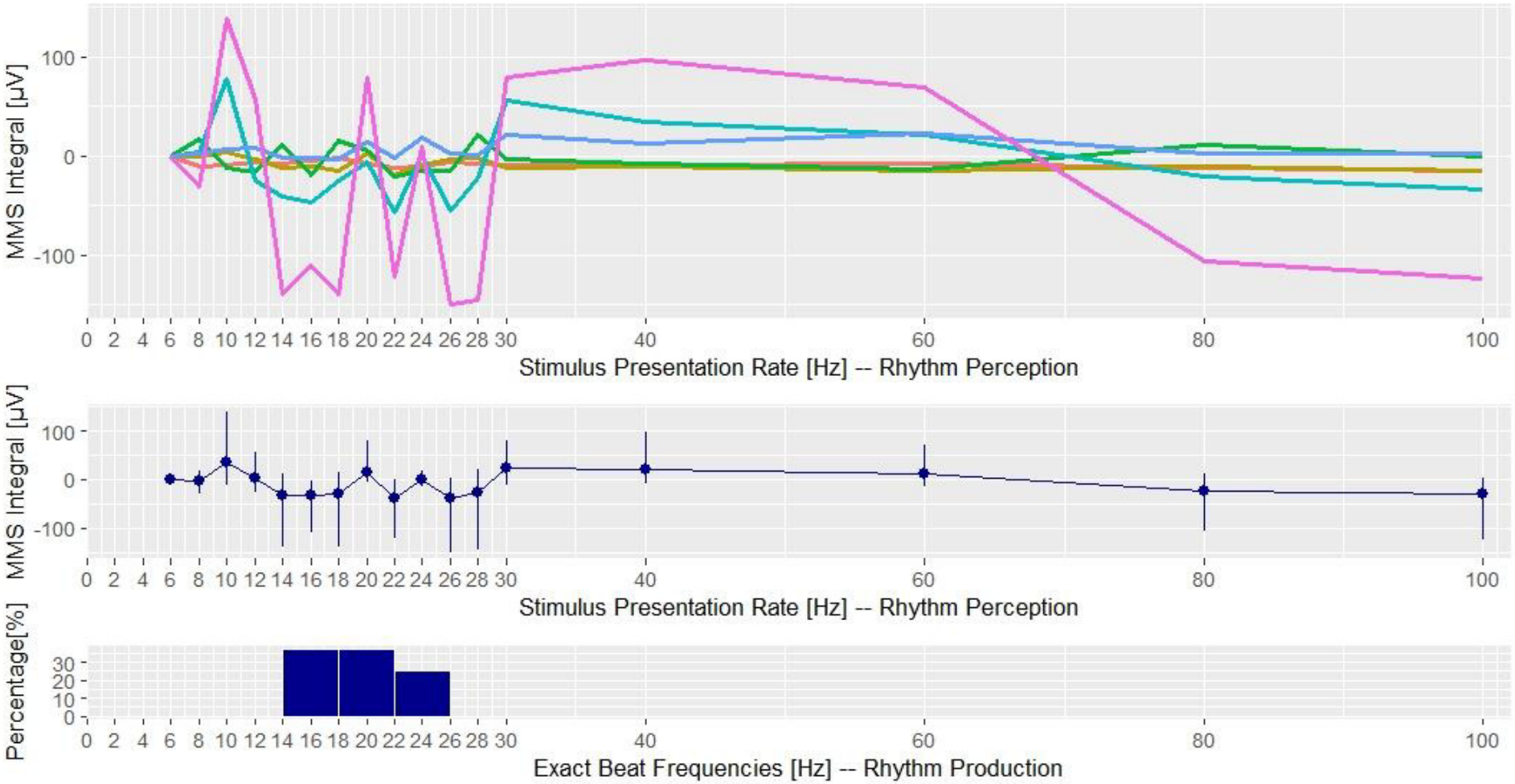
Figure Explanation:

Figure 37-48: Detailed Results of All Analysed Species – Experiment 1
(A) Integrated ABR response in µV per stimulus presentation rate. Every line represents one individual. Differences between individuals are visible in all species but are more pronounced in some. The general trend of a decrease in response strength between 6 Hz and 100 Hz is true for most individuals.

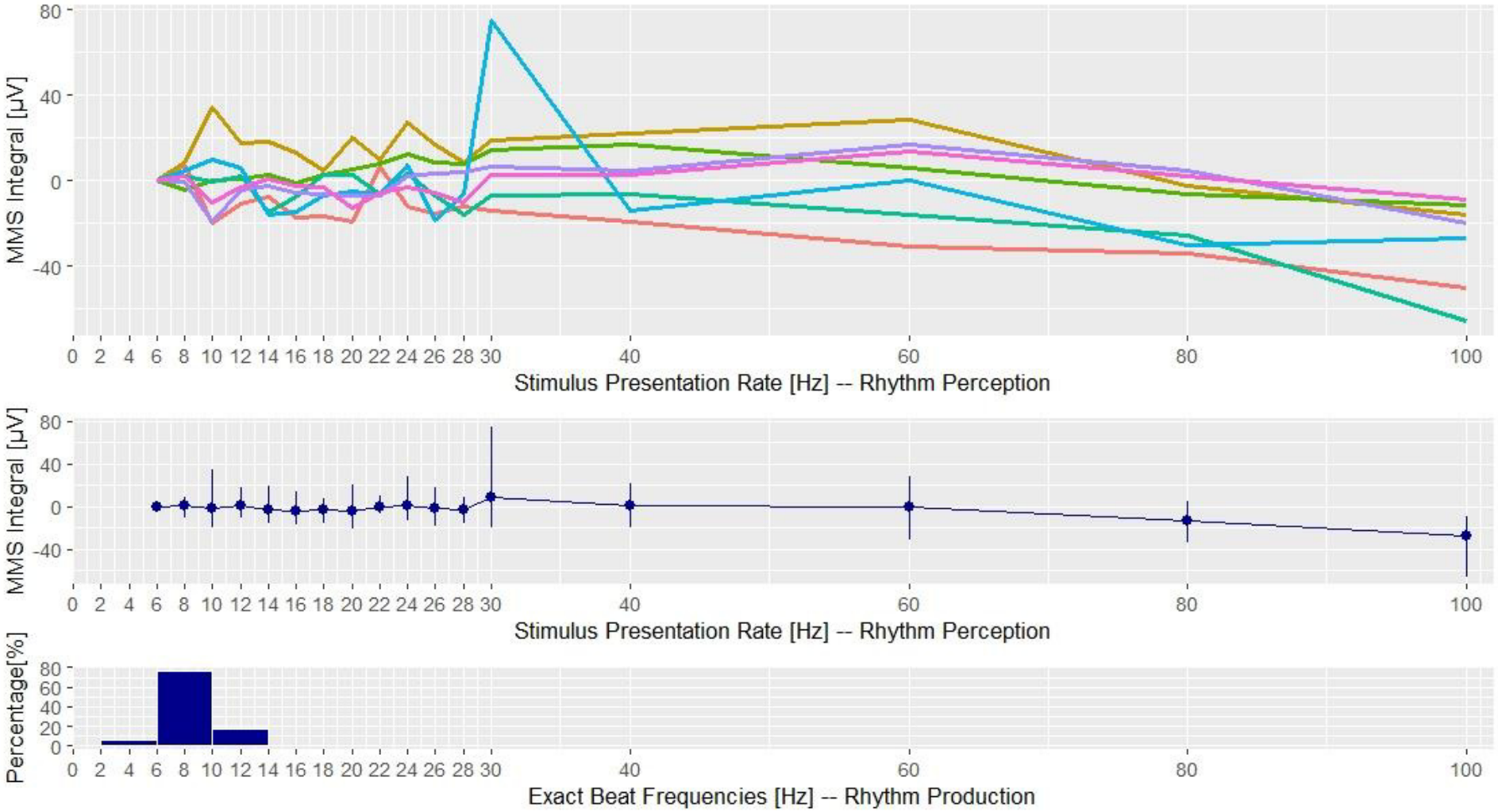
(B) The averaged integrated ABR response per species for all tested stimulus presentation rates. Shown is the average (black point) as well as the maximum and minimum as a blue line.

(C) The corresponding production rhythms of echolocation calls of the respective species. Shown is the distribution of production rhythms in Hz as calculated with the IOI approach.

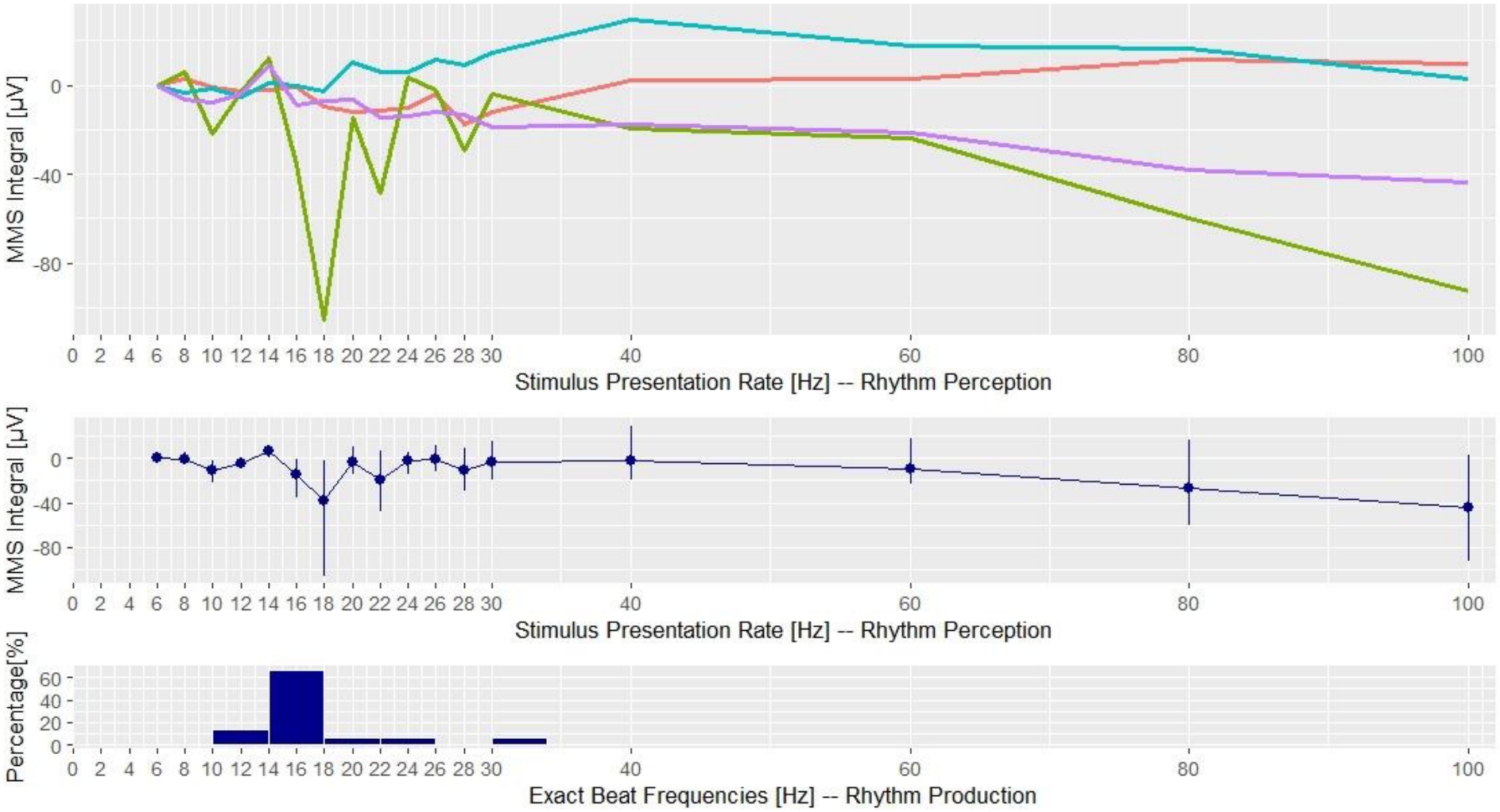
a. Figure 37: ***Rhynchonycteris naso***



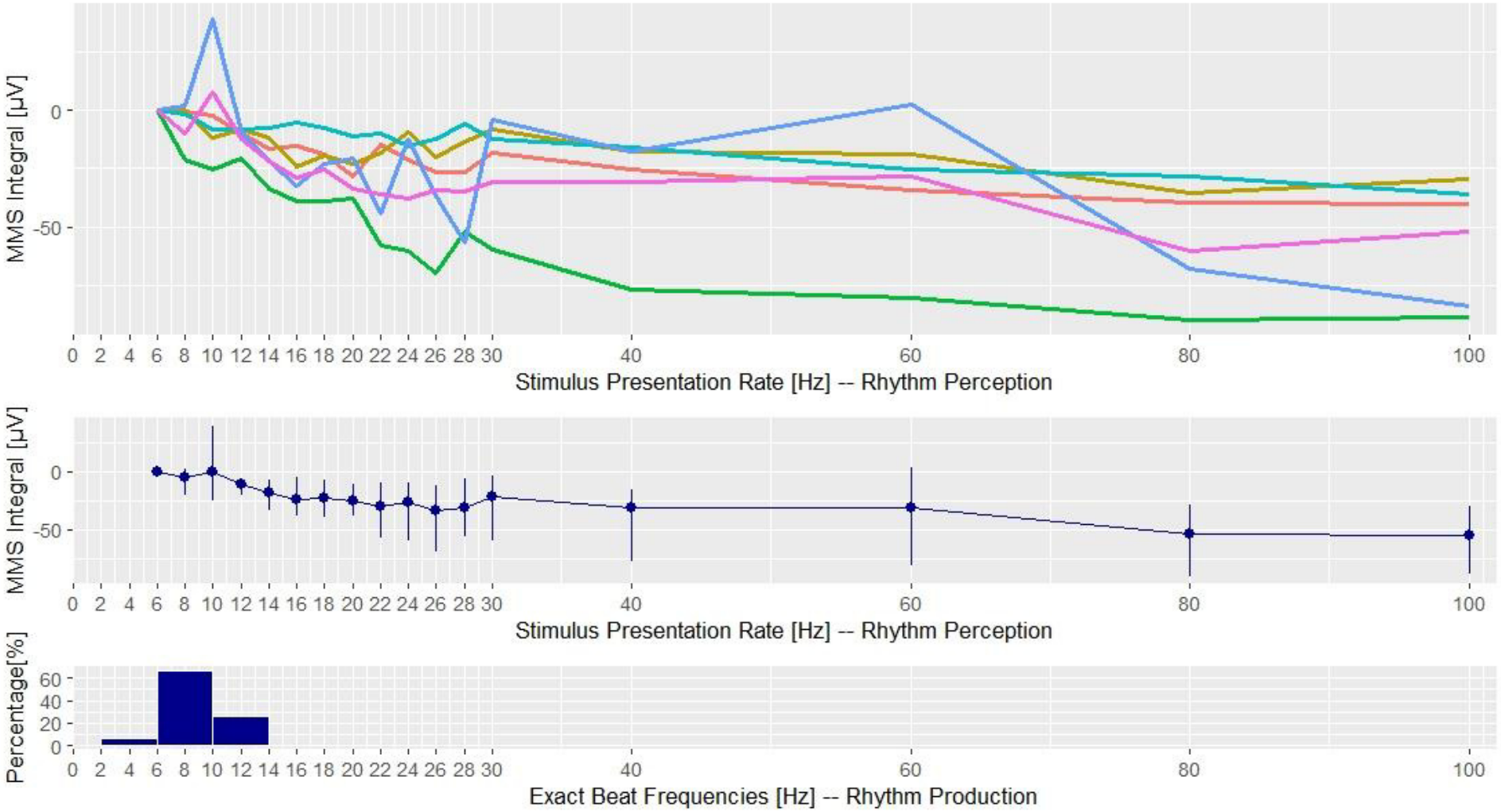
b. Figure 38: ***Saccopteryx bilineata***



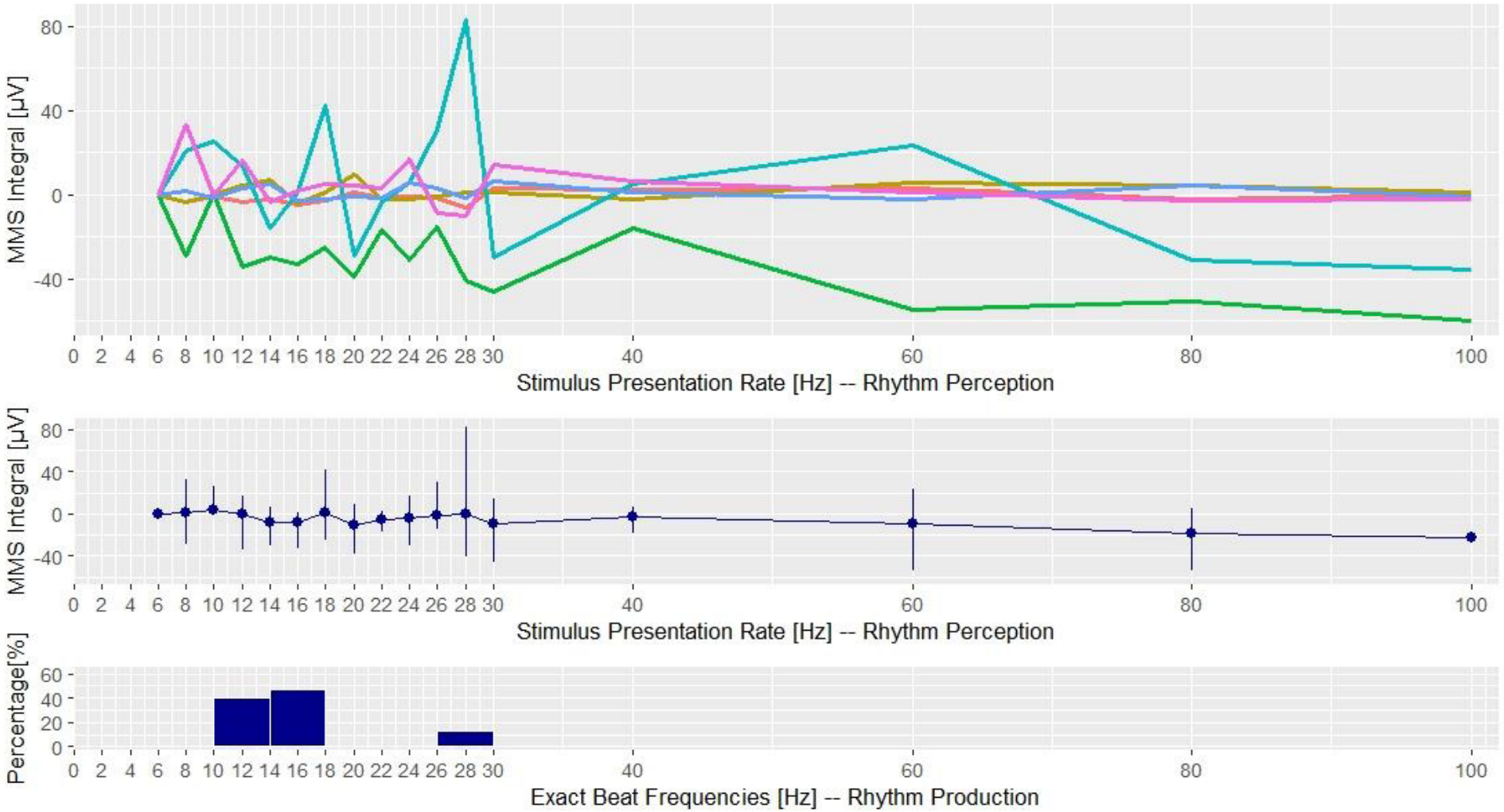
c. Figure 39: ***Saccopteryx leptura***



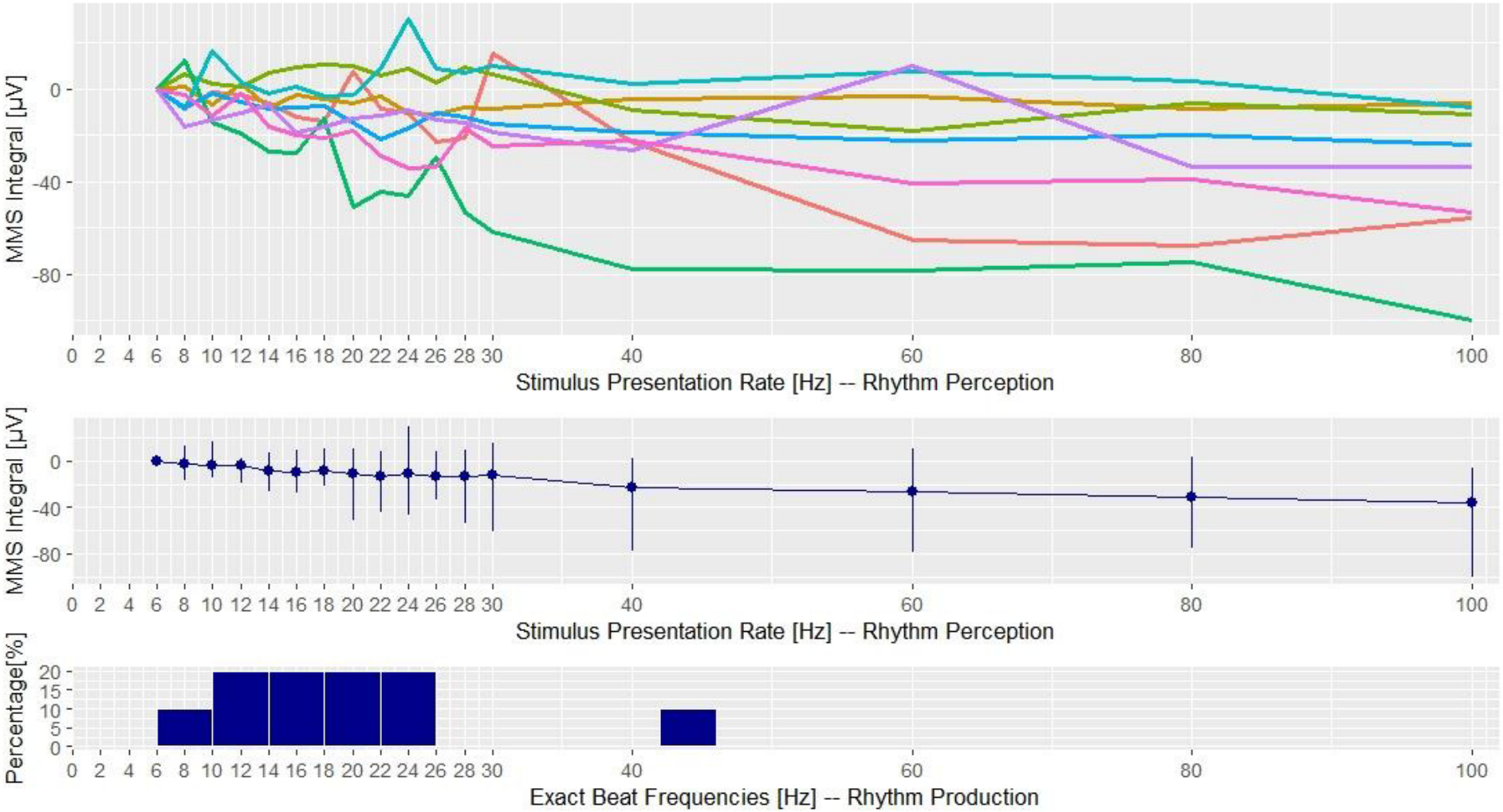
d. Figure 40: ***Molossus molossus***



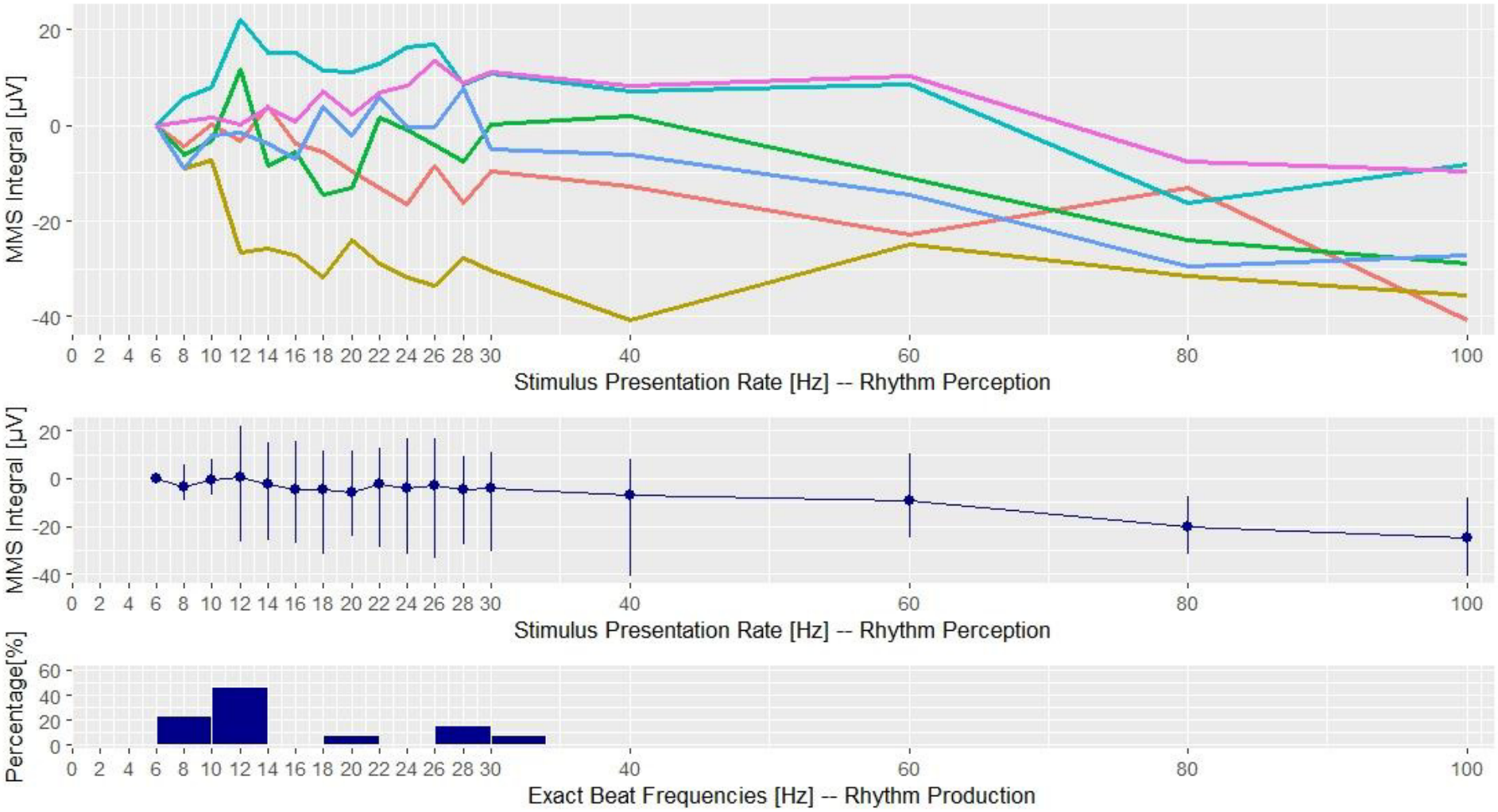
e. Figure 41: ***Pteronotus parnellii***



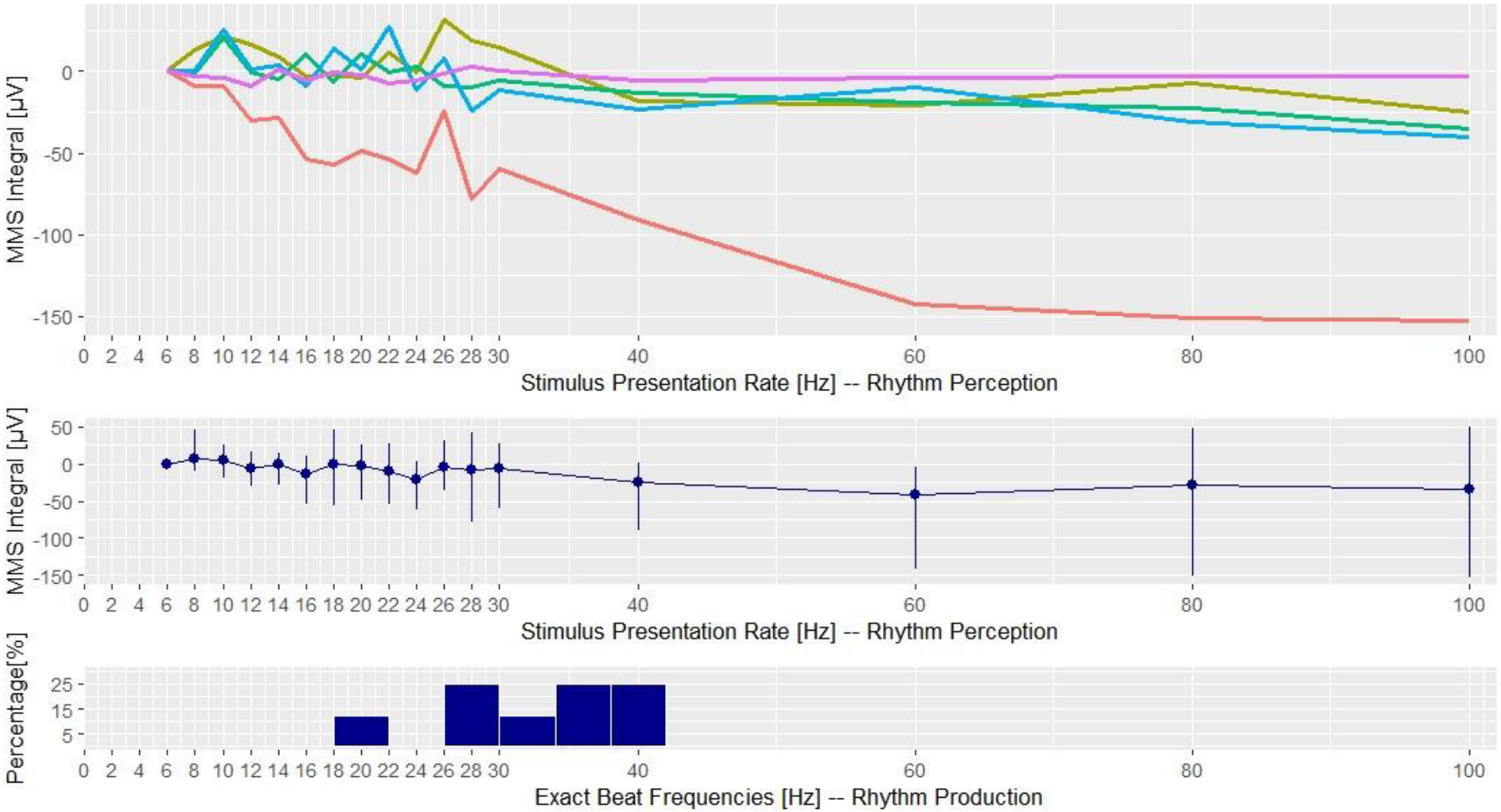
f. Figure 42: ***Carollia perspicillata***



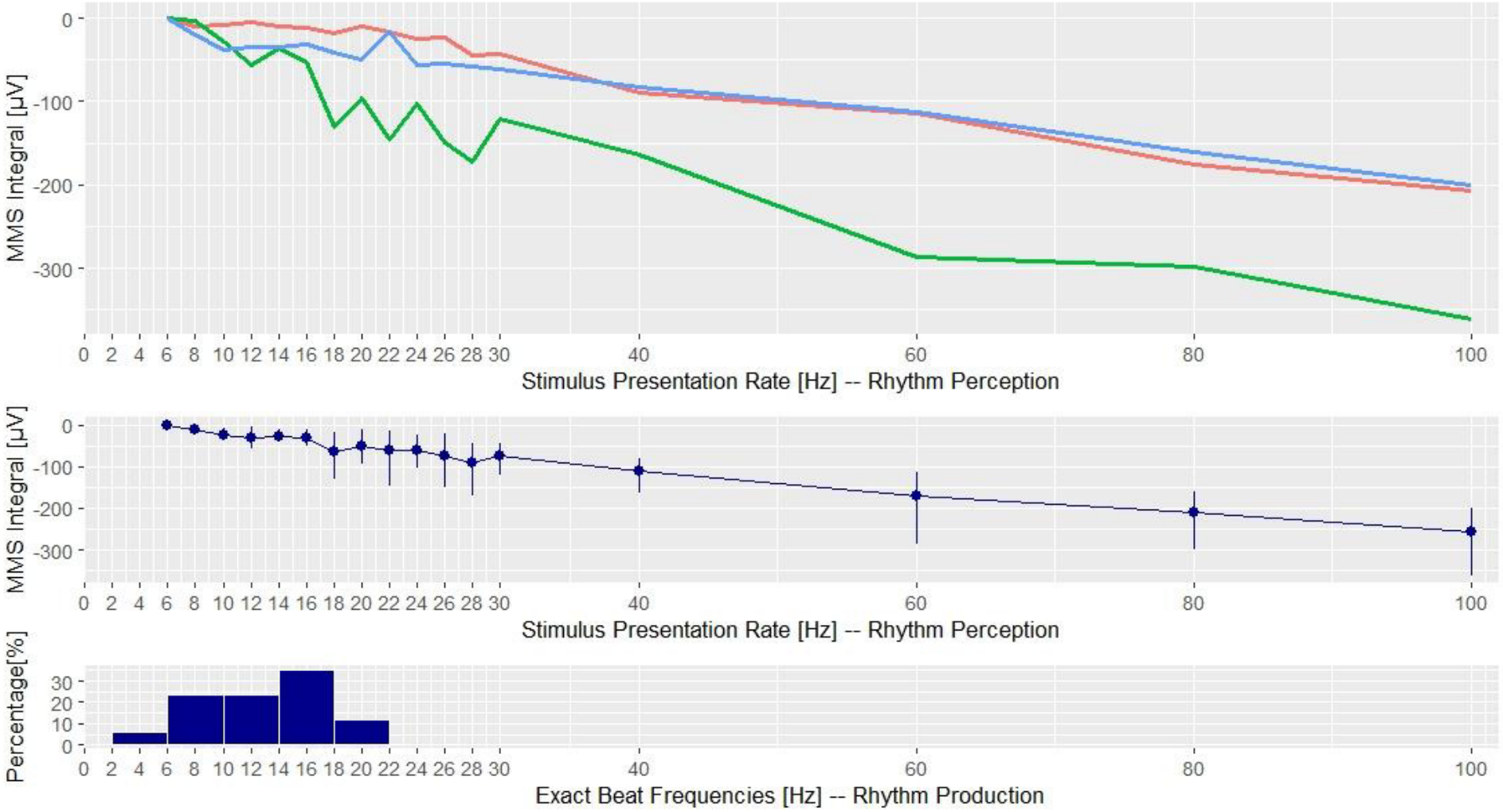
g. Figure 43: ***Desmodus rotundus***



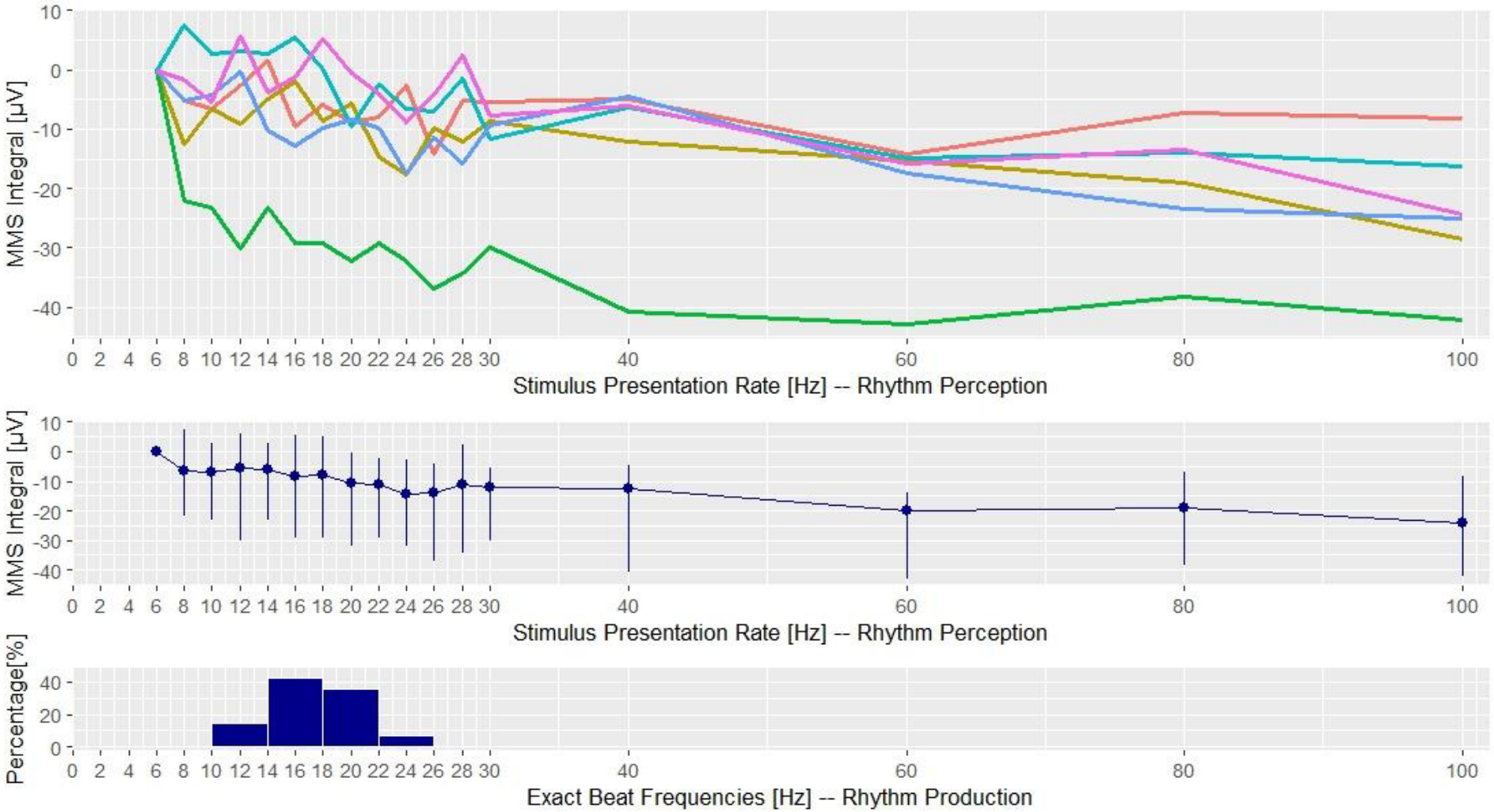
h. Figure 44: ***Glossophaga soricina***



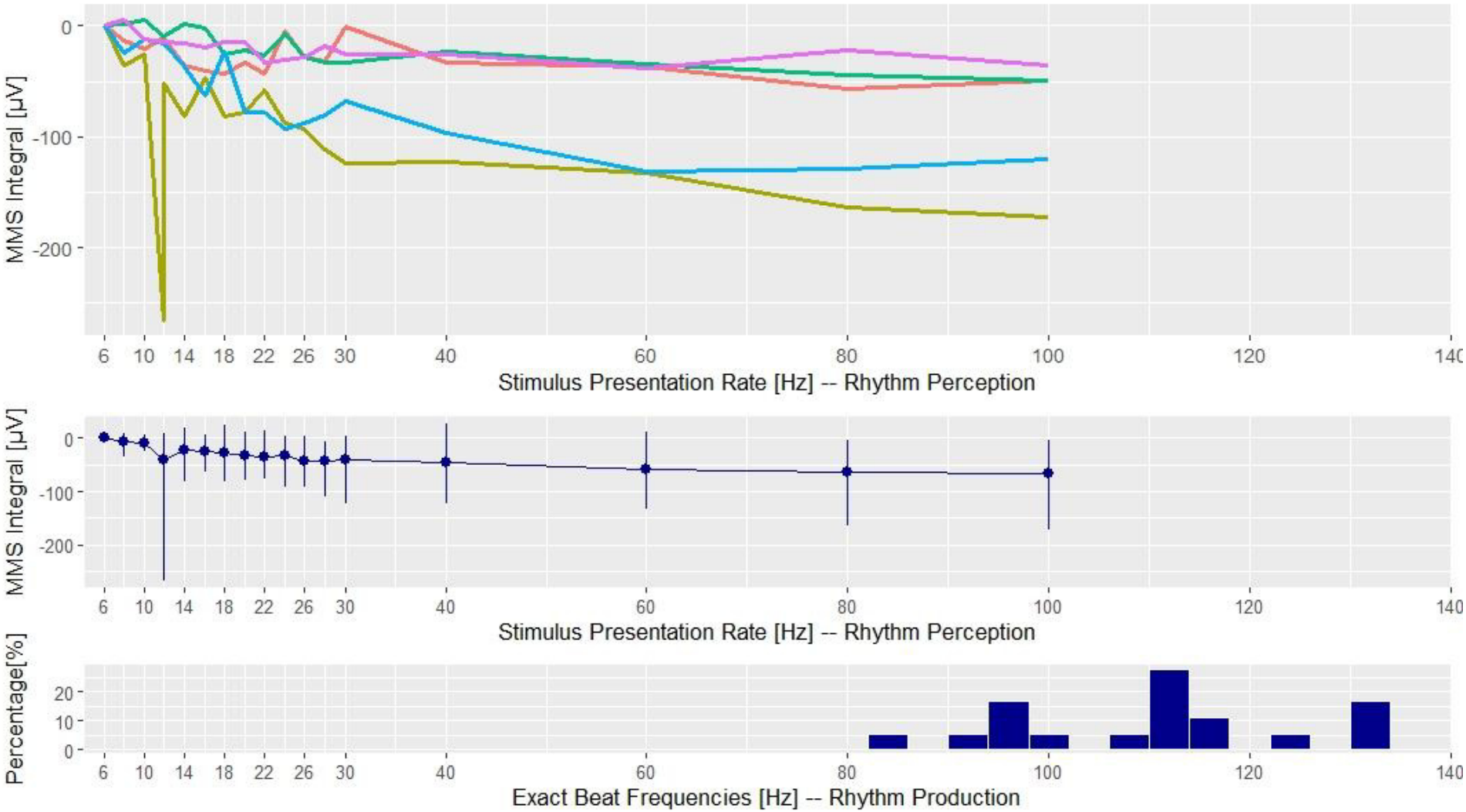
i. Figure 45: ***Lonchorhina aurita***



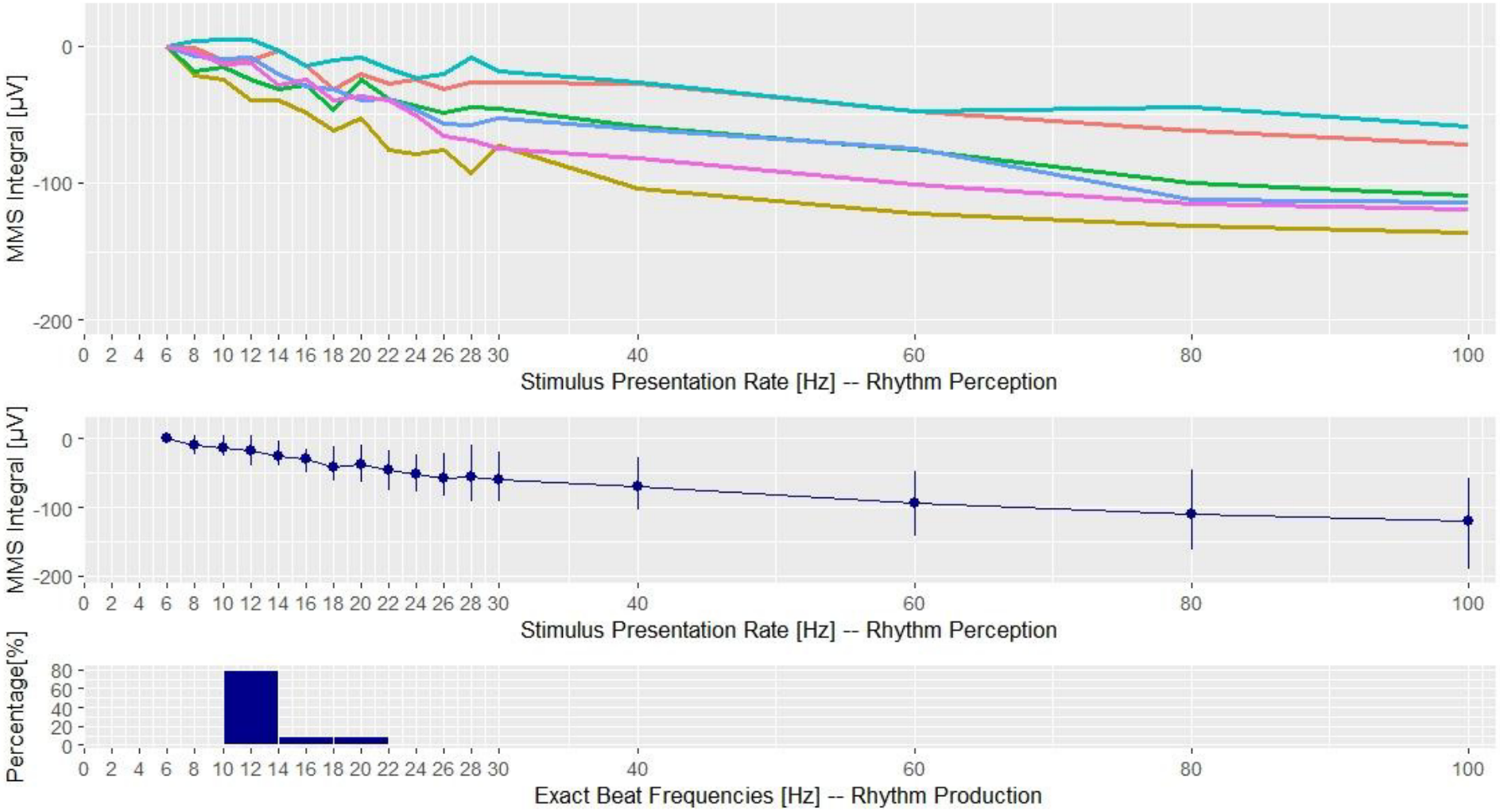
j. Figure 46: ***Phyllostomus hastatus***



k. Figure 47: ***Thyroptera tricolour***



l. Figure 48: ***Myotis nigricans***



Discussion

Rhynchonycteris naso

The very inconclusive results of *R. naso* could well be connected to an inappropriate decision regarding the frequency artificial stimuli were presented in, respectively the chosen sound pressure level. We chose a center frequency of 100 kHz (Jung, Kalko, & Helversen, 2007), which was presented at 90 dB SPL. Audiograms of that species show that this sound pressure might be too low for some individuals of *R. naso* to be able to perceive the stimuli at all, as the hearing threshold at a frequency of 100 kHz lies around 86 dB SPL on average. These are extrapolated values as pitch perception was measured for 87 kHz and 120 kHz (Lattenkamp et al., 2021). This could, of course, have a strong influence on the results, even though all measurements were significantly different from noise.

Thyroptera tricolour

A similar scenario could be assumed for *T. tricolour*. With peak frequencies of 140 kHz, this species calls at a higher pitch than the other studied species. The loudspeaker in our setup could only play frequencies of up to 120 kHz, which is why *T. tricolour* was presented with artificial stimuli at 100 kHz instead of their peak echolocation frequency of 140 kHz. In contrast to the other species *T. tricolour* was therefore not presented with stimuli in their best hearing range, and even though they should be able to hear the presented stimuli (Lattenkamp et al., 2021) we cannot exclude that this influenced the results.

References

Jung, K., Kalko, E. K. V., & Helversen, O. (2007). Echolocation calls in Central American emballonurid bats: signal design and call frequency alternation. *Journal of Zoology*(272), 125-137. https://doi.org/10.1111/J.1469-7998.2006.00250.X

Lattenkamp, E. Z., Nagy, M., Drexl, M., Vernes, S. C., Wiegrebe, L., & Knornschild, M. (2021). Hearing sensitivity and amplitude coding in bats are differentially shaped by echolocation calls and social calls. *Proceedings of the Royal Society B, 288*(1942), 20202600. https://doi.org/10.1098/rspb.2020.2600

Chapter IV

Novel Ideas to Further Expand the Applicability of Rhythm Analysis.

Abstract

The temporal structure of animals' acoustic signals can inform about the context, urgency, species, individual identity, or even geographical origin. We present three independent ideas to further expand the applicability of rhythm analysis. A description of a rhythm or beat is only as good as the knowledge about its' goodness-of-fit, meaning how well the rhythms describe a sequence. Existing goodness-of-fit values are not comparable between methods of rhythm analysis and not easily comparable between datasets. Furthermore, they are strongly correlated to certain parameters of the described sequence, e.g., the number of elements in the sequence. Here, we introduce a new, universal goodness-of-fit value, comparable across methods and datasets, which illustrates how well a certain beat frequency in Hz describes the temporal structure of a sequence of elements. This is the first goodness-of-fit value capable of giving the information per element, instead of only per sequence. This value can be used for the following methods: a generate-and-test approach, Fourier analysis, and the simple calculation of call rates or analysis of Inter-Onset-Intervals. We then describe two additional approaches to adapt already existing methods for analysing the rhythm of acoustic sequences of animals. The new additions, a slightly modified way to use the already established Fourier analysis and concrete examples on how to use the visualization with so-called recurrence plots, enable the analysis of more variable data, while also giving more details then previously. As most communication signals are quite variable and the structures/levels that are important for their perception are often not known, it is crucial to have methods that can analyse all possible structures and levels.

Introduction

In recent years, the temporal structure or rhythm of animal's acoustic signals has received increasing attention. Much emphasis lays on the development of methods to assess and quantify underlying temporal patterns (Burchardt & Knörnschild, 2020; Burchardt et al., 2019; Norton & Scharff, 2016; Ravignani & Norton, 2017; Saar & Mitra, 2008). The rhythm of an isochronous element sequence is termed a 'beat frequency' and is given in Hz. So far, three methods have been proposed for extracting exact beat frequencies in order to describe an isochronous element sequence; 1) Fourier analysis, which decomposes a signal into its sinus components (Burchardt & Knörnschild, 2020; Saar & Mitra, 2008),

2) Generate-And-Test approach (GAT), where a series of acoustic signals are overlaid with an artificial beat to test which artificial beat frequencies resembles the series best (Norton & Scharff, 2016; Ravignani & Norton, 2017), and 3) Inter-Onset-Intervals (IOI), which allows the calculation of beat frequencies by averaging IOIs and transforming this rate into a frequency (Burchardt & Knörnschild, 2020).

Until now, studies on temporal structure or rhythm of animal's acoustic signals have often focused on quite simple sequences with an underlying isochronous structure (e.g., only one element type, visually uniform temporal structures, or short sequences (Burchardt & Knörnschild, 2020; Ravignani, 2018)). Such a structure resembles a metronome sound, with constant beat and gap lengths. The above-mentioned methods, GAT and Fourier analysis, together with the commonly used calculation of rates or frequency-transformed rates (in Hz as in beats per second) describe these isochronous sequences well. However, for sequences containing various element types, sub-units and a strong variability between element duration and/or gap durations, such as skylark song (Briefer, Osiejuk, et al., 2010), nightingale song (Hultsch & Todt, 1981), or whale song (Payne & McVay, 1971), the interpretation of results of exact beat frequency calculations as described above becomes more difficult. Arising problems include that all methods will always give a "best-fitting" beat frequency. In case an isochronous beat is not suitable to describe the sequence, this beat frequency can therefore be very misleading. Also, interpretation for most analyses to determine isochrony are very clear for small values (i.e., nPVI or Coefficient of variation analysis, where low values are explicitly indicating low variability (Burchardt & Knörnschild, 2020; Ravignani & Norton, 2017)) but higher values are not as easily interpreted, as they could indicate a different rhythmic pattern than isochrony or indeed a random succession of elements. Analyses of the rhythm on such vocalizations require the refinement of established methods or the development of new ones, in order to allow a description of sequences in a meaningful and comparable way between species. Current problems related to existing methods are two-fold. The first issue, which is independent of the complexity of the structure, is the limitations with which so-called goodness-of-fit values quantifying how well a certain beat frequency describes an element sequence can be compared between species as well as between methods. These values exist for all three above-mentioned methods to extract exact, best-fitting beats (Burchardt & Knörnschild, 2020), but they inflict

three problems; a) they show complex correlations to, among other parameters, the number of elements in a sequence, b) values differ depending on the method used, which precludes any comparison between studies using different methods, and c) only one value can be obtained for the whole sequence that is being analysed, without any information at the element level. The second issue that becomes important regarding the analysis of more complex sequences is that, so far, existing methods provide only one best-fitting beat frequency when, in fact, the sequence might be best described by more than one beat frequency. Directly related, it might be interesting to look for sub-patterns and analyse different parts of a sequence separately, to be able to depict rhythm changes within a complex sequence. The next challenge thus becomes to know where or what these sub-patterns might be.

In this study, we propose three new ideas on how to extend the existing analyses options, as well as how to bypass certain limitations. First, we introduce a new universal goodness-of-fit value. Second, we suggest that reporting the ten most prominent beat frequencies in a sequence instead of only the best-fitting beat frequency in Fourier analysis, which implies the assumption that one beat frequency is enough, is essential to describe a complex temporal structure. Third, we encourage the use of recurrence plots to identify the sub-structures and sub-units that could be of interest for further analysis.

Introducing improvements to the existing methods

We performed a rhythm analysis on a dataset of flight songs of the skylark, *Alauda arvensis* (for details on recordings see (Briefer et al., 2008; Briefer, Rybak, et al., 2010)). The song produced by males of this species during the breeding season while in flight is very complex: each individual can combine more than 300 different syllables in its song, giving rise to a lot of variation (Aubin, 1982; Briefer et al., 2008; Briefer et al., 2008). The use of existing methods on such song, namely reporting only the one best-fitting beat frequency per sequence as calculated in Fourier analysis and the resulting goodness-of-fit values, proved to be insufficient for describing the rhythmic structure of this system. We therefore developed a goodness-of-fit value and revaluated how to best report results of the Fourier analysis, in order to further facilitate comparability between methods and species, for example, through enabling the description of both simple and more complex patterns with the same methods, but also by making the various existing methods themselves more comparable. We introduce a newly established universal

goodness-of-fit value, and we discuss additions to existing methods (Fourier analysis and recurrence plots), with the overall purpose to further advance rhythm analysis, its applicability, and comparability.

Our universal goodness-of-fit value is tested both on the skylark dataset and five already published datasets, where beat frequencies [in Hz] were calculated (Burchardt & Knörnschild, 2020; Burchardt et al., 2019); acoustic signals of the tropical American bat *Saccopteryx bilineata*: 500 isolation calls (Knörnschild, Nagy, et al., 2012), 142 territorial songs (Behr et al., 2006) and 33 echolocation call sequences (Knörnschild, Jung, et al., 2012); as well as 49 isolation calls of the tropical American bat *Carollia perspicillata* (Knörnschild et al., 2013) and 60 echolocation sequences of *Physeter macrocephalus* (Bøttcher et al., 2018; Tønnesen et al., 2018). For the revaluation of Fourier analysis results and recurrence plots, we focus on the dataset of skylark flight songs.

Universal Goodness-of-Fit Value

We propose a new, universal goodness-of-fit value that can be applied to any possible description of a temporal structure relying on frequencies; we term it *ugof* for "universal goodness-of-fit value". It is a value that is calculated for every element in a sequence and can then be summarized for a whole sequence or any other desired grouping (e.g., individuum, group, sequence type, etc.). A theoretical beat describes a sequence well when there are only small deviations between the original elements and the theoretically beats of the best-fitting beat frequency.

One element always lies between two theoretical beats. Therefore, the maximum deviation possible equals to half of the theoretical beat length since one will always search for the deviation to the next closest beat (Figure 49A). We can therefore describe a particular deviation as the ratio between the actual deviation to the next theoretical beat and the maximum deviation for the calculated best-fitting beat (Equation 1).

$$ugof = \frac{|\Delta|}{\Delta_{max}}$$

$ugof: universal\ goodness\ of fit\ value$ $\quad |\Delta|: absolut\ deviation\ to\ closest\ theoretical\ beat$

$\Delta_{max}: maximum\ possible\ deviation\ (half\ a\ beat\ duration)$

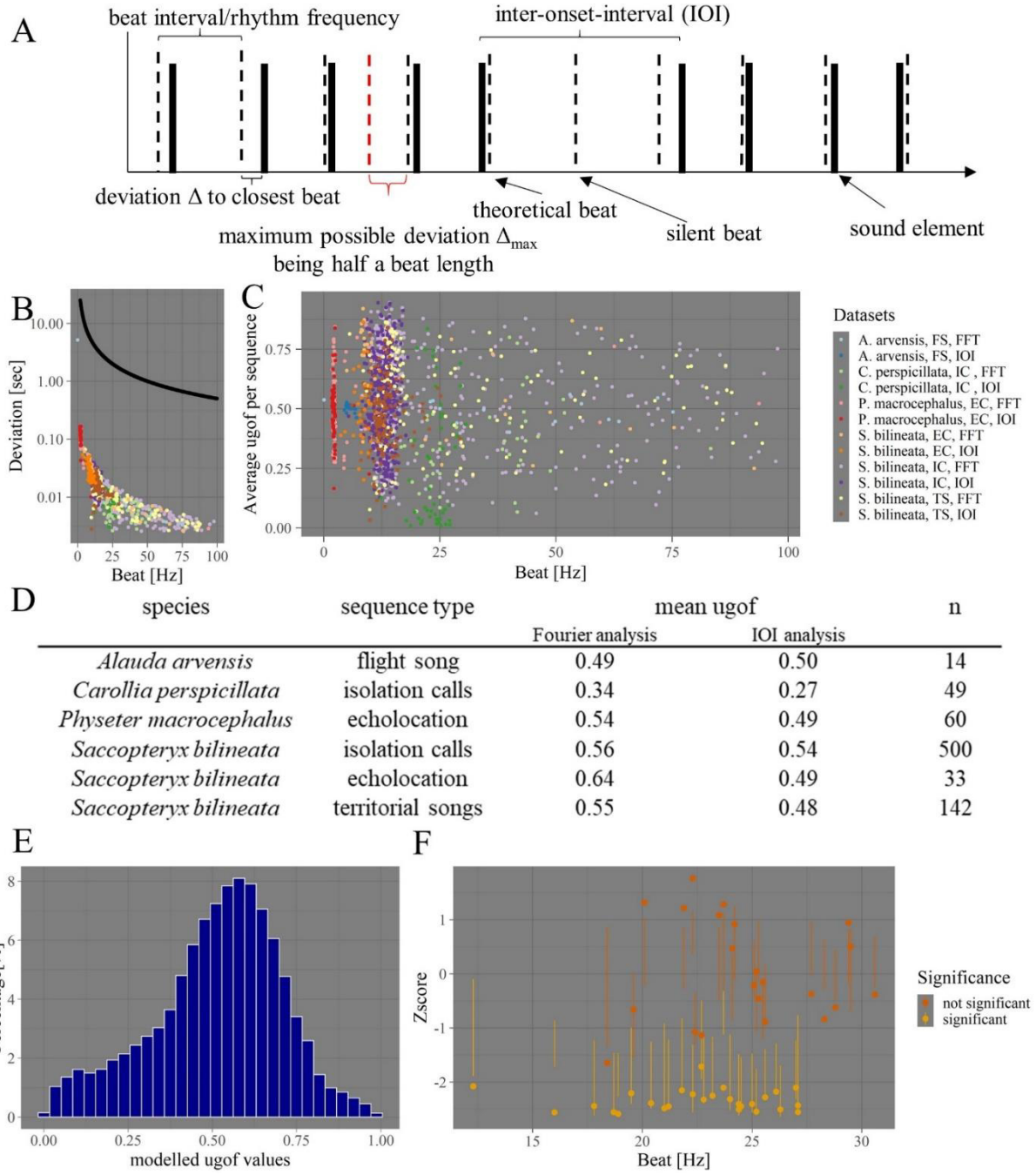
Both parameters, the maximum possible deviation to the next beat (Δ_{max}) and the actual deviation ($|\Delta|$), need to change in the same way depending on the corresponding frequency, for the method to be universally applicable. to be useful. This is depicted in Figure 49B, showing the theoretical maximum possible deviations (Δ_{max}) for beats of beat frequencies of up to 100 Hz (in black) and the actual deviations ($|\Delta|$) we measured for a total of 804 sequences and of two beat frequencies each (one as calculated with Fourier analysis, the second as calculated with the IOI approach), resulting in 1608 datapoints (in colour). The actual deviations are indeed much lower than the maximum possible deviations. By dividing the actual deviation ($|\Delta|$) by the maximum possible deviation (Δ_{max}) as shown in the equation, we get the ugof as a ratio and can be easily transformed into a percentage (by multiplying with 100) if required. The smaller *ugof* is, the closer the original elements of a sequence are to the theoretical beats. The resulting value (*ugof*) is independent of the number of elements in the sequence, the sampling length, or the number of silent beats in a sequence (Figure 49A). It is also independent of the best-fitting beat frequency it is describing (Figure 49C).

We calculated the *ugof* of six datasets, it was calculated for two beat frequencies per sequence: the best-fitting beat frequency as calculated with Fourier analysis and the best-fitting beat frequency as calculated with the IOI approach (Figure 49C and D, as calculated with Fourier analysis and IOI analysis in previous studies) (Burchardt & Knörnschild, 2020; Burchardt et al., 2019). To be able to evaluate and interpret a single *ugof* we modelled the distribution of *ugofs* for a dataset. To this aim, we calculated *ugof* from 0.1 Hz to 100 Hz in 0.1 Hz increments for all element sequences in the dataset. To illustrate what we mean by this, let's assume we have a sequence A. For this sequence A, for which we know when each element in the sequence starts, we calculate *ugof* for 1000 beat frequencies (0.1 Hz to 100 Hz in 0.1 Hz increments), by calculating the actual deviations ($|\Delta|$) as well as the maximum possible deviations (Δ_{max}) for each frequency. Figure 49E shows the results of these calculations for a dataset of 49 isolation calls of the bat *Carollia perspicillata*. Shown are therefore 1000 *ugof* values for 49 sequences, giving us a distribution of 49000 values. We used the mean and standard deviation of this gaussian distribution to evaluate any single best-fitting beat frequency as calculated with the IOI approach. Then we can calculate z-scores for every *ugof* by subtracting the mean of the distribution of *ugof* values for the given dataset (i.e., isolation calls of *C. perspicillata)* from the *ugof* in question (i.e.,

the ugof as calculated for the best-fitting beat frequency with the IOI approach) and dividing the difference by the standard deviation of the distribution. A calculated z-score can than easily be matched to the corresponding p-value using z-score tables (i.e., (Fisher & Yates, 1964; Rohatgi & Saleh, 2015)). This allows us to investigate if a calculated beat frequency fits significantly better with the element sequence than what could be expected depending on the calculated distribution of *ugof* for a specific dataset. We only considered negative *z*-scores as possibly significant, as a negative z score indicates that the corresponding value is *below* the distributions mean (Figure 49E). A positive z-score would on the other hand indicate, that the corresponding ugof is *above* the distribution mean, and could also be significant, but would then fit significantly worse than expected by chance.

This approach of using *z*-scores, and therefore the possibility to calculate p-values for different production rhythms, is mainly useful for comparability reasons, in order to assess which animal or individual can better keep a stable (theoretical) beat. We do *not* want to propose that a sequence would only be well described by an isochronous beat that results in a significant *ugof*. To illustrate the methods, we calculated *z*-scores for beat frequencies in the dataset at hand based on the IOI approach, as these resulted in on average smaller *ugof* compared to beat frequencies calculated with Fourier analysis (Figure 49D). The values were calculated for beat frequencies with a resolution of one decimal point (i.e., 12.1 Hz or 28.8 Hz). This analysis revealed that *ugof* can change rapidly within small increments. We thus suggest that it is reasonable to have a look at *ugof* within 1 or 2 Hz of the detected best-fitting beat frequency to be aware of the possible sensitivity of the method. We obtained *ugof* in a range of ± 1 Hz around the best-fitting beat frequency based on the IOI approach, calculated the minimum and maximum *ugof* in that range, and transformed them into *z*-scores. The resulting *z*-scores for the best-fitting beat frequencies are shown as well as the z-scores for the minimum and maximum *ugof* within ± 1 Hz of the best-fitting beat frequency (Figure 49F). We found that the best-fitting beat frequency, especially when significant, was also the one with the best (i.e., lowest) *z*-score/ best *ugof* value within ± 1 Hz (indicated by the vertical lines in Fig. 49F: for significant results, the lines are almost exclusively above the datapoint showing the z-score for the corresponding best-fitting beat frequency). This might illustrate differences in beat production abilities between individuals, as some individuals might produce sounds in a more consistent/rhythmic way than others, which, in turn, could constitute a fitness indicator.

In addition, differences between significant and not significant production rhythms could be related to different situations, i.e., different arousal/motivation or urgency levels.



D

species	sequence type	mean ugof		n
		Fourier analysis	IOI analysis	
Alauda arvensis	flight song	0.49	0.50	14
Carollia perspicillata	isolation calls	0.34	0.27	49
Physeter macrocephalus	echolocation	0.54	0.49	60
Saccopteryx bilineata	isolation calls	0.56	0.54	500
Saccopteryx bilineata	echolocation	0.64	0.49	33
Saccopteryx bilineata	territorial songs	0.55	0.48	142

Figure 49: The ugof Value.
(A) Theoretical element series (solid black elements) with an overlaid beat (dashed lines) of a certain beat frequency in Hertz. The maximum possible deviation for any element is half the beat duration (Δ_{max}). It is set in relation to the absolute deviation of an element to its closest beat (Δ). Other important concepts visualized are: Inter-Onset-Intervals and silent beats. (B) The theoretical maximum deviation per beat (in black) and actual deviations (as mean per sequence) measured from six datasets and for two calculated beat frequencies each. Both deviations change in the same way depending on the corresponding frequency and actual deviations are much smaller than maximum possible deviations. (C) *ugof* calculated for best-fitting beat frequencies

based on Fourier analysis and IOI analysis for six datasets. No correlation can be seen between *ugof* and beat frequency. (D) Tabular comparison of mean *ugof* per dataset for both beat calculation methods. Fourier analysis yields better results (lower *ugof*) *only* for the complex skylark song. (E) Distribution of *ugof* calculated for beat frequencies from 0.1 to 100 in 0.1 Hz increments for all sequences of *C. perspicillata* isolation calls, to be able to calculate *z*-scores. (F) *Z*-scores as calculated based on the modelled *ugof* for beat frequencies of 49 isolation calls of *C. perspicillata* using IOI analysis. Significant values are in yellow, and non-significant values in orange. The minimum and maximum *ugof* of beat frequencies ± 1 Hz around the best-fitting beat frequency are shown as vertical lines. The differences between significant and not significant beat frequencies could correlate to different individuals, and potentially be connected to the relevance of beat production as a fitness indicator. Abbreviations: IOI: beat frequencies calculated with IOI analysis, FFT: beat frequencies calculated with a fast Fourier transformation (Fourier Analysis), FS: flight song, IC: Isolation calls, EC: echolocation calls, TS: territorial songs.

Additions to existing methods of rhythm analysis

Ten highest Peaks of Fourier Analysis

Especially in more complex signals, such as bird song comprised of various motifs or phrases (Aubin, 1982; Hultsch & Todt, 1981; Kroodsma, 2005), it seems inappropriate to assume that one beat frequency could be enough to describe a sequence. The IOI approach seems particularly unsuitable here, as it simplifies the structure (Burchardt & Knörnschild, 2020). The Fourier analysis, on the other hand, gives a very detailed picture of all beat frequencies that make up the sequence. It decomposes any signal into its sinus components, which are nothing else but frequencies. A sequence of an animal's acoustic signals is transformed into a binary sequence, where an element onset is encoded as '1' and everything else encoded as '0'. A fast Fourier transformation is then conducted on this binary sequence (Ravignani & Norton, 2017; Saar & Mitra, 2008). In a recent publication, we settled to describe a sequence by the beat frequency that contributed the most to the description of a sequence, i.e., the one beat frequency that gave the highest amplitude in the Fourier analysis's frequency domain, and only reported this most prominent frequency (Burchardt & Knörnschild, 2020). We now propose as an alternative, to report the ten most prominent frequencies. This would allow the detection of frequency "clusters" (red circles in Figure 50A). None of the frequencies in a cluster might have the highest peak. However, when combined (summed up), they surely describe a series better than a single, slightly higher peak. Therefore, it could also be an option to report a summary, average, or range of a particular cluster to describe a particular sequence. This gives a much more detailed result, which can be used as basis for decisions about how to proceed or what to report. We suggest looking at the ten highest peaks, as it is a reasonably high

number to find possible clusters, without reporting beat frequencies that have only very small explanatory values for the sequence. Nevertheless, for certain sequences, it may be most informative to report only the five highest peaks or the twenty highest peaks.

Recurrence Plots

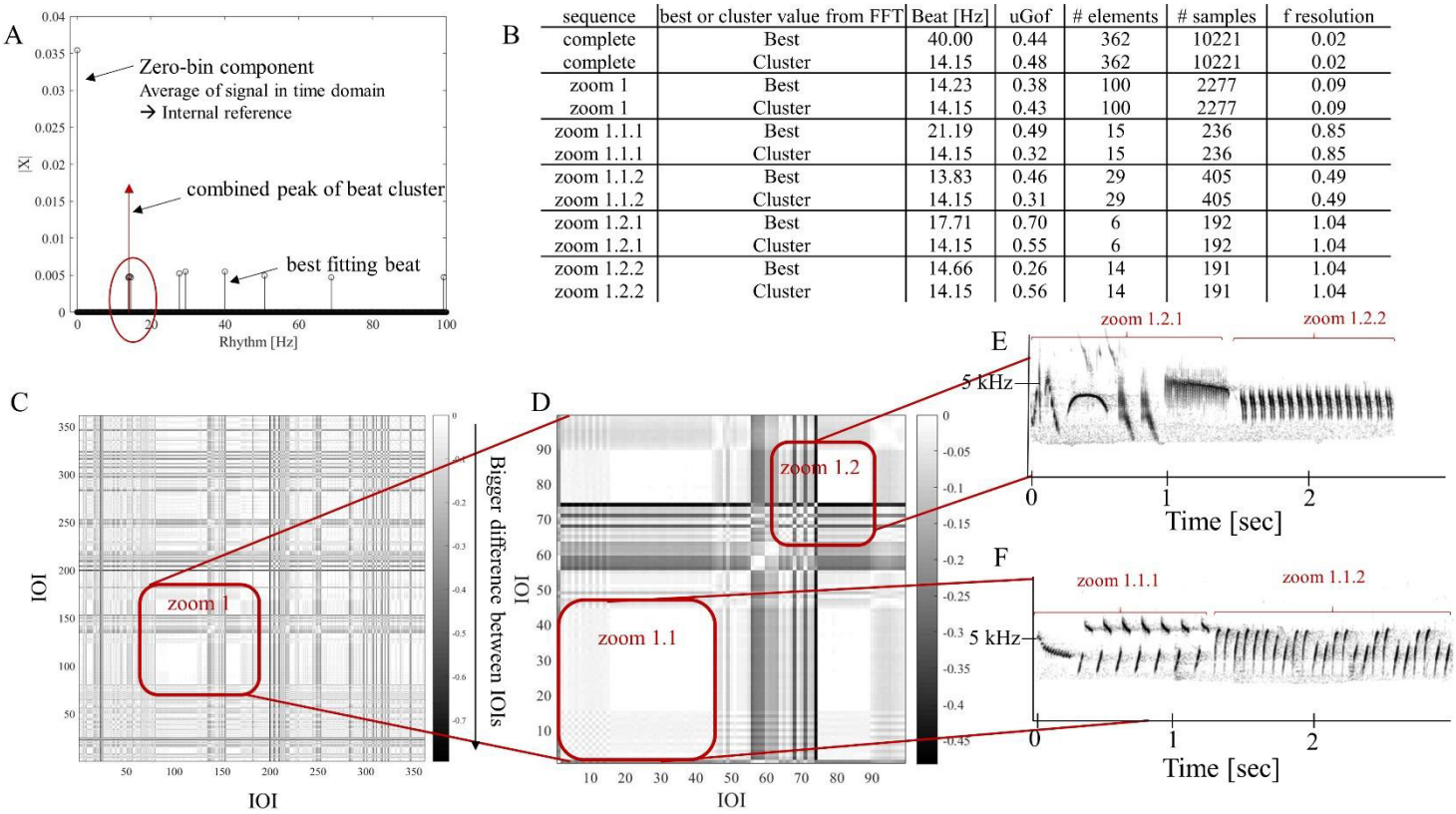
The recurrence plot, originally used in chaos theory (Eckmann et al., 1987; Marwan, 2008), is an easy way to visualize the overall temporal structure of a sequence and to find sub-units (Ravignani & Norton, 2017). It depicts the distance between any IOI pair in the sequence that is to be analysed. Every possible IOI pair is compared, the Euclidean distance is measured and plotted (Burchardt & Knörnschild, 2020; Ravignani & Norton, 2017). Differences are colour-coded in the plots; the darker a comparison, the more different are the two compared IOIs. Subunits with very different temporal structures can be easily spotted in such a plot, namely as a "break" in the pattern (Figure 50D). When analysing new acoustic signals, where knowledge about functional units such as motifs is scarce, such temporal breaks could easily show where a new motif or phrase starts. Furthermore, different sub-units might have different beat frequencies that convey meaning but that cannot be resolved with an overall best-fitting beat frequency. As we show in Figure 50B, identified sub-units can then be analysed to extract their specific best-fitting beat frequency, in order to see whether they fit the overall temporal structure or not.

Results of an exemplary analysis of skylark flight song

To illustrate the proposed additions, we analysed an excerpt from the complex flight song of a skylark. The specific sequence has a duration of 51.3 seconds and contains 362 elements (for details on recording see (Briefer et al., 2008; Briefer, Rybak, et al., 2010). We calculated the best-fitting beat frequency of the whole sequence and of five exemplary sub-units (there are more in the entire sequence), which we identified via recurrence plots (Figure 50C-F). As can be seen in Figure 50A, there is a strong cluster of beat frequencies with high descriptive value for this sequence. To calculate the universal goodness-of-fit, we used not only the single best-fitting beat frequency (indicated as "Best" in 50B), but also the beat frequency of the cluster mean (indicated as "Cluster" in 50B). The recurrence plot zoom 1.1 visualizes the subsequent switching between two element types followed by a series of very similar IOIs, which correspond to a single element type. In the recurrence plot zoom 1.2, on the other hand, we see more

variability, looking at both the recurrence plot and the spectrogram in 50E; there we can see various element types, then a "break" (black line) in the recurrence plot indicating a high difference between the two adjacent IOIs, followed again by a very stereotyped sub-unit (zoom 1.2.2). The calculated *ugofs* showed some interesting input. 1) The *ugof* for the best-fitting beat frequency of the whole sequence was indeed higher than the one for the cluster-beat frequency, but 2) the cluster-beat frequency described some of the sub-units well, sometimes even better than the best-fitting beat frequency calculated for the sub-units themselves. This was true especially for the four small sub-units (zoom 1.1.1 & 1.1.2 and zoom 1.2.1 & 1.2.2). Only one out of four sequences here showed a better (i.e., smaller) *ugof* for the best-fitting beat frequency (zoom 1.2.2) compared to the ugof calculated with the cluster beat frequency. However, here, both the cluster-beat frequency and the best-fitting beat frequency were very similar and, due to a lower frequency resolution in the Fourier analysis, the better fitting cluster-beat frequency could mathematically not be found (see (Burchardt & Knörnschild, 2020) for an explanation on frequency resolution in Fourier analysis).

Using the interpretation established in section 2.1, these *ugofs* range from fitting extremely well (zoom_1-2-2, "Best" or zoom_1-1-2, "Cluster") to fitting poorly (zoom_1-2-1, "Best" or zoom_1-2-2, "Cluster"). For the poor fits, the issue of a low frequency resolution needs to be kept in mind though, further proving that beat frequencies calculated by Fourier analysis with a low frequency resolution need to be handled and interpreted with care (Burchardt & Knörnschild, 2020).



sequence	best or cluster value from FFT	Beat [Hz]	uGof	# elements	# samples	f resolution
complete	Best	40.00	0.44	362	10221	0.02
complete	Cluster	14.15	0.48	362	10221	0.02
zoom 1	Best	14.23	0.38	100	2277	0.09
zoom 1	Cluster	14.15	0.43	100	2277	0.09
zoom 1.1.1	Best	21.19	0.49	15	236	0.85
zoom 1.1.1	Cluster	14.15	0.32	15	236	0.85
zoom 1.1.2	Best	13.83	0.46	29	405	0.49
zoom 1.1.2	Cluster	14.15	0.31	29	405	0.49
zoom 1.2.1	Best	17.71	0.70	6	192	1.04
zoom 1.2.1	Cluster	14.15	0.55	6	192	1.04
zoom 1.2.2	Best	14.66	0.26	14	191	1.04
zoom 1.2.2	Cluster	14.15	0.56	14	191	1.04

Figure 50: Exemplary Results of Rhythm Analysis of an Excerpt from Complex Flight Song of *A. arvensis*.
(A) Amplitude Plot of Fourier analysis. Beat frequency is depicted on the x-axis and the amplitude of the ten highest peaks – as calculated by a Fast Fourier Transformation – on the y-axis. The highest peak is always the zero-bin-component at 0 Hz; it is the average of the signal in the time domain, where elements were encoded in a binary sequence. One very strong cluster can be identified; a summary of this cluster might depict the temporal structure better than the detected single highest peak. (B) The table reports relevant parameters of the rhythm analysis for the five units depicted in the figure. (C) Recurrence plots of the complete example sequence: all IOI pairings in the sequence are compared to each other, forming a symmetric comparison of every IOI to every other IOI in the sequence. The Euclidean distance between any IOI pairing is colour coded. More different pairs of IOIs are characterized by longer distances and darker colours. (D) Zoom into a section of 100 elements (11.2 seconds) of the song sequence. A very consistent series of IOIs can be observed at the beginning, followed by some slight changes, and in the end again, a very consistent pattern. (E) Spectrogram of the zoom 1.2 section, which can further be divided into a variable pattern (zoom 1.2.1) and a very consistent pattern (zoom 1.2.2). (F) Spectrogram of the zoom 1.1 section, which can further be divided into two consistent patterns (zoom 1.1.1 and 1.1.2).

Discussion

Analysing the temporal structure of animals' acoustic signal is relevant for addressing many research questions, such as species discrimination, physiological correlates like couplings to wingbeat or respiration, mating preferences or arousal coding (Burchardt et al., 2019; David et al., 2003; Manser, 2001; McRae, 2020; Norton & Scharff, 2016). Other questions include duetting or the development of temporal structures during ontogeny (Pika et al., 2018; Sasahara et al., 2015; Yoshida & Okanoya,

2005). Many analyses conducted by bioacousticians include temporal parameters. We already indicated in an earlier paper (Burchardt & Knörnschild, 2020) that information, such as small scale inter-individual differences, might be lost by focussing only on the commonly used 'element rates', mostly called 'syllable rates' (Douglas et al., 2005; Manser, 2001; McRae, 2020). Using element rates or calculating beat frequencies per sequence by transforming the element rate into a frequency, could be described as a "spyglass" approach, mostly useful for studying highly temporally consistent communication signals (i.e., echolocation of bats or whales). It is useful when investigating a species' rhythm or other analyses that only require this level of detail. For more complex communication signals, or in cases when fine scale intra-individual differences or fine scale differences between contexts might play a role, the "magnifying glass" approach of the Fourier analysis should be used. Our newly established *ugof* clearly supports this claim, as our analyses revealed better results (indicated by lower *ugofs*) when using the Fourier analysis compared to the IOI approach *only* for the very complex skylark flight song.

Our suggested additions to already established methods, make these aims, of not losing relevant and interesting information during the analysis of temporal parameters, easier to reach. These new methods allow a comparison of rhythmicality both between studies and species, which was not easy beforehand. Analysing recurrence plots to make an educated decision on which sub-patterns to analyse can also be of interest when facing completely new acoustic signals. Clear temporal breaks, as can be seen in the recurrence plots shown above (Figure 50E), could easily indicate where a new motif or phrase starts. Distinguishing contexts or analysing syntax could be backed up by such analyses of the underlying temporal structure. An example for this could be research on dialects. Microgeographic differences between male skylarks' flight song is mostly based on differences in the syllable and phrase repertoire composition (Briefer et al., 2008). Such phrases could show a clearly distinguishable temporal patterning, shared phrases could then be automatically detected using this knowledge.

Our newly established universal goodness-of-fit value enables every researcher, whether reading such a study or conducting it, to grasp the rhythmicity of an individual, a single sequence, or a species, by looking at one number alone, which is accompanied by a p-value. A number between 0 and 1, with smaller numbers indicating a better fit, is easy to interpret. No understanding of correlations within the

data is needed. Furthermore, which of the methods used to describe a sequence (i.e., Fourier analysis or IOI approach) captures most of the underlying temporal structure, or whether a sub-unit has a beat frequency different from the beat frequency of the whole sequence, can be easily determined. It is to be noted, that a value of 1 is not expected, as this would mean that all elements of the sequence lie exactly between two beats, which would indicate that they all perfectly fit the theoretical beat, but phase displaced. Furthermore, it is the first proposed goodness-of-fit value that is calculated per element and not per sequence, therefore enabling bioacousticians to answer even more interesting questions about sub-level structures. Such questions could be about which elements "drive" a beat frequency or break it, which could then shed light on the accentuation of elements.

On another note, these methods are not only useful when analysing acoustic signals. They can be used on the temporal structure of anything, may it be a certain behaviour or physiological processes such as wingbeat, heartbeat, or respiration. All processes of interest can easily be transformed in a way to enable the analysis; for example, instead of interpreting the result in Hz, which is one beat per second, we could interpret it as beats/occurrences per hour, day, or more abstract processes such as a reproductive cycle. We can subsequently calculate our GAT, IOI or Fourier analysis on that particular time scale and retransform the results back to the original time scale. Rhythm analysis methods that have been developed for acoustic analysis could thus allow an even wider range of researchers in answering questions such as sleep cycle analysis, circadian rhythms, or various other research areas, possibly even in economics or engineering where temporal structures of processes are of utmost importance as well.

Funding

LSB was funded by an Elsa-Neuman Scholarship of the city of Berlin. MK was supported by a Heisenberg Fellowship (DFG KN935 5-1) from the German Research Foundation. EFB was funded by the French Minister of Research and Technology during the data collection. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We thank Iva Visnakova for labelling huge parts of the skylark flight song and Ahana Aurora Fernandez for inspiration and discussion on the universal goodness-of-fit value.

Data Availability Statement

All analysed data was already published in earlier studies (Elodie Briefer et al., 2008; Briefer, Rybak, et al., 2010; Burchardt & Knörnschild, 2020; Burchardt et al., 2019).

References

Aubin, T. (1982). Habituation au chant territorial chez l'alouette des champs (*Alauda arvensis* L.). *Biology of Behaviour, 7*, 353-362.

Behr, O., Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., & Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behavioural Ecology*(17), 810-817. https://doi.org/10.1093/beheco/arl013

Bøttcher, A., Gero, S., Beedholm, K., Whitehead, H., & Madsen, P. T. (2018). Variability of the inter-pulse interval in sperm whale clicks with implications for size estimation and individual identification. *The Journal of the Acoustical Society of America, 144*(1), 365-374. https://doi.org/10.1121/1.5047657

Briefer, E., Aubin, T., Lehongre, K., & Rybak, F. (2008). How to identify dear enemies: the group signature in the complex song of the skylark *Alauda arvensis*. *Journal of Experimental Biology, 211*, 317-326. https://doi.org/10.1242/jeb.013359

Briefer, E., Osiejuk, T. S., Rybak, F., & Aubin, T. (2010). Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. *Journal of Theoretical Biology, 262*(1), 151-164. https://doi.org/10.1016/j.jtbi.2009.09.020

Briefer, E., Rybak, F., & Aubin, T. (2008). When to be a dear enemy: flexible acoustic relationships of neighbouring skylarks, *Alauda arvensis*. *Animal Behaviour, 76*, 1319-1325. https://doi.org/10.1016/j.anbehav.2008.06.017

Briefer, E., Rybak, F., & Aubin, T. (2010). Are unfamiliar neighbours considered to be dear-enemies? *PLoS One, 5*(8), e12428. https://doi.org/10.1371/journal.pone.0012428

Burchardt, L. S., & Knörnschild, M. (2020). Comparison of methods for rhythm analysis of complex animals' acoustic signals. *PLOS Computational Biology, 16*(4), e1007755. https://doi.org/10.1371/journal.pcbi.1007755

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science, 6*. http://doi.org/10.1098/rsos.181076

David, J. A. d. O., Zefa, E., & Fontanetti, C. S. (2003). Cryptic species of *Gryllus* in the light of bioacoustic (Orthoptera: Gryllidae). *Neotropical Entomology, 32*, 75-80. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-566X2003000100010&nrm=iso

Douglas, L. A., Dawson, S. M., & Jaquet, N. (2005). Click rates and silences of sperm whales at Kaikoura, New Zealand. *The Journal of the Acoustical Society of America, 118*(1), 523-529. https://doi.org/10.1121/1.1937283

Eckmann, J.-P., Kamphorst, S., Oliffson, S., & Ruelle, D. (1987). Recurrence Plots of Dynamical Systems. *Europhysics Letters, 4*(9), 973-977. https://doi.org/10.1209/0295-5075/4/9/004

Fisher, R. A., & Yates, F. (1964). *Statistical Tables for Biological, Agricultural and Medical Research, Sixth Edition* (6th ed.). Oliver and Boyd.

Hultsch, H., & Todt, D. (1981). Repertoire sharing and song-post distance in nightingales (*Luscinia megarhynchos B*.). *Behavioural Ecology and Sociobiology, 8*(3), 183-188. https://doi.org/10.1007/BF00299828

Knörnschild, M., Feifel, M., & Kalko, E. K. V. (2013). Mother-offspring recognition in the bat *Carollia perspicillata*. *Animal Behaviour, 86*, 941-948. https://doi.org/10.1016/j.anbehav.2013.08.011

Knörnschild, M., Jung, K., Nagy, M., Metz, M., & Kalko, E. K. V. (2012). Bat echolocation calls facilitate social communication. *Proceedings of the Royal Society B, 279*, 4827-4835. https://doi.org/10.1098/rspb.2012.1995

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Animal Behaviour, 84*, 761-769. http://dx.doi.org/10.1016/j.anbehav.2012.06.029

Kroodsma, D. E. (2005). *The Singing Life of Birds: The Art and Science Listening to Birdsong*. Houghton Mifflin Company.

Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B, 268*(1483), 2315-2324. https://doi.org/10.1098/rspb.2001.1773

Marwan, N. (2008). A historical review of recurrence plots. *The European Physical Journal Special Topics, 164*(1), 3-12. https://doi.org/10.1140/epjst/e2008-00829-1

McRae, T. R. (2020). A review of squirrel alarm-calling behaviour: What we know and what we do not know about how predator attributes affect alarm calls. *Animal Behaviour and Cognition, 7*(2), 168-191. https://doi.org/10.26451/abc.07.02.11.2020

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

Payne, R. S., & McVay, S. (1971, Aug 13). Songs of humpback whales. *Science, 173*(3997), 585-597. 10.1126/science.173.3997.585

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B, 285*, 20180598. https://doi.org/10.1098/rspb.2018.0598

Ravignani, A. (2018). Spontaneous rhythms in a harbor seal pup calls. *BMC research notes, 11*(1). https://doi.org/10.1186/s13104-017-3107-6

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speech, movement and animal vocalizations. *Journal of Language Evolution*. https://doi.org/10.1093/jole/lzx002

Rohatgi, V. K., & Saleh, A. K. M. E. (2015). Statistical Tables. In V. K. Rohatgi & A. K. M. E. Saleh (Eds.), *An Introduction to Probability and Statistics* (pp. 647-665). https://doi.org/https://doi.org/10.1002/9781118799635.oth1

Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One, 3*(1), e1461. https://doi.org/10.1371/journal.pone.0001461

Sasahara, K., Tchernichovski, O., Takahasi, M., Suzuki, K., & Okanoya, K. (2015). A rhythm landscape approach to the developmental dynamics of birdsong. *Journal of the Royal Society Interface, 12*(112). https://doi.org/10.1098/rsif.2015.0802

Tønnesen, P., Gero, S., Ladegaard, M., Johnson, M., & Madsen, P. T. (2018). First-year sperm whale calves echolocate and perform long, deep dives. *Behavioural Ecology and Sociobiology, 72*(10), 165. https://doi.org/10.1007/s00265-018-2570-y

Yoshida, S., & Okanoya, K. (2005). Evolution of Turn-Taking: A Bio-Cognitive Perspective. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 12*(3), 153-165. https://doi.org/10.11225/jcss.12.153

General Discussion

Rhythms can be found anywhere in the world: our hearts have rhythms, circadian rhythms are all around, music across all cultures shares certain components such as rhythm, and the rhythms of acoustic communication get more and more attention. The rhythms of animals' acoustic communication and the issue of how to analyse them were under study in my thesis. I looked at rhythm production and perception in mammals and birds. A major part of this work was devoted to the evaluation of existing methods and their combination into a clear workflow. Another key finding of my thesis was the development of a universal goodness-of-fit value (ugof) that can evaluate how well a beat frequency (in Hertz as in beats per second) describes any element sequence. I also investigated how bats perceive different isochronous rhythms by adjusting an ABR procedure. Apart from that I reported evidence for a common isochronous rhythm in three vocalization types of the greater sac-winged bat *Saccopteryx bilineata*.

In the following, I will briefly summarise the findings presented in the previous chapters before summarising and discussing the methods used in this work. I will discuss the limitations of the presented work and provide an outlook for future research, including possible applications of methods in the context of the literature and current discussions in the field.

Summary of Findings

Isochrony: a common Pattern with Need for careful Distinctions

17 datasets from 14 different species and a total of 940 sequences were analysed in my thesis and in most cases, a reasonably well-fitting isochronous beat frequency was found to describe the sequence. Underlying isochronous patterns can often be found in animal communication, which might not be surprising as it is argued to be a very simple rhythmic pattern that could simply result from a motoric behaviour entraining to a neural oscillator (Ravignani, in press).

In Chapter I, I found that three sequence types of the greater sac-winged bat *Saccopteryx bilineata* share a common rhythm of around 6 to 24 Hz, that can be linked to the wingbeat (around 12 Hz) frequencies of that particular species, even though two of the three sequence types were uttered while bats were perching (Burchardt et al., 2019). The isochronous beats were analysed with the GAT approach and

validated by comparison with random data. In Chapter II two more sequence types from two more species were analysed and showed an isochronous rhythmic pattern: isolation call sequences of Seba's short-tailed bat *Carollia perspicillata* and echolocation call sequences of the sperm whale *Physeter macrocephalus* are also well described by isochronous beats (Burchardt & Knörnschild, 2020). The same was true for echolocation call sequences of 10 additional bat species analysed for Chapter III (see Table 1 for species). Even the very complex flight song of male skylarks (*Alauda arvensis*) could be described with isochronous beat frequencies. Recommendations on which methods to use for the analysis of such complex signals were presented in Chapter IV.

Some key issues need to be addressed when discussing the isochrony of animals' acoustic signals, though. A very important distinction we must make is the one between signal isochrony and induced isochrony (Honing, 2012; Honing & Ploeger, 2012; Ravignani, in press). Signal isochrony might be what most people think of when they talk about an element sequence having an isochronous rhythm: a signal that itself is isochronous. Opposed to that, we also have induced isochrony, where a sequence is well described by an underlying isochronous beat, but not every beat is accompanied by an element, and isochrony is only induced when perceived. Sequences showing signal isochrony can well be analysed with the simple IOI approach, while this approach might prove insufficient for sequences with induced isochrony. We found signal isochrony for echolocation call sequences of bats and the sperm whale and isolation calls of *C. perspicillata* (i.e., (Burchardt & Knörnschild, 2020) and Chapter III). Isolation calls and territorial song of *S. bilineata* were described well with the IOI approach, assuming signal isochrony, but were also well described by beat frequencies calculated with the GAT approach or Fourier analysis, that indicated signal isochrony in some and induced isochrony in other sequences (Burchardt et al., 2019). Generally, while signal isochrony is characterized by a low variability between IOIs, for induced isochrony the variability can be much higher. Thus, simple distributional parameters such as the nPVI for example would yield results that could easily be interpreted as indicating a random patterning in the case of induced isochrony (Burchardt & Knörnschild, 2020; Cameron et al., 2019; Duffy & Pearce, 2018). The distinction between signal isochrony and induced isochrony that has not yet been made consequently would add strongly to the clarity of results and help to illustrate the differences between rhythmic patterns in different sequences or species. Principally, it is very important, that the variability

of sequences gets reported reliably and comparably. Many studies report variability parameters only sparsely or in the form of parameters that yield problems themselves (i.e., standard deviation and variance, critically discussed in Chapter II). A clear recommendation would be to report not only standardized variability parameters such as the coefficient of variation, which has been introduced for this kind of data in Chapter II, but also the raw data, i.e., the IOI sequences, or even the start and end duration of labelled elements of interest. Meeting these recommendations will enable broad cross-species comparison as if need be, data could just be re-analysed.

Beat Perception in Bats: too fast to follow?

The more we know about the production of rhythmic patterns, the more important it gets to have a clear understanding of how these patterns are perceived and whether they carry meaning for the animals as in the northern Elephant Seal (Mathevon et al., 2017)) or are merely production correlates or depending on other physiological or external processes such as in some cricket species (Doherty, 1985).

In Chapter III, we investigated the perception of isochronous beats in 12 bat species in the wild and captivity for natural and artificial stimuli and stimulus presentation rates of 6 to 100 Hz. The first key finding was that slower presentation rates (i.e., 6 Hz) on average elicit higher reactions than faster presentation rates (i.e., 100 Hz). This was true for all analysed species. The second key finding was that natural stimuli that elicited a significant response, consequently elicited higher responses than artificial stimuli tested with the same setup in the same species (*C. perspicillata*). We tried to correlate the response strength for different stimulus presentation rates to production rhythms of search flight echolocation calls of the respective species. In many species, the production rhythms of these calls correlated well with the stimulus presentation rates, which elicited higher reactions (i.e., in *Lonchorhina aurita*, *Saccopteryx bilineata,* or *Myotis nigricans,* see Chapter III). It is to be noted that echolocation calls need to be produced very flexibly in terms of their timing, as production frequencies can rapidly increase in situations with a lot of clutter or for feeding attempts right before prey capture (Moss & Surlykke, 2001; Ratcliffe et al., 2013). At first sight, our results suggest that production rhythms in such situations are almost too fast to follow for the auditory apparatus, which is in line with electrophysiological studies in the auditory cortex of *C. perspicillata*, where cortical neurons could track frequencies of only up to 22 Hz (Martin et al., 2017). On the other hand, our experiments disregard an

important process: attention. The results of this study need to be backed up with behavioural experiments to include the effect of attention on the perception of faster beats that elicit lower reactions in the auditory brainstem in our setup. Production rhythms of social communication calls match the stimulus presentation rates eliciting higher reactions (i.e., lower presentation rates) as well for *S. bilineata* and *C. perspicillata*; thus, it might be interesting to disentangle whether temporal perception processes in bats are driven by social communication or echolocation.

Another process to be considered when interpreting the results is neuronal adaptation. The faster the stimulus presentation rates, the higher is the overall stimulus energy, and thus the possibility for neuronal adaptation. This is especially likely to affect the results, as a quite similar setup was used to study exactly these adaptation processes in the auditory brainstem of cats, gerbils, mice, chicken, humans, and echolocating bottle-nosed and common dolphins before (Burkard et al., 1994; Burkard & Sims, 2002; Burkard et al., 1997; Burkard et al., 2017; Burkard et al., 1996a; Burkard et al., 1996b; Burkard & Voigt, 1989; Ridgway et al., 1981). It is unclear how big the effect of adaptation is in our results, as there are some prominent differences between methodology and analysis: a) in former studies stimuli were presented more often (500 times vs. 256 times) and thus longer, which would increase effects of adaptation, and b) former studies analysed data not by calculating the overall response strength but analysed peak latency and peak amplitudes of single peaks. Some result characteristics differ as well between our study and the above-mentioned studies. For example, neuronal adaptation is supposed to lead to an almost linear decline of response strength towards higher stimulus presentation rates (Wiegrebe & Schmidt, 1996), as it was also observed in the bottle-nosed dolphin (Burkard et al., 2017). We could not always observe such a linear decline especially not on the level of single individuals. Nevertheless, adaptation is an important process to consider when studying beat perception through ABRs, and the experimental setup should be adjusted in follow-up studies to account for that. The ultimate solution would be to account for the energy differences between i.e., 6 Hz presentation rates and 100 Hz presentation rates, by adjusting the sound pressure accordingly. As the aim of this study was to investigate whether we would find differences in the general response strength for different stimulus presentation rates, the analysis approach of calculating the integrals under the whole curve was chosen, but the peak frequency and peak latency as commonly used parameters analysed in ABR experiments

(Brittan-Powell et al., 2002; Burkard et al., 2017) could also be calculated in future studies. In follow up experiments, it would furthermore be interesting to not only test more stimulus types, for example, more natural stimuli, and natural stimuli in more species, but also to re-run experiments using broadband clicks as stimuli instead of artificial stimuli of the peak frequency of the respective species' echolocation calls. Another interesting option, especially when studying bats, would be to apply so-called "maximum-length-sequences" (Burkard et al., 1994). Here, pseudorandom stimulus sequences with changing intervals between stimuli (i.e., between 0.5 and 6 ms in one sequence) are presented. This might mimic situations better in which echolocation calls are produced with flexible inter-onset-intervals and therefore give insights into another component of the temporal perception of stimuli in a species, where tempo perception is so important.

By and large, the results of these experiments are of general importance for the implementation of ABR experiments, as most studies running ABR experiments are using artificial stimuli (i.e., (Brittan-Powell et al., 2002; Burkard et al., 2017; Land et al., 2016; Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019; Wetekam et al., 2020), while our study suggests the use of natural stimuli. The same was suggested by electrophysiological studies in the auditory cortex of *C. perspicillata*, where natural stimuli also elicited higher reactions compared to artificial stimuli, indicating a higher relevance for the animals (Beetz et al., 2016). Furthermore, our results need to be considered when choosing stimulus presentation rates as many studies use quite high presentation rates of 40 to 50 Hz (i.e., (Brittan-Powell et al., 2002; Lattenkamp et al., 2021; Linnenschmidt & Wiegrebe, 2019; Wetekam et al., 2020), whereas our results suggest to use very low stimulus presentation rates between 6 and 10 Hz.

Summary of Methods: existing and newly established.

Various methods were used during my work to analyse and evaluate temporal structures of animals' acoustic signals on different levels. All methods that were used are summarised in Table 15.

While I used already established methods to answer a biological question in Chapter I (GAT approach: (Norton & Scharff, 2016)) and Fourier analysis: (Saar & Mitra, 2008)), Chapter II was aiming at developing a clear workflow and at filling gaps in the current methodological framework. Missing was a goodness-of-fit value for the Fourier analysis. Furthermore, the commonly used call rates or IOIs

(Manser, 2001; Moss et al., 2006; Ravignani, 2018) were reported in a different unit than the results of the Fourier analysis and the GAT approach, which is why I introduced the IOI approach. It transforms averaged IOIs of a sequence into a frequency in Hertz. A goodness-of-fit value was then also needed for this IOI approach and I introduced the use of the coefficient of variation for that purpose. Chapter II aimed also at giving exemplary results for many of the discussed methods for biological data, as so far only artificial data had been used, to illustrate possible results of the different methods (Ravignani & Norton, 2017). The key result of this work was a workflow with clear instructions on how to analyse what kind of data, to be able to evaluate the temporal structure of a given sequence reasonably. It got apparent that depending on the data at hand and assuming isochrony, all three methods to extract exact beat frequencies (IOI approach, GAT approach, and Fourier analysis) prove valuable for different data and/or different questions. The more consistent a pattern is, the better suited is the IOI approach to describe a sequence. It is also valuable to describe very short sequences. The IOI approach proves furthermore useful for overview questions and could be described as a "spyglass" approach. If more details are required and small differences potentially needed to be disentangled, the IOI approach could oversimplify results. Especially for questions regarding inter-individual differences the GAT approach and Fourier analysis are more suitable, as they can detect smaller differences. Furthermore, these two can find induced isochrony (i.e., (Burchardt et al., 2019; Norton & Scharff, 2016)), while the IOI approach, due to its nature of using an average, is prone to almost always suggest signal isochrony. The GAT approach has a disadvantage, though, that is separating it from the Fourier analysis: it is sensitive to tempo changes. While at first this sounds like an advantage, it is not. As the result of the GAT approach is a beat frequency that on average fits best to all element onsets, this is problematic in a situation, where e.g., the ten first elements are perfectly isochronous while the last four elements increase in tempo rapidly. The method now still finds the one beat, with the least frequency normalized root-mean-square deviation (FRMSD) considering *all* elements (Norton & Scharff, 2016). This might result in a situation where a beat frequency is presented as "best-fitting" that is matching neither the perfectly isochronous 10 first elements, nor the faster elements in the end (Burchardt & Knörnschild, 2020). Fourier analysis is stable for such situations, and to be preferred for sequences with tempo changes. It can also be described as the opposite to the IOI approach as such that it can detect small differences and

could therefore be described as a “magnifying glass” approach in contrast to the “spyglass” approach. Even though goodness-of-fit values were established for Fourier analysis and the IOI approach, the three values indicating goodness-of-fit for the three methods to extract exact beat frequencies were still difficult to compare between each other for the same sequence and thus even more difficult to compare between species or studies. To circumvent these issues, I set out to develop a goodness-of-fit value that would be applicable for any of the described methods, therefore yielding comparable results (Chapter IV). The value I developed and termed “*ugof*” for “*universal goodness-of-fit value*” has even more advantages than being easily comparable: it can without difficulty be calculated for a whole range of sequences and beat frequencies, enabling us to see the true distribution of the ugof in a specific dataset. Using the population mean and standard deviation, z-scores can be calculated for exact beat frequencies that can, in turn, be matched to their corresponding p-values (Fisher & Yates, 1964; Rohatgi & Saleh, 2015), enabling us for the first time to put a p-value on an isochronous beat frequency describing an element sequence. Another important novelty for this specific value is the fact, that in contrast to previously described goodness-of-fit values, it can be calculated for single elements in a sequence. While the FRMSD, nGOF and GOF as well as the coefficient of variation (see Table 15 for explanation) are calculated for a whole sequence, the ugof is calculated per element and can then be reported as required per sequence, per individual, per species or any other level of analysis that is interesting. All this opens a whole new field of possible investigations that will be further discussed in the Outlook section.

Table 15: Overview of Rhythm Analysis Methods Used in this Thesis.

Method	Type and Aim	Chapter	Usage & Development
Histograms of IOIs	Visualisation; assess isochrony	I, II	Used
Coefficient of Variation	Variability parameter/ goodness-of-fit value; assess isochrony, assess goodness-of-fit of IOI approach	II, III, IV	Developed in Chapter II for this purpose
nPVI	Variability parameter; assess isochrony	II	Used
IOI approach	Exact beats, spyglass approach	II, III, IV	Developed in Chapter II
GAT approach	Exact beats	I, II	Used
Fourier Analysis	Exact beats; magnifying glass approach	I, II, IV	used, additions proposed in Chapter IV
Recurrence Plots	Visualisation; find substructures	II, IV	Used, intensely discussed in Chapter IV
FRMSD	Goodness-of-fit value; assess goodness-of-fit of GAT approach	I, II	Used
GOF/nGOF	Goodness-of-fit value; assess goodness-of-fit of Fourier analysis	II	Developed in Chapter II
Ugof	Goodness-of-fit value; assess goodness-of-fit for any of the three methods	IV	Developed in Chapter IV

Conclusion and Outlook

The established methods and obtained results open the field to investigate many new questions or hold the potential to improve yet other analysis approaches. I will discuss open questions and possible future applications in the last paragraph of this thesis.

Overall, analysing the temporal structure of animals' acoustic signals can be relevant for addressing many research questions, such as questions on species discrimination, physiological correlates like couplings to wingbeat or respiration, mating preferences, or arousal coding (Burchardt et al., 2019; David et al., 2003; Manser, 2001; McRae, 2020; Norton & Scharff, 2016). Other questions include duetting or the development of temporal structures during ontogeny (Pika et al., 2018; Sasahara et al.,

2015; Yoshida & Okanoya, 2005). It can also inform about vocal production learning processes (Wirthlin et al., 2019) and aims at helping to understand the evolution of both music and language in humans (Honing, 2012; Ravignani et al., 2014). In that capacity, it could also enable us to better understand various speech and vocalization impairments as well as perception deficits, connected to rhythmic structures of signals in humans and other animals (Norton et al., 2019; Wieland et al., 2015). Moreover, rhythm analysis should in the future be used more intensively, to set up behavioural experiments including any form of acoustics in a more informed way, therefore ensuring that for example in playback experiments the study animal is presented with relevant stimuli in terms of the temporal structure (as it was done in (Ravignani, 2018)).

Furthermore, the aim of rhythm analysis remains to uncover possible functions and reasons for the rhythmic production of sounds. For that further investigation into the perception of stimuli might prove valuable. Another line of investigation could try to link the production of rhythms to motor-correlates or neuronal correlates. A good example of this are the results from the Egyptian fruit bat (*Rousettus aegyptiacus*). Here wingbeat and tongue clicks are tightly coupled around 10 Hz (Yartsev & Ulanovsky, 2013), as we also found for *S. bilineata* between wingbeat and laryngeal sound production of three sequence types. These beat frequencies show a resemblance to the frequency of theta brain waves. Thought to be important for spatial memory, movements, and active stimulus intake (Colgin, 2013) amongst others, theta waves might be a promising neural correlate explaining the production of the detected rhythms in *S. bilineata* and *R. aegyptiacus*.

Rhythm analysis also has great potential in furthering research on social interactions. Temporal plasticity can be argued to play an important role in animal communication. Dynamics that would be interesting here include chorus situations resulting in synchrony or antiphony[5], where signallers adjust their signal emission either to synchronise with other individuals (Greenfield, 1994; Greenfield & Schul, 2008; Nityananda & Balakrishnan, 2006), interrupt other signallers, or specifically avoid to overlap with them (i.e., (Gochfeld, 1978; Greenfield, 1994; Hultsch & Todt, 1982; Martínez-Rivera & Gerhardt, 2008; Naguib, 1999; Tárano & Carballo, 2016)). We also know from singing mice that the motor cortex

[5] Describing a phase relation between elements of 180°

influences the pace of the singing behaviour on a moment-by-moment basis, depending on the song of another present and singing individual, thus shaping the rhythm of the communication attempt (Okobi et al., 2019). A different example of social interactions shaping the temporal structure of animals' acoustic signals was found in two frog species in Australia, where beat frequencies of calls in a chorus increase significantly for both frog species when a competitor species is present and calling as well (Filer et al., accepted).

Currently, more and more ideas are being suggested on how to study the interactions between neural underpinnings, rhythm production, and rhythm perception and how to investigate rhythms as the multimodal feature they are (Anichini et al., 2020; Pouw et al., 2021), which I can only support. Again, the description of clear methodologies, interpretation examples, and parameters supporting comparison will help in that regard, as all of that enables non-experts to use different methods or interpret published results, and thus foster interdisciplinary studies.

Especially the ugof value holds great potential for future applications, interdisciplinary studies, cross-species comparison, and the further improvement of already existing methods. It would for example be interesting to analyse the ugof throughout ontogeny for various settings. It would be interesting to see whether we can find changes in the goodness-of-fit of rhythms during ontogeny in general, and more specific to compare between for example learned and innate vocalizations, or between vocal learning species and non-vocal learners. This could shed light on the importance of rhythms during the learning process and could prove/disprove the hypothesis that rhythm and melody serve as guiding mechanisms for each other during the ontogeny of sound production: as producing sounds on an isochronous grid, making them predictable, would leave cognitive capacity to focus on the spectral domains of a call-template an individual is trying to learn (through a tutor or rehearsing on auditory feedback (Ravignani, in press)). Even though we did not find significant differences in the exact beats in isolation calls of *S. bilineata* during ontogeny (Burchardt et al., 2019), it might well be that the goodness-of-fit value does change during that time. This might be even more interesting to analyse in the conspicuous babbling behaviour of *S. bilineata* pups (Fernandez, 2020; Knörnschild et al., 2006). It could be analysed whether the ugof differs for specific element types uttered by pups compared to the same element types uttered by adults.

Getting back to the definition for rhythm used in the introduction, that rhythm is as 'systematic patterning of sound in terms of timing, accent, and grouping' (Patel, 2008), methods used and discussed in my thesis were mostly focusing on describing the timing of such systematic patterns (Chapter I to III). The introduced ugof on the other hand has the potential to analyse the other two aspects of this definition as well. Both grouping and accentuation could be investigated employing the ugof, through a detailed analysis of single elements and their ugof. To illustrate that with an example, we might consider the isolation calls of *S. bilineata* analysed in Chapter I. One could try to analyse the ugof of different element types in these calls. We might find that a certain element type always shows lower ugof, compared to another element type, leading to the possible interpretation that this particular element type is stressed. The same could be done for whole sequence types, i.e., for different social communication sequence types, but also different situations in terms of for example arousal or urgency.

Directly connected to that (arousal and urgency) rhythm analysis in combination with ugof calculations could also benefit emotions research, relying amongst others on the measurement of heart rates. Heart rates are an often used parameter in emotions research (i.e., (Forkman et al., 2007)) and especially the heart rate variability that is tested in some studies (i.e.,(Briefer et al., 2015)) could benefit from using the ugof to assess this variability in more detail.

Another possible application that was discussed in Chapter IV is the use of the ugof to back up analyses to distinguishing contexts or to analyse syntax. One concrete example here could be research on dialects[6], that can be found for example in different bird and bat species (Boughman & Wilkinson, 1998; Boughmann, 1998; Briefer et al., 2008; Davidson & Wilkinson, 2002; Nelson, 2000; Nelson & Poesel, 2007; Prat et al., 2017). Microgeographic differences between male skylarks' flight songs or male *S. bilineatas'* territorial songs for example are in parts based on differences in the element type (syllable) and sequence type (phrase) repertoire of a certain population (Briefer et al., 2008; Davidson & Wilkinson, 2002). Such sequence types could show a distinguishable temporal patterning, shared

[6] Local vocal variation within an animal species (i.e., Henry, L., Barbu, S., Lemasson, A., & Hausberger, M. (2015). Dialects in Animals: Evidence, Development and Potential Functions. *Animal Behavior and Cognition, 2*(2), 132-155. https://doi.org/10.12966/abc.05.03.2015)

sequence types could then be automatically detected using the knowledge about temporal structures and ugof.

As mentioned above, the long-term aim should be to combine approaches on different levels, such as behavioural characteristics and neural underpinnings to investigate rhythm as a single piece in the large puzzle of communication. To achieve that, we need a clear terminology that includes terms used in ecology and evolutionary research, terms used in electrophysiological studies and bioacoustics to terms used in musicology and linguistics. That is why we included a glossary in Chapter II, explaining different terms that different fields might use in slightly different meanings. Another recent study added to that: they compared terms from different fields and combined them in an overall terminology framework (Pouw et al., 2021). Furthermore, we need well-documented data and codes accessible for future studies. All of this will in turn foster cross-species comparison as well as cross-modality analyses and comparisons, which are both argued to be of utmost importance (Pouw et al., 2021; Ravignani et al., 2014) but have received little to no attention so far. Different species, species groups, or levels of rhythmicality (i.e., behavioural or neural) are currently most often investigated as separate entities but need to be investigated in a connected way (Pouw et al., 2021). Only then will we be able to truly disentangle the adaptive functions as well as the evolution of rhythmicality in acoustic communication.

My thesis showed examples of rhythmic structures of animals' acoustic signals. Isochronous patterns were found in bats, birds, and whales. Multiple methods to analyse rhythms were tested and evaluated as well as new methods developed. Clear recommendations were given on how to perform rhythm analysis for different datasets and questions and finally, open questions were raised and future applications for newly established methods proposed. All of which in the end will hopefully contribute to a better understanding of rhythms in animals' acoustic signals and help in the quest to uncover adaptive functions and the evolution of rhythmicality in acoustic communication in humans and other animals alike, furthering research on the evolution of language as well as music, and the general knowledge about the different aspects and their importance of acoustic communication in animals.

References

Anichini, M., de Heer Kloots, M., & Ravignani, A. (2020). Interactive rhythms in the wild, in the brain, and in silico. *Canadian Journal of Experimental Psychology 74*(3), 170-175. https://doi.org/10.1037/cep0000224

Beetz, M. J., Hechavarria, J. C., & Kössl, M. (2016). Temporal tuning in the bat auditory cortex is sharper when studied with natural echolocation sequences. *Scientific Reports, 6*. https://doi.org/10.1038/srep29102

Boughman, J. W., & Wilkinson, G. S. (1998). Greater spear-nosed bats discriminate group mates by vocalizations. *Animal Behaviour, 55*(6), 1717-1732. https://doi.org/10.1006/anbe.1997.0721

Boughmann, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society B, 265*, 227-233. https://doi.org/10.1098/rspb.1998.0286

Briefer, E., Aubin, T., Lehongre, K., & Rybak, F. (2008). How to identify dear enemies: the group signature in the complex song of the skylark *Alauda arvensis*. *Journal of Experimental Biology, 211*, 317-326. https://doi.org/10.1242/jeb.013359

Briefer, E. F., Tettamanti, F., & McElligott, A. G. (2015). Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour, 99*, 131-143. https://doi.org/10.1016/j.anbehav.2014.11.002

Brittan-Powell, E. F., Dooling, R. J., & Gleich, O. (2002, Sep). Auditory brainstem responses in adult budgerigars (*Melopsittacus undulatus*). *The Journal of the Acoustical Society of America, 112*(3), 999-1008. https://doi.org/10.1121/1.1494807

Burchardt, L. S., & Knörnschild, M. (2020). Comparison of methods for rhythm analysis of complex animals' acoustic signals. *PLOS Computational Biology, 16*(4), e1007755. https://doi.org/10.1371/journal.pcbi.1007755

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science, 6*. http://doi.org/10.1098/rsos.181076

Burkard, R., Jones, S., & Jones, T. (1994). Conventional and cross-correlation brain-stem auditory evoked responses in the white leghorn chick: rate manipulations. *The Journal of the Acoustical Society of America, 95*(4), 2136-2144. https://doi.org/10.1121/1.408675

Burkard, R., & Sims, D. (2002). The Human Auditory Brainstem Response to High Click Rates. *American journal of audiology, 10*(2), 53-61. https://doi.org/10.1044/1059-0889(2001/008)

Burkard, R., Trautwein, P., & Salvi, R. (1997). The effects of click level, click rate, and level of background masking noise on the inferior colliculus potential (ICP) in the normal and carboplatin-treated chinchilla. *The Journal of the Acoustical Society of America, 102*(6), 3620-3627. https://doi.org/10.1121/1.420149

Burkard, R. F., Finneran, J. J., & Mulsow, J. (2017). The effects of click rate on the auditory brainstem response of bottlenose dolphins. *The Journal of the Acoustical Society of America, 141*(5), 3396. https://doi.org/10.1121/1.4983447

Burkard, R. F., McGee, J., & Walsh, E. J. (1996a). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. I. Peak latencies. *The Journal of the Acoustical Society of America, 100*(2), 978-990. https://doi.org/10.1121/1.416209

Burkard, R. F., McGee, J., & Walsh, E. J. (1996b). Effects of stimulus rate on the feline brain-stem auditory evoked response during development. II. Peak amplitudes. *The Journal of the Acoustical Society of America, 100*(2), 991-1002. https://doi.org/10.1121/1.416210

Burkard, R. F., & Voigt, H. F. (1989). Stimulus dependecies of the gerbil brain-stem auditory-evoked response (BAER). I: Effects of click level, rate, and polarity. *The Journal of the Acoustical Society of America, 85*(6), 2514-2525.

Cameron, D. J., Zioga, I., Lindsen, J. P., Pearce, M. T., Wiggins, G. A., Potter, K., & Bhattacharya, J. (2019). Neural entrainment is associated with subjective groove and complexity for performed but not mechanical musical rhythms. *Experimental Brain Research, 237*(8), 1981-1991. https://doi.org/10.1007/s00221-019-05557-4

Colgin, L. L. (2013). Mechanisms and functions of theta rhythms. *Annual Review of Neuroscience, 36*, 295-312. https://doi.org/10.1146/annurev-neuro-062012-170330

David, J. A. d. O., Zefa, E., & Fontanetti, C. S. (2003). Cryptic species of *Gryllus* in the light of bioacoustic (Orthoptera: Gryllidae). *Neotropical Entomology, 32*, 75-80. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-566X2003000100010&nrm=iso

Davidson, S. M., & Wilkinson, G. S. (2002). Geographic and Individual Variation in Vocalizations by Male *Saccopteryx bilineata* (Chiroptera: Emballonuridae). *Journal of Mammalogy, 83*(2), 526-535. https://doi.org/10.1644/1545-1542(2002)083<0526:GAIVIV>2.0.CO;2

Doherty, J. A. (1985). Temperature Coupling and 'Trade-Off' Phenomena in the acoustic Communication System of the Cricket, *Gryllus Bimaculates* De Geer (Gryllidae). *The Journal of Experimental Biology, 114*, 17-35.

Duffy, S., & Pearce, M. (2018). What makes rhythms hard to perform? An investigation using Steve Reich's Clapping Music. *PLoS One, 13*(10), e0205847. https://doi.org/10.1371/journal.pone.0205847

Fernandez, A. A. (2020). *Vocal ontogenetic processes in bat pups, From babbling behavior to the interplay of social and vocal complexity* [Doctoral thesis, Freie Universität Berlin].

Filer, A., Burchardt, L. S., & van Rensburg, B. J. (accepted). Assessing acoustic competition between sibling frog species using rhythm analysis. *Ecology and Evolution*.

Fisher, R. A., & Yates, F. (1964). *Statistical Tables for Biological, Agricultural and Medical Research, Sixth Edition* (6th ed.). Oliver and Boyd.

Forkman, B., Boissy, A., Meunier-Salaün, M. C., Canali, E., & Jones, R. B. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior, 92*(3), 340-374. https://doi.org/10.1016/j.physbeh.2007.03.016

Gochfeld, M. (1978). Intraspecific Social Stimulation and Temporal Displacement of Songs of the Lesser Skylark, *Alauda gulgula*. *Zeitschrift für Tierpsychologie, 48*(4), 337-344. https://doi.org/10.1111/j.1439-0310.1978.tb00264.x

Greenfield, M. D. (1994). Synchronous and Alternating Choruses in Insects and Anurans: Common Mechanisms and Diverse Functions. *American Zoologist, 34*(6), 605-615. https://doi.org/10.1093/icb/34.6.605

Greenfield, M. D., & Schul, J. (2008). Mechanisms and evolution of synchronous chorusing: Emergent properties and adaptive functions in *Neoconocephalus katydids* (Orthoptera: Tettigoniidae). *Journal of Comparative Psychology, 122*(3), 289-297. https://doi.org/10.1037/0735-7036.122.3.289

Henry, L., Barbu, S., Lemasson, A., & Hausberger, M. (2015). Dialects in Animals: Evidence, Development and Potential Functions. *Animal Behavior and Cognition, 2*(2), 132-155. https://doi.org/10.12966/abc.05.03.2015

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*, 85-91. https://doi.org/10.1111/j.1749-6632.2011.06402.x

Honing, H., & Ploeger, A. (2012). Cognition and the evolution of music: pitfalls and prospects. *Topics in Cognitiv Sciences, 4*(4), 513-524. https://doi.org/10.1111/j.1756-8765.2012.01210.x

Hultsch, H., & Todt, D. (1982). Temporal performance roles during vocal interactions in nightingales (*Luscinia megarhynchos B.*). *Behavioral Ecology and Sociobiology, 11*(4), 253-260. https://doi.org/10.1007/BF00299302

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften*(93), 451-454.

Land, R., Burghard, A., & Kral, A. (2016). The contribution of inferior colliculus activity to the auditory brainstem response (ABR) in mice. *Hearing Research, 341*, 109-118. https://doi.org/10.1016/j.heares.2016.08.008

Lattenkamp, E. Z., Nagy, M., Drexl, M., Vernes, S. C., Wiegrebe, L., & Knornschild, M. (2021). Hearing sensitivity and amplitude coding in bats are differentially shaped by echolocation calls and social calls. *Proceedings of the Royal Society B, 288*(1942), 20202600. https://doi.org/10.1098/rspb.2020.2600

Linnenschmidt, M., & Wiegrebe, L. (2019). Ontogeny of auditory brainstem responses in the bat, *Phyllostomus discolour*. *Hearing Research, 373*, 85-95. https://doi.org/10.1016/j.heares.2018.12.010

Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B, 268*(1483), 2315-2324. https://doi.org/10.1098/rspb.2001.1773

Martin, L. M., Garcia-Rosales, F., Beetz, M. J., & Hechavarria, J. C. (2017). Processing of temporally patterned sounds in the auditory cortex of Seba's short-tailed bat, *Carollia perspicillata. European Journal of Neuroscience, 46*(8), 2365-2379. https://doi.org/10.1111/ejn.13702

Martínez-Rivera, C. C., & Gerhardt, H. C. (2008). Advertisement-call modification, male competition, and female preference in the bird-voiced treefrog *Hyla avivoca*. *Behavioral Ecology and Sociobiology, 63*(2), 195-208. https://doi.org/10.1007/s00265-008-0650-0

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals Voices. *Current Biology, 27*(15), 2352-2356.e2352. https://doi.org/10.1016/j.cub.2017.06.035

McRae, T. R. (2020). A review of squirrel alarm-calling behavior: What we know and what we do not know about how predator attributes affect alarm calls. *Animal Behavior and Cognition, 7*(2), 168-191. https://doi.org/10.26451/abc.07.02.11.2020

Moss, C. F., Bohn, K., Gilkenson, H., & Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *PLOS Biology, 4*(4), e79. https://doi.org/10.1371/journal.pbio.0040079

Moss, C. F., & Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *The Journal of the Acoustical Society of America, 110*(4), 2207-2226. https://doi.org/10.1121/1.1398051

Naguib, M. (1999). Effects of song overlapping and alternating on nocturnally singing nightingales. *Animal Behaviour, 58*(5), 1061-1067. https://doi.org/10.1006/anbe.1999.1223

Nelson, D. A. (2000). Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour, 60*(6), 887-898. https://doi.org/10.1006/anbe.2000.1560

Nelson, D. A., & Poesel, A. (2007). Segregation of information in a complex acoustic signal: individual and dialect identity in white-crowned sparrow song. *Animal Behaviour, 74*(4), 1073-1084. https://doi.org/10.1016/j.anbehav.2007.01.018

Nityananda, V., & Balakrishnan, R. (2006). Synchrony during acoustic interactions in the bushcricket Mecopoda 'Chirper' (Tettigoniidae: Orthoptera) is generated by a combination of chirp-by-chirp resetting and change in intrinsic chirp rate. *Journal of Comparative Physiology A, 193*(1), 51. https://doi.org/10.1007/s00359-006-0170-1

Norton, P., Barschke, P., Scharff, C., & Mendoza, E. (2019). Differential Song Deficits after Lentivirus-Mediated Knockdown of FoxP1, FoxP2, or FoxP4 in Area X of Juvenile Zebra Finches. *The Journal of Neuroscience, 39*(49), 9782-9796. https://doi.org/10.1523/jneurosci.1250-19.2019

Norton, P., & Scharff, C. (2016). "Bird Song Metronomics": Isochronous Organization of Zebra Finch Song Rhythm. *Frontiers in Neuroscience, 10*(309). https://doi.org/10.3389/fnins.2016.00309

Okobi, D. E., Banerjee, A., Matheson, A. M. M., Phelps, S. M., & Long, M. A. (2019). Motor cortical control of vocal interaction in neotropical singing mice. *Science, 363*(6430), 983. https://doi.org/10.1126/science.aau9480

Patel, A. D. (2008). *Music, Language and the Brain*. NY: Oxford University Press.

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B, 285*, 20180598. https://doi.org/10.1098/rspb.2018.0598

Pouw, W., Proksch, S., Drijvers, L., Gamba, M., Holler, J., Kello, C. T., Schaefer, R. S., & Wiggins, G. A. (2021). Multilevel rhythms in multimodal communication. *Proceedings of the Royal Society B*. https://doi.org/10.1098/rstb.2020.0334

Prat, Y., Azoulay, L., Dor, R., & Yovel, Y. (2017). Crowd vocal learning induces vocal dialects in bats: Playback of conspecifics shapes fundamental frequency usage by pups. *PLOS Biology, 15*(10), e2002556. https://doi.org/10.1371/journal.pbio.2002556

Ratcliffe, J. M., Elemans, C. P., Jakobsen, L., & Surlykke, A. (2013). How the bat got its buzz. *Biology Letters, 9*(2). https://doi.org/10.1098/rsbl.2012.1031

Ravignani, A. (2018). Spontaneous rhythms in a harbor seal pup calls. *BMC research notes, 11*(1). https://doi.org/10.1186/s13104-017-3107-6

Ravignani, A. (in press). Isochrony, vocal learning and the acquisition of rhythm and melody. *Behavioral and Brain Science, in press*.

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology, 5*, 1118. https://doi.org/10.3389/fpsyg.2014.01118

Ravignani, A., & Norton, P. (2017). Measuring ryhtmic complexity: A primer to quantify and compare temporal structure in speech, movement and animal vocalizations. *Journal of Language Evolution*. https://doi.org/10.1093/jole/lzx002

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., & Galambos, R. (1981). Auditory brainstem response in dolphins. *Proceedings of the National Academy of Science, 78*(3). https://doi.org/10.1073/pnas.78.3.1943

Rohatgi, V. K., & Saleh, A. K. M. E. (2015). Statistical Tables. In V. K. Rohatgi & A. K. M. E. Saleh (Eds.), *An Introduction to Probability and Statistics* (pp. 647-665). https://doi.org/https://doi.org/10.1002/9781118799635.oth1

Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One, 3*(1), e1461. https://doi.org/10.1371/journal.pone.0001461

Sasahara, K., Tchernichovski, O., Takahasi, M., Suzuki, K., & Okanoya, K. (2015). A rhythm landscape approach to the developmental dynamics of birdsong. *Journal of the Royal Society Interface, 12*(112). https://doi.org/10.1098/rsif.2015.0802

Tárano, Z., & Carballo, L. (2016). Call intercalation in dyadic interactions in natural choruses of Johnstone's whistling frog *Eleutherodactylus johnstonei* (Anura: Eleutherodactylidae). *Behavioural Processes, 126*, 55-63. https://doi.org/10.1016/j.beproc.2016.03.004

Wetekam, J., Reissig, C., Hechavarria, J. C., & Kössl, M. (2020). Auditory brainstem responses in the bat *Carollia perspicillata*: threshold calculation and relation to audiograms based on otoacoustic emission measurement. *Journal of Comparative Physiology A, 206*(1), 95-101. https://doi.org/10.1007/s00359-019-01394-6

Wiegrebe, L., & Schmidt, S. (1996). Temporal integration in the echolocating bat, *Megaderma lyra*. *Hearing Research, 102*, 35-42. https://doi.org/10.1016/s0378-5955(96)00139-6

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and Language, 144*, 26-34. https://doi.org/10.1016/j.bandl.2015.03.008

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., & Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron, 104*(1), 87-99. https://doi.org/10.1016/j.neuron.2019.09.036

Yartsev, M. M., & Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. *Science, 340*(6130), 367-372. https://doi.org/10.1126/science.1235338

Yoshida, S., & Okanoya, K. (2005). Evolution of Turn-Taking: A Bio-Cognitive Perspective. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 12*(3), 153-165. https://doi.org/10.11225/jcss.12.153

Eidesstaatliche Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertation gemäß §7 Abs. 4 der Promotionsordnung des Fachbereichs Biologie, Chemie, Pharmazie der Freien Universität Berlin vom 31. Mai 2018 eine eigenständig verfasste Forschungsleistung ist und ich keine anderen als die angegebenen Hilfsmittel benutzt habe. Diese Arbeit ist weder in dieser noch in ähnlicher Form einem anderen Promotionsausschuss vorgelegt worden.

Ort, Datum	Lara Sophie Burchardt