Vocal ontogenetic processes in bat pups
From babbling behavior to the interplay of social and vocal complexity

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin
by

Ahana Aurora Fernandez

Berlin, January 2020

Die vorliegende Doktorarbeit wurde von Mai 2015 bis Januar 2020 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.

1. Gutachterin:	PD. Dr. Mirjam Knörnschild
2. Gutachterin:	Prof. Constance Scharff, Ph.D.

Disputation am 22. April 2020

Danksagung

Viele Menschen haben dazu beigetragen, dass diese Arbeit entstanden ist. Einigen mag vielleicht nicht einmal bewusst sein, wie wertvoll sie für mich sind und wie ihre Unterstützung, - in ganz unterschiedlichen Formen - dazu geführt hat, dass ich diese Doktorarbeit verwirklichen konnte. Deswegen an alle ein riesiges Dankeschön!

Meiner Doktormutter Mirjam Knörnschild gebührt riesiger Dank dafür, dass ich bei ihr lernen durfte, was gute Wissenschaft ausmacht und wie man zu einer guten Wissenschaftlerin wird. Ich habe von ihr gelernt, wie wichtig gute Fragen sind und dass gute Antworten auch mit Weitblick und Offenheit zu tun haben! Und wie spannend es sein kann, wenn man plötzlich „das ganze Bild“ sieht – und wie unfassbar toll es ist, wenn dadurch immer neue Fragen und Ideen auftauchen. Sie hat mich immer unterstützt, mich angeleitet und vor allem ermutigt und bestärkt – und mir auch Narrenfreiheit gewährt wenn ich selber Ideen hatte und Dinge auf eigene Faust ausprobieren wollte. Eine besser Mischung kann man sich nicht wünschen!

Ebenso möchte ich mich bei meiner Zweitgutachterin Constance Scharff herzlich bedanken. Ich habe viel von ihr gelernt in den wöchentlichen Gruppenmeetings, habe mich immer ernst genommen und unterstützt gefühlt (auch oft hinter den Kulissen des Geschehens) und sie ist mir ein tolles Vorbild dafür, wie Wissenschaft durch Frauen geprägt werden kann und sollte!

Ganz wichtig waren während meiner Doktorarbeitszeit auch alle Menschen meiner Arbeitsgruppe. Liebe Martina, ich danke dir für deine immerwährende Unterstützung, dein Wissen, deinen Humor und vor allem deine Fähigkeit, Dinge (egal was) aus einem komplett anderen Blickwinkel zu betrachten! Ich habe so viel von dir gelernt! Ein besonderer Dank geht an Lara, die nicht nur Arbeitskollegin ist (und mich dazu angestachelt hat viel besser in Statistik und R zu werden), sondern vor allem eine Freundin! Danke für deine stete Unterstützung und Freundschaft. Danke Anja und Sophie, dass ihr meinen Büroalltag lustiger macht und wir auf Konferenzen weinselige Abende zusammen verbringen können! Linus, mit dir hatte ich tolle Gespräche über Gott und die Welt – egal ob beim joggen, Bier trinken oder im Büro. Danke dafür. Special thanks also to Hans who came here for an internship and instantly fascinated me with his deep knowledge about - and passion for vocal ontogeny. It´s great, Hans, having you as my friend and colleague and thanks for the awesome long conversations and discussions about vocal ontogenetic processes and your support with my English texts!

Danke auch an alle Mitglieder der FU-Arbeitsgruppe, vor allem Stefan und Jon die meine Anfangszeit in Berlin erheblich verschönert haben, weil sie mir unter anderem beibrachten, wo und wie man in Berlin feiern geht!

Lieber Stefan, lieber Thomas, tausend Dank, dass ich meine ersten Fledermausfangnächte mit euch verbringen und mein „Fanghandwerk" bei den Besten lernen durfte! Ihr wart immer geduldig und habt euer enormes Wissen über all die tropischen Fledermäuse gerne mit mir geteilt. Lieber Stefan – was hätte ich in meiner ersten BCI Zeit ohne dich getan! Danke für deine immerwährende Freundschaft und vor allem für unsere großartigen „Babbelstunden". Liebe Gloria – wie schön, dass wir beide uns auch in der ersten BCI Zeit kennen lernen durften. Unvergessen sind unsere Gespräche über Fledermäuse und Avisoft Funktionen, aber auch über die anderen wichtigen Dinge des Lebens auf dem Balkon von BCI – oder in der Rooftopbar in Casco Viejo. Und tausend Dank für die wundervolle Sacco-Zeichnung die ich als Titelbild meiner Dissertation verwenden durfte. Es waren noch unzählige andere Menschen mit mir im Feld und haben mir unvergessliche Momente geschenkt; danke an Matt, Luis, Darjia, David, Leonie, Inga und Gregg. Querida Jineth, nunca me imaginaba que encontraría una alma gemela en Gamboa! Gracias, por soportarme, por las conversaciónes sobre los murcis y ciencia en general y por compartir tu vida conmigo. Y por las noches en la Rana dorada.

Liebe Linda, Du bist für mich in Berlin ein wunderbarer Ausgleich zu meinem Büroalltag und hast mich mit deiner Begeisterung für AcroYoga, Bouldern und Klettern angesteckt. Dank dir habe ich so oft meinen Kopf „freibouldern" und „freiturnen" können – danke für deine Freundschaft!

A mini geliebti Schwizerfraktion: Obwou ig doch Wauberlinerin bi blibeni im Härzä immer äs Bärner Meitschi. Auso: Liebi Lea, ig danke dir us ganzem Härzä für dini Fründschaft wo scho eewig währt. Immer bringsch Du mi wider zum lachä – mit dine Gschichte us dim spannende Autag und mit dim Optimismus! Liebs Pili, tuusig dank dass Du immer für mi da bisch, mir zämä di beschtä Wiiäbätä hei (mit oder ohni Elias), di legendärschte Usgäng und sogar es Zimmer mit Styroporwand chöi teilä! Liebi Corä, immer bisch du da gsi und hesch mi begleitet uf mim Wäg vom Schadoupark bis nach Berlin…danke, dass Du immer a mi gloubsch und immer da bisch mit Rat und Tat! Michu – abgseh dervo dasmer eifach legendär si egau wasmer zämä machä – findis dr Hammer, dass üsi Fründschaft scho so lang aduuert und mir üs trotz Entfärnig immer wider gseh und di entspannteschte Äbätä mitenand verbringä woni aubes Buchweh ha vor Lachä. Simu, du hesch mi so oft grettet mit R und genereu dim enormä Wüssä über Computer und vorauem übers Programmiere. Danke! Ä unfassbari Unterstützig für mi isch üsi Fründschaft und ou di Optimismus womi immer wider asteckt – sogar weni muess versuehce uf dinä Schultere z balanciere ufem Tämpelhofer Fäud! Maria, dankä dass Du immer wider uf Berlin chunnsch und mit mir hie unvergässlechi Wuchenändi verbringsch wosech afühele wi mindeschtens ei Monet ferie, gracias Linda! Märu auts Huus – unvergässä si üsi Äbätä in Costa Rica ir Hängematte mit em einzig wahre Rum – Voumond – philospophische Gspräch übers Läbä oder üsne legendäre Baukon-tanz-Verastautige. Dankä dassi immer eifach bi dir darf uftouchä und du mi so nimsch und akzeptiersch wini bi. Nicolas, du halbe Schuhe – je te dois beaucoup, ton enthousiasme constant pour la science et surtout pour les chauves-souris m´a contaminé. Je te remercie de m´avoir permis de faire mes premiers pas dans la science avec toi.

Liebe Ma, querido Pa, euch verdanke ich so viel! Ustedes siempre me apoyan y creen en mi camino. Und habt euch während meiner Doktorarbeitszeit von der Begeisterung für Fledermäuse anstecken lassen. Mil gracias por siempre estar a mi lado y ofrecer la posibilidad de seguir mi sueño mas grande: estudiar y comprender la comunicación social de los animales.

Lieber Anand, egal wo Du gerade steckst Du bist der beste kleine große Bruder der Welt und hast mich zu dem Menschen gemacht, der ich heute bin!

Hal aunt, I owe you so much! Thanks for always being supportive, believing in the stuff I am doing, providing me with the world best sphags and evenings full of funny stories. I am so glad that you accompanied my way – from history lessons as an undergrad until now!

Liebster Simon, Du hast mit deinem steten Optimismus, deiner Fröhlichkeit und deinem Wissen so vieles möglich und vor allem einfacher gemacht! Danke, dass Du mich immer zum lachen bringst und stets an mich glaubst, auch und vor allem wenn ich das selbst nicht mehr tue. Schön bist Du da!

Für meinen großen kleinen Bruder. In Liebe.

Index

Publication list and author contributions

Chapter IV and Chapter V are published in peer-revised journals (see DOI) and are presented as published. Chapters I, II and III are in manuscript form.

Chapter I: Currently in rebuttal status at the journal *Proceedings of the National Academy of Sciences of the United States* (PNAS).

Fernandez*, AA., Burchardt, LS. & Knörnschild M.

Babbling in bat pups and human infants: common ontogenetic features of vocal production learners

Author contributions:

AAF and MK designed the study. AAF collected the data. AAF analyzed all data with the exception of the rhythm analysis conducted by LB. AAF and MK wrote the manuscript and revised it.

Chapter II

Fernandez* AA., Serve N., Fabian S-C., & Knörnschild M.

Social influences shape the babbling behavior of greater sac-winged bat pups, *Saccopteryx bilineata*

Author contributions:

AAF designed the study. AAF and NS collected the data. AAF analyzed the data, FSC participated in data analysis. AAF and MK wrote the manuscript and revised it.

Chapter III: in preparation for the journal *Frontiers in Ecology and Evolution*.

Fernandez* AA. & Knörnschild M.

Pup-directed vocalizations of adult females and males in a vocal learning bat

Author contributions:

AAF and MK designed the study. AAF collected the data. AAF analyzed the data. AAF and MK wrote the manuscript and revised it.

Chapter IV: published

Fernandez*, A. A., & Knörnschild, M. (2017).

Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages.

Animal Behavior and Cognition, *4*, 169-186. https://doi.org/10.12966/abc.04.05.2017

Author contributions:

AAF and MK designed the study. AAF collected and analyzed the data. AAF and MK wrote and revised the manuscript.

Chapter V: published

Knörnschild*, M., **Fernandez, A. A.**, & Nagy, M. (2019).

Vocal information and the navigation of social decisions in bats: Is social complexity linked to vocal complexity?

Functional Ecology. https ://doi.org/10.1111/1365-2435.13407

Author contributions:

MK and MN designed the study. AAF collected the data. MK, MN and AAF analyzed the data. MK and MN wrote the manuscript. MK, MN and AAF revised the manuscript.

Notes:

AAF	Ahana Aurora Fernandez
MK	Mirjam Knörnschild
LB	Lara Sophie Burchardt
MN	Martina Nagy
NS	Nora Serve
FSC	Sarah-Cecil Fabian

*Corresponding author

Summary

The human language faculty is a complex system involving key components such as speech, syntax and semantics. Speech is the motor vocal output of language and our main communication modality in daily life. Speech acquisition requires the ability of vocal production learning and is characterized by a pronounced vocal practice behavior called canonical babbling. During canonical babbling the infant acquires the basic speech subunits – usually consonant-vowel pairs – by matching the own oral output to the auditory input. Hence, canonical babbling is a prerequisite to acquire the phonetic repertoire of language. So far, babbling behavior appears to be a relatively rare trait in the animal kingdom and even rarer among mammals.

In the thesis presented here, I investigated a conspicuous vocal practice behavior, called babbling, in a vocal production learner, the bat species *Saccopteryx bilineata.* The main finding of my dissertation shows that *S. bilineata* pups exhibit a babbling behavior which is characterized by the same features that define human infant canonical babbling. During their ontogeny, pups engaged in a conspicuous eight-week long phase of vocal practice behavior that is highly reminiscent of infant canonical babbling. During babbling, pups acquire a part of their adult vocal repertoire through vocal production learning. The analyses of this babbling behavior demonstrated that features defining canonical babbling in infants are likewise characteristic of pup babbling in *Saccopteryx bilineata*, including early babbling onset (I), babbling sequences composed simultaneously of undifferentiated proto-syllables (i.e. non-adult syllable types) and adult-like ones (II), vocal overproduction (III), subset acquisition of the adult syllable repertoire (IV), a non-linear repertoire acquisition pattern (V), repetitiveness (VI), rhythmicity (VII), meaninglessness (VIII) and universality (IX). These findings suggest that similar features characterize a distinctive vocal practice phase across vocal production learners, probably arising from similar underlying mechanisms geared toward acquiring adult communication.

In human infants, canonical babbling is shaped by vocal and/or behavioral feedback provided by the social environment. In *S. bilineata,* the pups´ social environment (i.e. the colony) is composed of their mothers, other females with their pups and adult males. The babbling behavior of pups was accompanied by behaviors and interactions restricted to mother-pup dyads whereas other adult individuals do not interact with pups. Adult males produce two song types and provide the main auditory input for pups. I investigated the influence of the social environment on three main factors of the pups` babbling behavior, namely the amount of vocal practice, the babbling bout composition (i.e. the syllabic composition of babbling) and

the pups´ final syllable repertoire size. The results showed that the maternal behavior influences the amount of vocal practice, both on a daily level (babbling bout duration) and over the entire ontogeny (number of days a pup engaged in babbling during ontogeny). Both the mothers´ behavior and the number of singing males present in a colony influenced the babbling bout composition (i.e. the syllabic composition of vocal practice), especially the production of syllable sequences that were mainly composed of song syllables acquired through vocal production learning. On the contrary, the pups' final syllable repertoire size was not affected by the social environment. This study demonstrates that, similar to infant speech development, social interactions impact the developmental processes during a pronounced vocal ontogeny phase.

Pup-directed vocalizations during vocal ontogeny are common and usually female-biased. Females of *S. bilineata* produce one pup-directed vocalization, the maternal directive call. Production of this call was restricted to the babbling phase. Other vocalizations produced by females are directed at adults. In humans, infant-directed speech – also known as "motherese" – supports linguistic learning and has social benefits. Infant-directed speech is characterized by universal features such as increased pitch, expanded intonation contours and a slower cadence. Moreover, recent research has shown that the timbre (i.e. the spectral quality of sound) of motherese is distinctly different than the timbre of adult-directed speech in humans. This led to the question if the timbre of pup-directed female vocalizations differed from the timbre of adult-directed vocalizations. My results demonstrated that the timbre of the pup-directed female vocalization was significantly different from the timbre of adult-directed female vocalizations. This thesis thus includes the first description of a phenomenon reminiscent of motherese in bats. In the course of investigating the pups` vocal ontogeny, I observed an unusual male acoustic behavior directed at pups: Adult males produced isolation calls, a call which is otherwise only emitted by pups. This multisyllabic call is usually produced to solicit maternal care and encodes information about individual identity and social group affiliation (i.e. vocal signatures). Until now, it was assumed that the group signature develops through call convergence of pups from the same colony towards a social group (i.e. colony) mean. The observation of adult male isolation call production raised the question of whether males may serve as tutors for the vocal group signature. The observations revealed that adult males produce isolation calls in reaction to pup isolation call production; in most cases male isolation call production succeeded pup isolation call bouts (i.e. a series of isolation calls). My acoustic analysis demonstrated that the isolation calls of pups are acoustically more similar to those of males from their own social group than to those of males from other social groups. These findings allow the speculation that the group signature in isolation calls could be transmitted horizontally across generations.

The aforementioned pup isolation call, which is produced throughout the entire vocal ontogeny, is composed of different syllable types (i.e. composite- and short stereotyped syllables). As aforementioned, this call encodes vocal signatures on individual identity and social group affiliation. The possibility to encode information in a vocalization is limited by the available acoustic space (i.e. the information in bit which can be encoded in an acoustic signal). Therefore, vocalizations that are composed of multiple elements have the potential to distribute information among those elements and hence increase the amount of encoded information. The multiple messages hypothesis states that different elements of a vocalization can convey different information. In my study, I investigated if the multisyllabic pup isolation call encoded further information and if so – how the information content was distributed over the different syllables of the isolation call. My results demonstrated that the individual signature was encoded in the composite syllable (corroborating the findings of former studies) and included the novel finding that acoustic features of the short stereotyped syllables significantly changed during ontogeny, providing information about pup age. Hence, the multisyllabic isolation call encodes multiple messages which are temporally segregated.

The "social complexity hypothesis for communication" states that increasing social complexity should lead to increased vocal complexity, ultimately resulting in a positive feedback loop. An example of a measurable feature of vocal complexity is the amount of information encoded in specific vocalizations (e.g. vocal signatures). Bats are highly gregarious animals, exhibiting a variety of social systems and display a similar diversity concerning their vocal repertoires. Bats' social vocalizations can encode a variety of information, amongst which information about identity and social group affiliation (i.e. vocal signatures) is crucial in social interactions. Individual signatures in social vocalizations are the best-studied vocal signature type in bats. Considering the aforementioned arguments, bats provide an ideal study system to investigate the relation between social and vocal complexity. To investigate a potential positive feedback loop between social and vocal complexity in bats, I analyzed how the information content of social vocalizations encoding an individual signature was related to the complexity of the social group. A phylogenetic comparative analysis across species revealed that increased social complexity (i.e. relevant social group size) was correlated with an increase in information content encoded in social vocalizations.

Zusammenfassung

Die menschliche Sprache ist ein komplexes System welches sich aus Schlüsselkomponenten wie der Fähigkeit zur Lautproduktion, Syntax und Semantik zusammensetzt. Gesprochene Sprache ist unsere am häufigsten genutzte Kommunikationsmodalität im täglichen Leben. Der Spracherwerb erfordert die Fähigkeit des vokalen Produktionslernens und zeichnet sich durch ein ausgeprägtes Übungsverhalten aus, das kanonisches Lallen oder Babbeln genannt wird. Während des kanonischen Lallens erwirbt das Kind die grundlegendsten Lauteinheiten der gesprochenen Sprache – üblicherweise Silben - indem es die eigene Lautproduktion mit dem Höreindruck abgleicht. Kanonisches Lallen ist daher eine Voraussetzung für den Erwerb des phonetischen Sprachrepertoires. Bislang scheint das Lall-Verhalten ein relativ seltenes Merkmal im Tierreich und noch seltener bei Säugetieren zu sein.

In der hier vorgestellten Doktorarbeit untersuchte ich ein auffälliges Übungsverhalten eines vokalen Produktionslerners der Fledermausart *Saccopteryx bilineata*. Vokales Produktionslernen beinhaltet das Neuerlernen eines akustischen Signals oder die Modifikation seiner Struktur. Das Hauptergebnis meiner Dissertation zeigt, dass Fledermausjungtiere von *S. bilineata* ein Übungsverhalten zeigen, das sich durch dieselben Merkmale auszeichnet, welche das kanonische Lallen des menschlichen Kleinkindes definieren. Während ihrer Ontogenie durchlaufen die Jungtiere eine achtwöchige Phase auffälligen Übungsverhaltens, die stark an das kanonische Lallen von Kleinkindern erinnert. Während dieser sogenannten Babbelphase erwerben die Jungtiere einen Teil des adulten Lautrepertoires durch vokale Imitation. Meine Analysen dieses Babbelverhaltens zeigten, dass Merkmale, die das kanonische Lallen bei Kleinkindern definieren, ebenfalls für das Babbeln von *S. bilineata* Jungtiere charakteristisch sind: der frühe Beginn des Babbelverhaltens während der Ontogenie (I), Babbelsequenzen, die gleichzeitig aus sogenannten undifferenzierten so genannten Proto-Silben bestehen (d.h. ein Silbentyp der nicht zum adulten Repertoire gehört) und solchen, die als Vorläufer von Silben des adulten Lautrepertoires erkennbar sind (II), vokale Überproduktion (III), Teilerwerb des adulten Lautrepertoires (IV), ein nichtlinearer Erwerb des Silbenrepertoires (V), Reduplikation (VI), rhythmische Lautproduktion (VII), Bedeutungslosigkeit (VIII) und Universalität (IX). Meine Ergebnisse zeigen, dass die Übungsphase bei menschlichen und nichtmenschlichen vokalen Produktionslernern durch dieselben Merkmale charakterisiert ist, was wahrscheinlich auf ähnliche zugrundeliegende Mechanismen zurückzuführen ist, die nötig sind um komplexe vokale Kommunikation zu erwerben.

Bei menschlichen Kleinkindern wird das kanonische Lallen durch stimmliches und/oder verhaltensbezogenes Feedback aus dem sozialen Umfeld geprägt. Bei *S. bilineata* setzt sich das soziale Umfeld der Jungtiere (d.h. die Fledermauskolonie) aus der Mutter, anderen Weibchen mit ihren Jungtieren und adulten Männchen zusammen. Das Babbelverhalten der Jungtiere wurde von Verhaltensweisen und Interaktionen begleitet, die ausschließlich zwischen Mutter und Jungtier stattfinden, während andere adulte Individuen nicht mit den Jungtieren interagierten. Die adulten Männchen produzieren zwei Gesangstypen und liefern somit den wichtigsten auditiven Input für die Jungtiere. Ich untersuchte den Einfluss des sozialen Umfelds auf drei Hauptfaktoren des Babbelverhaltens der Jungtiere, nämlich den Zeitaufwand der für das Babbeln aufgewendet wird, die Silbenzusammensetzung des Babbelns, und den Umfang des Lautrepertoires, welches sich ein Jungtier am Ende seiner Babbelphase angeeignet hat. Die Ergebnisse zeigten, dass das Verhalten der Mutter den Zeitaufwand beeinflusst, der auf Babbeln verwendet wird, sowohl täglich (Dauer einer kompletten Babbelsequenz) als auch über die gesamte Ontogenie (Anzahl der Tage, die ein Jungtier während der Ontogenie babbelt). Sowohl das Verhalten der Mütter als auch die Anzahl der singenden Männchen in einer Kolonie beeinflussten die Silbenzusammensetzung des Babbelns, insbesondere die Produktion von Silbenabfolgen, die hauptsächlich aus den Silben zusammengesetzt sind, welche durch Imitation erlernt werden. Der Umfang des Silbenrepertoires der Jungtiere wurde hingegen nicht durch das soziale Umfeld beeinflusst. Diese Studie zeigt, dass ähnlich wie bei der kindlichen Sprachentwicklung soziale Interaktionen die Entwicklungsprozesse während des Babbelns beeinflussen.

An Jungtieregerichtete Vokalisationen während der vokalen Ontogenie sind häufig und werden meist von den Müttern produziert. *Saccopteryx bilineata* Weibchen produzieren während der Babbelphase einen sogenannten Richtruf (directive call), der an ihr jeweiliges Jungtier adressiert ist. Andere von Weibchen produzierte Vokalisationen richten sich an adulte Artgenossen. Beim Menschen unterstützt die sogenannte Kleinkindersprache (motherese) das sprachliche Lernen und ist außerdem ein positives soziales Feedback für das Kleinkind. Die Kleinkindersprache zeichnet sich durch universelle Eigenschaften wie erhöhte Tonhöhe, erweiterte Intonationskonturen und eine langsamere Kadenz aus. Darüber hinaus hat neuste Forschung gezeigt, dass sich die Klangfarbe der Kleinkindersprache im Vergleich zur Klangfarbe der an Erwachsene gerichteten Sprache beim Menschen unterscheidet. Dies führte zu der Frage, ob sich die Klangfarbe von Weibchen unterscheidet, wenn sie mit ihren Jungtieren oder adulten Artgenossen kommunizieren. Meine Ergebnisse zeigten, dass sich die Klangfarbe zwischen beiden Adressaten signifikant unterschied. Diese Doktorarbeit beinhaltet somit die

erste Beschreibung eines Phänomens, das an die Kleinkindersprache beim Menschen erinnert. Bei der Untersuchung der vokalen Ontogenie der Jungtiere beobachtete ich ein ungewöhnliches akustisches Verhalten der adulten Männchen, das sich an die Jungtiere richtete. Die Männchen erzeugten Isolationsrufe (d.h. Bettelrufe), die sonst nur von den Jungtieren produziert werden. Dieser multisyllabische Ruf wird in der Regel beim Betteln um mütterliche Fürsorge erzeugt und kodiert die individuelle Identität und soziale Gruppenzugehörigkeit (d.h. vokale Signaturen). Bisher wurde angenommen, dass sich während der Ontogenie die Gruppensignatur durch Rufkonvergenz der Jungtiere aus derselben Kolonie entwickelt. Da auch erwachsene Männchen Isolationsrufe produzierten und sie direkt an Jungtiere adressierten, ging ich der Frage nach, ob die Männchen die Rufkonvergenz zu einer Gruppensignatur beeinflussen könnten. Meine Beobachtungen ergaben, dass adulte Männchen Isolationsrufe als Reaktion auf die Produktion von Jungtier-Isolationsrufen produzieren. Meine akustische Analyse zeigte, dass die Isolationsrufe von Jungtieren denen der Männchen ihrer eigenen sozialen Gruppe akustisch ähnlicher sind als denen der Männchen anderer sozialer Gruppen. Diese Ergebnisse lassen Spekulationen über die Möglichkeit einer generationenübergreifenden Übertragung einer Gruppensignatur zu.

Der erwähnte Jungtier-Isolationsruf, der während der gesamten Ontogenie erzeugt wird, setzt sich aus verschiedenen Silbentypen zusammen (d.h. zusammengesetzte Silben und kurze, stereotype Silben). Wie bereits erwähnt, kodiert dieser Ruf die individuelle Identität und soziale Gruppenzugehörigkeit. Die Möglichkeit, Informationen in einer Vokalisierung zu kodieren, ist durch den verfügbaren akustischen Raum begrenzt. Daher haben Vokalisationen, die aus mehreren Elementen zusammengesetzt sind, das Potential, unterschiedliche Informationen auf diese Elemente zu verteilen und somit die Menge an kodierter Information zu erhöhen. Die „Multiple-messages Hypothesis“ besagt, dass verschiedene Elemente einer Vokalisation unterschiedliche Informationen vermitteln können welche zeitlich getrennt sind. In meiner Studie habe ich untersucht, ob der mehrsilbige Jungtier-Isolationsruf unterschiedliche Informationen kodiert und - falls ja - wie der Informationsgehalt auf die verschiedenen Silbentypen verteilt wurde. Meine Ergebnisse zeigten, dass die individuelle Signatur in der zusammengesetzten Silbe kodiert war (was die Ergebnisse früherer Studien bestätigte). Eine neue Erkenntnis ist, dass sich die akustischen Merkmale der kurzen stereotypen Silben während der Ontogenie signifikant veränderten, was Aufschluss über das Alter der Jungtiere gibt. Somit kodiert der multisilbische Isolationsruf unterschiedlichen Informationsgehalt, der zeitlich getrennt ist.

Die " social complexity hypothesis for communication " besagt, dass zunehmende soziale Komplexität zu einer erhöhten vokalen Komplexität führen sollte, was letztendlich zu einer positiven Rückkopplungsschleife führt. Ein Beispiel für ein messbares Merkmal vokaler Komplexität ist die Menge an Informationen, die in bestimmten Vokalisationen kodiert sind. Fledermäuse sind sehr gesellige Tiere, die eine Vielfalt an sozialen Systemen aufweisen und eine ähnliche Vielfalt an sozialen Lautrepertoires aufweisen. Individuelle Signaturen in Vokalisationen sind der am besten untersuchte Signaturtyp bei Fledermäusen. Unter Berücksichtigung der oben genannten Argumente bieten Fledermäuse ein ideales Studiensystem, um die Beziehung zwischen sozialer und vokaler Komplexität zu untersuchen. Um eine mögliche positive Rückkopplungsschleife zwischen sozialer und vokaler Komplexität bei Fledermäusen zu untersuchen, habe ich analysiert, wie der Informationsgehalt von sozialen Vokalisationen, die eine individuelle Signatur kodieren, mit der Komplexität der sozialen Gruppe zusammenhängt. Eine phylogenetisch vergleichende Analyse über verschiedenen Arten hinweg zeigte, dass eine erhöhte soziale Komplexität (d.h. Gruppengröße) mit einer Zunahme der individuellen Signaturstärke in Vokalisationen korreliert.

General Introduction

More than anything else, language defines human nature. Language, in all its different manifestations – spoken, signed or written – is a powerful communication tool allowing us to share information. The power of this communication system derives from the fact that we can turn thoughts into speech and share them with others. Our technological advances, innovations and our culture are based on the accumulation of knowledge which is transmitted mainly through the speech modality. Our maternal language and/or dialect also strongly influences the way we consider and perceive the world (Deutscher, 2010). Without language, our daily life would simply be unimaginable. That, and the fact that the faculty of language appears to be unique in the animal kingdom, makes the question of the evolution of language into one of science's most challenging and intriguing (Fitch, 2010; Hauser, Chomsky, & Fitch, 2002). Uncovering the evolutionary pathway of language is challenging because language does not fossilize. A promising approach to tackle the question of the origins of language is to investigate its biological foundations (Fitch, Hauser, & Chomsky, 2005). This so called biolinguistic approach investigates the mechanisms and cognitive abilities involved in the faculty of language in a variety of fields such as neurobiology, genetics and cognitive psychology (Martins & Boeckx, 2016). The multidisciplinary biolinguistics field investigates if and to what extend key elements of the language faculty evolved in non-human animals. This approach is promising because the faculty of language is a complex system composed of different elements and requires a multifaceted set of cognitive abilities such as vocal learning, auditory processing and memory, many of which are shared with non-human animals (Fitch, 2010; Hauser, Chomsky, & Fitch, 2002). Three key elements of language are speech, syntax and semantics (Fitch, 2010; Hauser, Chomsky, & Fitch, 2002; Margoliash & Nusbaum, 2009). Syntax is a set of rules which governs the composition of meaningless sounds into meaningful units (i.e. phonological syntax) and the production of recursive hierarchical structures (i.e. lexical syntax) (Margoliash & Nusbaum, 2009). Recursion allows unlimited expression based on a limited set of elements (Hauser, Chomsky, & Fitch, 2002). Semantics refers to the meaning encoded in speech sounds, and the ability to intentionally transmit this information and alter the behavior of the receiver (e.g. Fitch, 2010; Seyfarth, & Cheney, 2003; Suzuki, 2016). Speech refers to the ability to produce language-relevant speech subunits (phonemes) acquired through vocal learning (Fitch, Hauser, & Chomsky, 2005).

Speech: the vocal motor output of language

Speech is the vocal motor output of language (Fitch, 2018). In humans, the faculty of speech is based on the ability of vocal production learning and relies mainly on precise neuronal control of the vocal apparatus. In birds and mammals, including humans, the sound production system is composed of the respiratory system (i.e. lungs, air sacs) and the phonatory system (Bradbury & Vehrencamp, 2011; Ghazanfar & Rendall, 2008; Janik & Slater, 2000). The phonatory system is usually comprised of the larynx, (or syrinx, in birds). The upper vocal tract (i.e. the oral, nasal, and sometimes pharyngeal cavities) can act like a filter, and in humans it plays a crucial role in speech production (Fitch, 2000, 2018; Janik & Slater, 2000). Each of these different systems controls different aspects of sound production (Janik & Slater, 2000). The respiratory system is responsible for temporal and amplitude changes, controlled by the respiratory muscles. The phonatory system is responsible for fundamental frequency changes and frequency modulations produced by muscular control of the larynx or syrinx (Janik & Slater, 2000). The filter system is responsible for formant-related changes, i.e. the relative energy distributions in an acoustic signal (Fitch, 2000).

A prerequisite for speech: vocal production learning

The capacity of vocal production learning is a crucial prerequisite for speech (Fitch, 2018). Vocal production learning is defined as a process where acoustic signals are modified as a result of social interactions (Boughman & Moss, 2003; Janik & Slater, 2000). Vocal learning as a cognitive mechanism also includes contextual learning (i.e. the context in which an existing signal is used is learned) which is further divided into usage learning and comprehension learning (Janik & Slater, 2000). Vocal comprehension learning is the association of an existing acoustic signal with a context, action or outcome. Usage learning is learned production of an innate acoustic signal in a novel behavioral context (Janik & Slater, 2000). Contextual learning is relatively widespread in the animal kingdom (Schusterman, 2008), but vocal production learning is comparatively rare. Apart from humans, vocal production learning has been documented in parrots (e.g. Pepperberg, 1994), hummingbirds (e.g. Araya-Salas & Wright, 2013), songbirds (e.g. Marler, 1970), elephants (e.g. Poole, Tyack, Stoeger-Horwath, & Watwood, 2005), cetaceans (Janik, 2014), pinnipeds (e.g. Schusterman, 2008) and bats (Knörnschild, 2014). Vocal learning is thus a multifaceted complex trait (Wirthlin et al., 2019). Vocal learning processes ranges from modification of innate vocalizations to learned acquisition and culminates in the imitation of other species' acoustic signals (such as ”Hoover” the captive harbor seal who could imitate the speech of his human caregiver) and the invention of completely new signals (e.g. humans invent new words) (Boughman & Moss, 2003; Janik & Slater, 2000; Reichmuth & Casey, 2014).

New learned vocalizations can be added to an individual's repertoire through a number of mechanisms (Boughman & Moss, 2003). Acquisition learning involves auditory input of a tutor and auditory feedback from the vocalizer's own acoustic output. It is often – but not always – age dependent and usually critically depends on social interactions (Boughman & Moss, 2003). Social modification is a learning mechanism during which innate pre-existing signals are modified and altered in response to social interactions (Boughman & Moss, 2003). Mostly, this form of vocal learning results in call convergence – i.e. increased acoustic similarity between social partners (e.g. Boughman, 1998). But it can also result in more distinct vocalizations when individual identity is important for recognition purposes (Janik & Slater, 1997).

Human infant speech acquisition

Human children start acquiring speech sounds in early infancy. Considering the task's complexity, it is astonishing how seemingly effortlessly they seem to learn (Stoll, Lieven, Winskel, & Padakannaya, 2013; Vihman, 2014). Language acquisition has to occur within a critical period (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Newport, 1990). This period lasts until early puberty, as tragic cases of social deprivation have demonstrated (Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974). Infant speech acquisition can be divided into different phases, mostly defined by the characteristics of the infants' oral output (for general overview see Oller, 1980; Vihman, 2014). A very conspicuous and well-known phase during speech acquisition is the canonical babbling phase.

Infant canonical babbling – a crucial step in speech development

The power of spoken language relates to the mechanism of infinite expression through use of finite means (Fitch, 2010). Speech is constituted of language specific phonetic repertoires (i.e. a finite set of basic elements), which are composed into syllables, which in turn are the basis of words (Vihman, 1996). Hence, the acquisition of the language- specific phonetic repertoire during the so called canonical babbling stage is the first milestone in infant speech development (Oller, 1980; Vihman, 1996, 2014). While babbling children imitate and rehearse the relevant phonetic units and aim to match their own oral output to the engram of their perceptual input (Oller, 1980; Vihman, 1996). During canonical babbling the child produces repetitive vocal sequences composed of fully resonant adult-like syllables (usually consonant-vowel pairs) –

the speech sound subunits which are the basis of most languages and so-called protophones (i.e. non-speech sounds) (Oller, 1980; Vihman, 2014).

Canonical babbling is stringently required for speech acquisition. This is supported by the – thankfully few – reported cases about socially deprived children and studies on deaf or tracheotomized children where infants never entered the babbling stage (Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974; Locke & Pearson, 2009). The fundamental importance of canonical babbling for speech acquisition is the reason why babbling is universal, irrespective of cultural background and maternal language. Canonical babbling is a well-studied phenomenon and despite the immense variety of languages with different phonetic features (e.g. tonal versus atonal languages, different grammars), the literature suggests that canonical babbling is characterized by a nine universal features (for references see: Oller, 1980; Oller & Eilers, 1982; Vihman, Ferguson, & Elbert, 1986):

- Feature I: Onset of canonical babbling is early in infancy, usually around 6-7 months of age.

- Feature II: Canonical babbling is composed of canonical syllables and protophones. Canonical syllables are characterized by adult-speech like temporal transition between vowels and consonants and the vowels are fully resonant. Protophones comprise all non-speech sounds.

- Feature III: Canonical babbling includes the phenomenon of vocal overproduction, i.e. the production of sounds which are not part of the adult phonetic repertoire (e.g. the protophones).

- Feature IV: During canonical babbling, children acquire and rehearse only a subset of the phonetic adult repertoire.

- Feature V: This phonetic subset acquisition during babbling does not follow a linear pattern.

- Feature VI: Repetitiveness is – together with rhythmicity – the most characteristic and recognizable feature of canonical babbling. During babbled sequences, infants repeat the same syllable type which results in the characteristic reduplicated speech sound sequences, such as /ma ma ma/ /da da da/.

- Feature VII: Rhythmicity results from the repetition of the same syllable. It is often coupled with rhythmical movements of the body, for example shaking a rattle or banging the hands on a table.

- Feature VIII: Canonical babbling is meaningless. Syllables are per definition meaningless units and only their composition to the higher level unit, the word, gives rise to meaning. Hence, babbling does not transmit any meaning in the linguistic sense.

- Feature IX: And, as aforementioned, babbling is universal irrespective of the infants' gender, cultural background or maternal language.

So far, songbirds have been the preferred model system for comparative research on human speech development (Doupe & Kuhl, 1999; Marler, 1970). The reason for this preference is that songbirds share some important traits with humans: they are vocal production learners, and they share functional similarities in the organization of the neuronal circuitry for auditory perception and vocal production (Doupe & Kuhl, 1999; Jarvis, 2004). However, the utility of this model taxon is limited by its phylogenetic distance from humans and, to a minor degree, the fact that song learning is male biased. But note that so far, female song has been understudied and future studies about song learning in female birds will most likely reveal similar developmental stages as described for male birds (Odom, Hall, Riebel, Omland, & Langmore, 2014). For these reasons, a mammalian model species would be a crucial step forward in elucidating shared neuronal underpinnings in vocal production learners with a pronounced vocal practice phase. So far, babbling in mammals has only been reported in the non-vocal learning primate, the pygmy marmoset (Elowson, Snowdon, & Lazaro-Perea, 1998) and the greater sac-winged bat, *S. bilineata* (Knörnschild, Behr, & von Helversen, 2006).

The influence of the social environment on vocal ontogeny

Speech acquisition in infants is strongly influenced by the social environment (Kuhl, 2003, 2007). Many studies confirmed that social interactions between the child and its closest social interaction partners – usually the parents – shape the entire course of speech development (Goldstein & Schwade, 2008, 2010). The social environment provides the infant with auditory input, for example when adults talk to each other. But what strongly impacts the child's vocal production is the contingent social feedback directed at the vocalizing child. Social feedback equally includes vocal (e.g. talking to infant) and behavioral reactions (e.g. touching the infant) from the caregivers directed at the child (Gleitman & Newport, 2002; Goldstein & Schwade, 2008, 2010; Kuhl, 2003). Most studies about social influences on speech acquisition focus on the canonical babbling stage. A result of social interactions during canonical babbling is that children significantly increase the production of advanced vocalizations and the frequency of occurrence of fully resonant syllables (i.e. canonical syllables) when parents provide contingent social feedback (Goldstein & Schwade, 2008). Non-verbal responses such as smiling, touching the infant or providing a toy has the same effect on babbling behavior (Goldstein, King, & West, 2003). Hence, social interactions, regardless of the sensory modality, significantly shape babbling characteristics (Goldstein & Schwade, 2010).

Studies about vocal ontogeny in non-human animals have corroborated the finding that social interactions play an important role in vocal learning (Beecher & Burt, 2004). Some songbird species only learn their song (or learn it better) if they can visually or acoustically interact with a tutor (e.g. Adret, 2004; Chen, Matheson, & Sakata, 2016; Payne, 1981). Furthermore, social interactions are not restricted to the tutor. Studies in birdsong learning demonstrate that behavioral interactions with females shape song learning performance significantly. For example in zebra finches, Carouso-Peck and colleagues found that females produce a specific behavioral display (which they also produce during courtship with adult males) related to the pupil's plastic song (Carouso-Peck & Goldstein, 2019). If this social feedback is contingent on the vocal output of the juvenile, the learning success increases (i.e. the overall song similarity to the tutor song is higher). Another well-known study on cowbirds demonstrated that the wing stroke given by females in response to specific vocal elements led to their retention whereas other elements were discarded (King, West, & Goldstein, 2005; Smith, King, & West, 2000; West & King, 1988). Thus, females shaped the later song of adult male birds through behavioral displays. That social factors clearly affect the vocal development was also demonstrated in mammals, both vocal learning (e.g. Prat, Azoulay, Dor, & Yovel, 2017) and non-learning (e.g. Takahashi et al., 2015). To summarize, behavioral and vocal interactions provided by the social environment clearly affect the vocal ontogeny of both vocal learners and vocal non-learners.

Motherese: infant-directed speech in humans

A special social feedback during infant speech acquisition is infant-directed speech. Infant-directed speech (IDS), or "motherese," is a well-known phenomenon in humans, especially during early infant speech acquisition. When speaking to infants, caregivers intuitively alter the acoustic properties of their speech compared to when speaking to other adults. Particular acoustic features that characterize motherese include a higher overall pitch, an increased frequency range and a slower cadence (i.e. speech tempo) (Fernald & Kuhl, 1987; Grieser & Kuhl, 1988). The characteristic features of motherese are universal, and independent of culture and rules of language usage (Broesch & Bryant, 2015; Grieser & Kuhl, 1988). Infants prefer IDS over adult-directed speech (ADS) (Cooper & Aslin, 1990; Fernald, 1985), probably because it attracts their attention though the more salient acoustic features of IDS. The features of motherese facilitate the perception of characteristic phonetic features and support linguistic learning (Kuhl et al., 1997). A recent study of IDS reported a very interesting new result: the authors showed that the timbre of IDS is different from ADS (Piazza, Iordan, & Lew-Williams, 2017). Research on any phenomena similar to motherese in non-human animals is restricted to one descriptive study of general acoustic parameters of infant-directed call production between Squirrel Monkey mothers and their infants (Biben, Symmes, & Bernhards, 1989). But, to my knowledge, comparative studies of acoustic and timbral differences between pup- and adult-directed vocalizations of the same individuals is completely absent for non-human animals.

Pup-directed adult vocalizations

Vocal interactions range from song tutoring to specific pup/infant-directed vocalizations that occur exclusively during pup/infant ontogeny. Pup- or infant-directed vocalizations have been described for various non-human species, particularly for gregarious species where reproduction is synchronized and many young are born within a limited period of time (Beecher, Stoddard, & Loesche, 1985; Biben, Symmes, & Bernhards, 1989; Brown, 1976; Charrier, Mathevon, & Jouventin, 2001; Esser & Schmidt, 1989; Seyfarth & Cheney, 2003; Whitham, Gerald, & Maestripieri, 2007). Pup-directed vocalizations are often simple, monosyllabic vocalizations and are exclusively produced during pup ontogeny. They are always pup directed and usually function to facilitate parent-young reunions for subsequent parental care (e.g. nursing, thermoregulation, protection). They can also influence the development of a particular vocalization, though this has been described less often (Esser & Schmidt, 1989).

Bats are highly social and gregarious animals that typically have synchronized births (Neuweiler, 2000). In bats, pup-directed vocalizations are described in only a few species, namely *Phyllostomus discolor* (Esser & Schmidt, 1989), *Antrozous pallidus* (Brown, 1976), *Tadarida brasiliensis* (Balcombe & McCracken, 1992), *Hipposideros Pomona* (Jin et al., 2015), *Desmodus rotundus* (Schmidt, 1972) and *Saccopteryx bilineata* (Knörnschild & von Helversen, 2008). Pup-directed vocalizations are mainly produced by females, whereas descriptions of adult male pup-directed vocalizations are not reported for bats. The function of the maternal directive (i.e. pup-directed call) call is to facilitate mother-pup reunions and is produced when mothers are separated from pups, as a response to pup isolation calls and during retrieval of pups (Balcombe & McCracken, 1992; Brown, 1976; Esser & Schmidt, 1989; Jin et al., 2015). Hence, pup-directed calls often encode information about individual identity.

Information content of acoustic signals

Acoustic signals can encode many types of information, for example about the signaler's quality (e.g. hierarchy rank: Fischer, Kitchen, Seyfarth, & Cheney, 2004; e.g. size: Koren & Geffen, 2009), behavioral context (e.g. group movement: Jansen, Cant, & Manser, 2012; e.g. referential alarm calls: Templeton, Greene, & Davis, 2005), motivation (aggression: e.g. Fernandez, Fasel, Knörnschild, & Richner, 2014; courtship: e.g. Jahelková, Horáček, & Bartonička, 2008) and the individual itself (individual identity: e.g. Jansen, Cant, & Manser, 2012; sex: e.g. Ey, Hammerschmidt, Seyfarth, & Fischer, 2007; age: e.g. Fernandez & Knörnschild, 2017; social group affiliation: e.g. Boughman, 1998). During vocal ontogeny, different types of information encoded in vocalizations can develop. The information content in vocalizations is influenced by various factors, including genetics, maturation of the vocal tract, and vocal learning. An interesting vocalization to study information type change during ontogeny is the pup isolation call of *S. bilineata.* This call is produced directly after birth and is maintained throughout the pup's ontogeny (Knörnschild & von Helversen, 2008). It is a multisyllabic vocalization, composed of different acoustic parts and syllables. Thus, it is a perfect candidate for investigation of ontogenetic processes and the "multiple messages hypothesis".

The "multiple messages hypothesis"

A single acoustic signal can convey multiple messages to a receiver, and sometimes these different messages are directed to different receivers (e.g. Blumstein & Armitage, 1997; Charlton, Zhihe, & Snyder, 2009; Fischer, Kitchen, Seyfarth, & Cheney, 2004; Jansen, Cant, & Manser, 2012). How can signal reliability and information transfer remain successful despite the various types of information encoded in a signal? A solution to this problem is offered by the so-called multiple messages hypothesis *sensu* (Hebets & Papaj, 2004). A signal can be composed of different elements or acoustic units emitted in a sequence, each conveying one layer of information. Hence, information content is temporally segregated. This is beneficial because it alleviates difficulties arising from the simultaneous processing of multiple types of information. Information theory shows that vocalizations occupy an acoustic space which can only encode a limited size of information (i.e. the amount of information in a bit; Beecher, 1989). By encoding different types of information in different parts of the vocalization, signal reliability is enhanced and the decoding process of the receiver is probably facilitated (Koren & Geffen, 2009; Nelson & Poesel, 2007; Templeton, Greene, & Davis, 2005). Evidence for the use of temporal segregation to encode several messages exists for the black-capped chickadees (Templeton, Greene, & Davis, 2005), banded mongooses, (Jansen, Cant, & Manser, 2012), primates (e.g. baboons; Fischer, Kitchen, Seyfarth, & Cheney, 2004) and giant pandas (Charlton, Zhihe, & Snyder 2009). Bats are a promising taxon to study temporal segregation of information in vocal communication because their social vocalizations are often composed of distinct acoustic units and contain several different elements (e.g. Behr & von Helversen, 2004; Jahelková, Horáček, & Bartonička, 2008). Nevertheless, studies on potential temporal segregation of different information types have only been conducted in a few bat species (Bohn, Schmidt-French, Schwartz, Smotherman, & Pollak, 2009; Jahelková, Horáček, & Bartonička, 2008).

Information content: the link between social and vocal complexity

Messages encoded in acoustic signals are an important source of information for receivers (Bradbury & Vehrencamp, 2011). The decoding of information encoded in social vocalizations is highly advantageous for receivers in many different contexts, especially because decision making becomes more crucial the less predictable the social environment is (Font & Carazo, 2010). Therefore, the more complex the social environment and social interactions become, the more information content should be present in vocalizations used to mediate social interactions. The idea that the complexity of vocal communication might be influenced and driven by the complexity of social structures of species is termed the "social complexity hypothesis for communication" (Bouchet, Blois-Heulin, & Lemasson, 2013; Freeberg, Dunbar, & Ord, 2012; Krams, Krama, Freeberg, Kullberg, & Lucas, 2012). This hypothesis predicts a positive feedback loop between increasing social complexity and vocal complexity.

A promising approach to investigate the role of social complexity on vocal complexity is a phylogenetic comparison of species which face similar selective pressures (Bouchet, Blois-Heulin, & Lemasson 2013; Manser et al., 2014; Wilkinson, 2003). Bats are gregarious animals that exhibit a variety of different social systems, ranging from small social groups composed of up to six individuals (e.g. Spix´s disc-winged bat; Chaverri, Gillam, & Kunz, 2012) to huge aggregations with several thousand individuals (e.g. Mexican free-tailed bat; Bohn, Schmidt-French, Schwartz, Smotherman, & Pollak, 2009). Furthermore, bats also exhibit a variety of social vocalizations, ranging from simple calls to multisyllabic calls and songs (for review see Smotherman, Knornschild, Smarsh, & Bohn, 2016). Vocal repertoire size shows a similarly large range across species, ranging from small (e.g. Knörnschild, Kalko, & Feifel, 2014) to large (e.g. Behr & von Helversen, 2004). Therefore, bats are a promising taxon for comparative research investigating the "social-vocal complexity" conundrum.

One of the best understood and thoroughly studied types of information in bats' social vocalizations is the individual signature. Individual identity signatures are important in mother-pup communication (e.g. Bohn, Wilkinson, & Moss, 2007), for recognition of social group members (e.g. Carter, Skowronski, Faure, & Fenton, 2008), coordination of group cohesion and shelter allocation (e.g. Chaverri, Gillam, & Kunz, 2012), group foraging (Boughman, 1998) and aggressive male-male interactions (e.g. Fernandez, Fasel, Knörnschild, & Richner, 2014). Recognizing conspecifics or social interaction partners is crucial in all the aforementioned scenarios, and errors can be extremely costly. Therefore, social groups with more conspecifics which can be potentially confused during daily social interactions should promote more complex vocalizations (in this case, increased information content encoded in vocalizations). Social and vocal complexity are difficult to define and assess (Freeberg, Dunbar, & Ord, 2012; Manser et al., 2014). To determine the complexity of a social system, variables such as group size, frequency of repeated encounters, consistency of social relationships (e.g. cooperation among kin and non-kin), social structures (e.g. hierarchies, social roles such as reproductive status or dominance), context (e.g. mating, predation, foraging) and temporal consistency (e.g. number of years a group remains stable, fission-fusion dynamics, longevity of single individuals) have been used (Blumstein & Kenneth, 1997; Freeberg, Dunbar, & Ord, 2012; McComb & Semple, 2005). Defining and measuring vocal complexity is similarly challenging. Currently, vocal complexity is assessed based on variables such as vocal repertoire size, compositional structure, and information content (i.e. information in bits *sensu* Beecher, 1989) of specific vocalizations (Freeberg, Dunbar, & Ord, 2012).

The study species *Saccopteryx bilineata*

Saccopteryx bilineata is a highly social animal that lives in perennial social groups and exhibits a large vocal repertoire (Behr & von Helversen, 2004; Voigt et al., 2008). The species has a polygynous mating system in which males (harem males) defend a territory (Bradbury & Emmons, 1974; Tannenbaum, 1975). A harem is composed of a harem male and up to eight females with their respective pups (Bradbury & Vehrencamp, 1976; Tannenbaum, 1975). The species exhibits an unusual dispersion behavior: young females disperse to new colonies, whereas males remain in their natal colony, where they either queue for harem positions or found a new colony elsewhere (Voigt & Streich, 2003). Colonies are thus composed of few patrilines (Nagy, Heckel, Voigt, & Mayer, 2007). Females are larger than males, and therefore males cannot monopolize mating. Thus, female choice seems to be a crucial factor, resulting in high male-male competition (Heckel & von Helversen, 2003; Voigt et al., 2008). As mentioned previously, this species exhibits an unusually large vocal repertoire. The social and vocal behavior of this species has been studied in detail for decades, which has led to a complete description of the adult vocal repertoire supplemented by the corresponding behavioral contexts (Chapter 1; Behr & von Helversen, 2004; Davidson & Wilkinson, 2004; Knörnschild & von Helversen, 2008). Males produce two sexually selected song types: the territorial and courtship songs. The territorial song is produced daily at dusk and dawn to acoustically defend a territory against potential intruders (Eckenweber & Knörnschild, 2013). Pups acquire this song through vocal production learning (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010) and it encodes various types of identity information (Behr et al., 2006; Eckenweber & Knörnschild, 2013; Knörnschild, Blüml, Steidl, Eckenweber, & Nagy, 2017). Other vocalizations produced by both sexes in the day roost comprise a variety of vocalizations emitted during agonistic, affiliative and neutral contexts (Behr & von Helversen, 2004). The entire vocal repertoire consists of 25 distinct syllable types that are combined into different vocalization types.

Each year at the beginning of the rainy season, females give synchronized birth to a single pup (Tannenbaum, 1975). The pup remains with its mother for about 12-14 weeks. During vocal ontogeny, pups produce a conspicuous vocal practice behavior that is reminiscent of infant canonical babbling (Knörnschild, Behr, & von Helversen, 2006). This extraordinary behavior – which is scarce among non-human mammals – is my main interest and the focus of this thesis.

Thesis outline and aims

This thesis investigates various novel aspects of the pups´ vocal ontogeny in *S. bilineata.* The aims of this thesis are: 1) the in-depth analysis of an extraordinary vocal practice behavior, called babbling; 2) the investigation how the social environment influences main factors of babbling; and 3) the first investigation about a phenomenon reminiscent of motherese in non-human animals. Furthermore, this thesis explores if the temporal segregation of different types of information can be found in the multisyllabic pup isolation call, as predicted by the "multiple-messages hypothesis". In the final section of the thesis, I aim to investigate the proposed positive interplay between social and vocal complexity according to the "social complexity hypothesis for communication". Results of my thesis are presented in five chapters:

Chapter I: Babbling in bat pups and human infants: common ontogenetic features of vocal production learners.

This thesis begins with a study of a conspicuous pup vocal practice behavior reminiscent of infant canonical babbling. Infant canonical babbling is characterized by nine universal features. The aim of the first chapter is to investigate if pup babbling is characterized by the same features that define infant canonical babbling. Furthermore, I want to provide the first comparison of babbling features across vocal learning mammals

Chapter II: Social influences shape the babbling behavior of greater sac-winged bat pups, *Saccopteryx bilineata*

In human speech acquisition, social feedback significantly influences speech acquisition—especially the canonical babbling phase. The social environment of *S. bilineata* pups includes the mother, other females with pups and adult males. Mothers produce different behaviors during babbling bouts and interact with their pups. Adult males do not interact with pups, but provide the acoustic input. I want to investigate how the social environment influences main babbling factors, namely the amount of vocal practice, the syllabic babbling composition and the pups´ final syllable repertoire size.

Chapter III: Pup-directed vocalizations of adult females and males in a vocal learning bat

This chapter investigates pup-directed female and male vocalizations occurring during vocal ontogeny. Motherese is infant-directed speech in humans which is characterized by several universal acoustic features such as high pitch, increased frequency contour and slow cadence.

Furthermore, a recent study revealed that the timbre (i.e. the voice color) of infant-directed speech is significantly different from the timbre of adult-directed speech. In *S. bilineata,* females produce one pup-directed vocalization and a few adult-directed vocalizations in the day roost. In this study, I aim to investigate if the timbre of the pup-directed vocalization is different compared to the timbre of adult-directed vocalizations

Furthermore, during pup ontogeny, when pups´ isolation calls (used to solicit maternal care) are converging towards a social group mean to form a group signature, adult males also produce isolation calls in relation to pup call behavior. Therefore, I aim to answer the question whether male pup-directed isolation calls may serve as auditory input for learning a group signature across generations.

Chapter IV: Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages.

In contrast to other bat species, the isolation call of *S. bilineata* pups is composed of three syllable types, namely variable, composite and stereotyped syllables. It is known, that the composite syllable of *S. bilineata* isolation calls encodes two vocal signatures: one concerning individual identity and one concerning group affiliation. This chapter aims to investigate ontogenetic changes of the isolation call and if it encodes different types of information that are temporally segregated according to the "multiple-messages" hypothesis.

Chapter V: Vocal information and the navigation of social decisions in bats: Is social complexity linked to vocal complexity?

In the last chapter, I aim to investigate the predicted positive relation between social and vocal complexity proposed by the "social complexity hypothesis for communication." Vocal signatures about identity are important for decision making in various social interactions and errors can be extremely costly. Social groups with more conspecifics which can be potentially confused during daily social interactions should support more complex vocalizations (in this case increased information content encoded in vocalizations). We aim to answer the question if increased social complexity (i.e. social group size) is positively correlated with encoded information content (i.e. individual signature strength) in social vocalizations. If so, this would support the positive relation between social and vocal complexity in bats

References

Adret, P. (2004). Vocal imitation in blindfolded zebra finches (*Taeniopygia guttata*) is facilitated in the presence of a non-singing conspecific female. *Journal of ethology, 22*(1), 29-35.

Araya-Salas, M., & Wright, T. (2013). Open-ended song learning in a hummingbird. *Biology Letters, 9*(5), 20130625.

Balcombe, J. P., & McCracken, G. F. (1992). Vocal recognition in Mexican free-tailed bats: do pups recognize mothers? *Animal Behaviour, 43*(1), 79-87.

Beecher, M. D. (1989). Signaling Systems for Individual Recognition - an Information-Theory Approach. *Animal Behaviour, 38*, 248-261. doi:10.1016/S0003-3472(89)80087-9.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song learning. *Current Directions in Psychological Science, 13*(6), 224-228.

Beecher, M. D., Stoddard, P. K., & Loesche, P. (1985). Recognition of parents' voices by young cliff swallows. *The Auk, 102*(3), 600-605.

Behr, O., & von Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology, 56*(2), 106-115. doi:10.1007/s00265-004-0768-7.

Behr, O., Von 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). *Behavioral Ecology, 17*(5), 810-817.

Biben, M., Symmes, D., & Bernhards, D. (1989). Contour variables in vocal communication between squirrel monkey mothers and infants. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 22*(6), 617-631.

Blumstein, D. T., & Armitage, K. B. (1997). Alarm calling in yellow-bellied marmots .1. The meaning of situationally variable alarm calls. *Animal Behaviour, 53*, 143-171. doi: 10.1006/anbe.1996.0285.

Blumstein, D. T. A., Kenneth B. . (1997). Does Sociality Drive The Evolution Of Communicative Complexity - A Comparative Test With Ground-Dwelling Sciurid Alarm Calls. *The American Naturalist, 150*(2), 179-200.

Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and Stereotypy of Free-Tailed Bat Songs. *PLoS One, 4*(8). doi:ARTN e674610.1371/journal.pone.0006746.

Bohn, K. M., Wilkinson, G. S., & Moss, C. F. (2007). Discrimination of infant isolation calls by female greater spear-nosed bats, *Phyllostomus hastatus*. *Animal Behaviour, 73*, 423-432. doi:10.1016/j.anbehav.2006.09.003.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Frontiers in Psychology, 4*, 390. doi:10.3389/fpsyg.2013.00390.

Boughman, J. W. (1998). Vocal learning by greater spear–nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 265*(1392), 227-233.

Boughman, J. W., & Moss, C. F. (2003). Social sounds: vocal learning and development of mammal and bird calls. In *Acoustic communication* (pp. 138-224): Springer.

Bradbury, J. W., & Emmons, L. H. (1974). Social organization of some Trinidad bats: I. Emballonuridae. *Zeitschrift für Tierpsychologie, 36*(1-5), 137-183.

Bradbury, J. W., & Vehrencamp, S. L. (1976). Social-Organization and Foraging in Emballonurid Bats .2. Model for Determination of Group-Size. *Behavioral Ecology and Sociobiology, 1*(4), 383-404. doi:10.1007/Bf00299400.

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication* (Vol. Second edition): Sinauer Associates Inc., Sunderland USA.

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development, 16*(1), 31-43.

Brown, P. (1976). Vocal communication in the pallid bat, *Antrozous pallidus*. *Zeitschrift für Tierpsychologie, 41*(1), 34-54.

Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology, 29*(4), 631-636. e633.

Carter, G. G., Skowronski, M. D., Faure, P. A., & Fenton, B. (2008). Antiphonal calling allows individual discrimination in white-winged vampire bats. *Animal Behaviour, 76*(4), 1343-1355.

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009). The information content of giant panda, Ailuropoda melanoleuca, bleats: acoustic cues to sex, age and size. *Animal Behaviour, 78*(4), 893-898. doi:10.1016/j.anbehav.2009.06.029.

Charrier, I., Mathevon, N., & Jouventin, P. (2001). Mother's voice recognition by seal pups. *Nature, 412*(6850), 873.

Chaverri, G., Gillam, E. H., & Kunz, T. H. (2012). A call-and-response system facilitates group cohesion among disc-winged bats. *Behavioral Ecology, 24*(2), 481-487. doi:10.1093/beheco/ars188.

Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences, 113*(24), 6641-6646.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*(5), 1584-1595.

Davidson, S. M., & Wilkinson, G. S. (2004). Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Animal Behaviour, 67*(5), 883-891.

Deutscher, G. (2010). Through the language glass: How words colour your world.

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

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

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998). Infant 'babbling' in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour, 135*(5), 643-664.

Esser, K. H., & Schmidt, U. (1989). Mother-Infant Communication in the Lesser Spear-Nosed Bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae) - Evidence for Acoustic Learning. *Ethology, 82*(2), 156-168. doi: 10.1111/j.1439-0310.1989.tb00496.

Esser, K. H., & Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae) - evidence for acoustic learning. *Ethology, 82*(2), 156-168.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., & Fischer, J. (2007). Age- and Sex-Related Variations in Clear Calls of *Papio ursinus*. *International Journal of Primatology, 28*(4), 947-960. doi:10.1007/s10764-007-9139-3.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*(2), 181-195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279-293.

Fernandez, A. A., Fasel, N., Knörnschild, M., & Richner, H. (2014). When bats are boxing: aggressive behaviour and communication in male Seba's short-tailed fruit bat. *Animal Behaviour, 98*, 149-156.

Fernandez, A. A., & Knörnschild, M. (2017). Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages. *Animal Behavior and Cognition, 4*(2), 169-186.

Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology, 56*(2), 140-148. doi:10.1007/s00265-003-0739-4.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences, 4*(7), 258-267.

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

Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annual Review of Linguistics, 4*, 255-279.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi:10.1016/j.cognition.2005.02.005.

Font, E., & Carazo, P. (2010). Animals in translation: why there is meaning (but probably no message) in animal communication. *Animal Behaviour, 80*(2), e1-e6.

Freeberg, T. M., Dunbar, R. I., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transcations of the Royal Society B: Biological Sciences, 367*(1597), 1785-1801. doi:10.1098/rstb.2011.0213.

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the “critical period”. *Brain and language, 1*(1), 81-107.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology, 18*(11), R457-R460.

Gleitman, L. R., & Newport, E. L. (2002). The invention of language by children: Environmental and biological influences.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological science, 19*(5), 515-523.

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. *The Oxford handbook of developmental and comparative neuroscience*, 708-729.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*(1), 14.

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), 1569-1579. doi:10.1126/science.298.5598.1569.

Hebets, E. A., & Papaj, D. R. (2004). Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology, 57*(3), 197-214. doi:10.1007/s00265-004-0865-7.

Heckel, G., & Von Helversen, O. (2003). Genetic mating system and the significance of harem associations in the bat *Saccopteryx bilineata*. *Molecular Ecology, 12*(1), 219-227.

Jahelková, H., Horáček, I., & Bartonička, T. (2008). The advertisement song of *Pipistrellus nathusii* (Chiroptera, Vespertilionidae): a complex message containing acoustic signatures of individuals. *Acta Chiropterologica, 10*(1), 103-126. doi:10.3161/150811008x331144.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology, 28*, 60-65.

Janik, V. M., & Slater, P. J. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour, 26*, 59-100.

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

Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biology, 10*, 1741-7007.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016*(1), 749-777.

Jin, L., Yang, S., Kimball, R. T., Xie, L., Yue, X., Luo, B., . . . Feng, J. (2015). Do pups recognize maternal calls in pomona leaf-nosed bats, *Hipposideros pomona*? *Animal Behaviour, 100*, 200-207.

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: parallels to human speech development. *Ethology, 111*(1), 101-117.

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

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*(9), 451-454. doi:10.1007/s00114-006-0127-9.

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

Knörnschild, M., Kalko, E. K. V., & Feifel, M. (2014). Male courtship displays and vocal communication in the polygynous bat *Carollia perspicillata*. *Behaviour, 151*(6), 781-798. doi:10.1163/1568539x-00003171.

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. doi:10.1098/rsbl.2009.0685.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Koren, L., & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. *Behavioral Ecology and Sociobiology, 63*(4), 581-590. doi:10.1007/s00265-008-0693-2.

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1597), 1879-1891. doi:10.1098/rstb.2011.0222.

Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(17), 9645-9646. doi:10.1073/pnas.1733998100.

Kuhl, P. K. (2003). Human speech and birdsong: communication and the social brain. *Proceedings of the National Academy of Sciences, 100*(17), 9645-9646.

Kuhl, P. K. (2007). Is speech learning ‘gated’by the social brain? *Developmental science, 10*(1), 110-120.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., . . . Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684-686. doi:10.1126/science.277.5326.684.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the" critical period". *Language Learning and Development, 1*(3-4), 237-264.

Locke, J. L., & Pearson, D. M. (2009). Linguistic significance of babbling: evidence from a tracheostomized infant. *Journal of Child Language, 17*(01). doi:10.1017/s0305000900013076.

Manser, M. B., Jansen, D. A., Graw, B., Hollén, L. I., Bousquet, C. A., Furrer, R. D., & le Roux, A. (2014). Vocal complexity in meerkats and other mongoose species. In *Advances in the Study of Behavior* (Vol. 46, pp. 281-310): Elsevier.

Manser, M. B., Jansen, D. A. W. A. M., Graw, B., Hollén, L. I., Bousquet, C. A. H., Furrer, R. D., & le Roux, A. (2014). Vocal Complexity in Meerkats and Other Mongoose Species. In *Advances in the Study of Behavior, 46*, 281-310.

Margoliash, D., & Nusbaum, H. C. (2009). Language: the perspective from a organismal biology. *Trends in Cognitive Sciences, 13*(12), 505-510. doi:10.1016/j.tics.2009.10.003.

Marler, P. (1970). Birdsong and Speech Development - Could There Be Parallels. *American Scientist, 58*(6), 669-673.

Martins, P. T., & Boeckx, C. (2016). What we talk about when we talk about biolinguistics. *Linguistics Vanguard, 2*(1).

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biol Lett, 1*(4), 381-385. doi:10.1098/rsbl.2005.0366.

Nagy, M., Heckel, G., Voigt, C. C., & Mayer, F. (2007). Female-biased dispersal and patrilocal kin groups in a mammal with resource-defence polygyny. *Proceedings of the Royal Society B: Biological Sciences, 274*(1628), 3019-3025. doi:10.1098/rspb.2007.1008.

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. doi:10.1016/j.anbehav.2007.01.018.

Neuweiler, G. (2000). *The biology of bats*: Oxford University Press on Demand.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive science, 14*(1), 11-28.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nature Communications, 5*, 3379.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In J. F. K. a. C. A. F. G. H. Yeni-Komshian (Ed.), *Child phonology* (Vol. 1, pp. 93-112). New York: Academic Press, Inc.

Oller, D. K., & Eilers, R. E. (1982). Similarity of babbling in Spanish-and English-learning babies. *Journal of Child Language, 9*(3), 565-577.

Payne, R. B. (1981). Song learning and social interaction in indigo buntings. *Animal Behaviour, 29*(3), 688-697.

Pepperberg, I. M. (1994). Vocal learning in grey parrots (*Psittacus erithacus*): effects of social interaction, reference, and context. *The Auk, 111*(2), 300-313.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology, 27*(20), 3162-3167. e3163.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature, 434*(7032), 455.

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.

Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology, 28*, 66-71.

Schmidt, U. (1972). Die sozialen Laute juveniler Vampirfledermäuse (*Desmodus rotundus*) und ihrer Mütter.

Schusterman, R. J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. *The evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication*, 41-70.

Seyfarth, R., & Cheney, D. L.(1986). Vocal development in vervet monkeys. *Animal Behaviour*, 1640-1658.

Seyfarth, R. M., & Cheney, D. L. (2003). Meaning and emotion in animal vocalizations. *ANNALS-NEW YORK ACADEMY OF SCIENCES*, 32-55.

Smith, V. A., King, A. P., & West, M. J. (2000). A role of her own: female cowbirds, *Molothrus ater*, influence the development and outcome of song learning. *Animal Behaviour, 60*(5), 599-609.

Smotherman, M., Knörnschild, M., Smarsh, G., & Bohn, K. (2016). The origins and diversity of bat songs. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 202*(8), 535-554. doi:10.1007/s00359-016-1105-0.

Stoll, S., Lieven, E., Winskel, H., & Padakannaya, P. (2013). Studying language acquisition cross-linguistically. In *South and Southeast Asian Psycholinguistics,* 19-35.

Suzuki, T. N. (2016). Semantic communication in birds: evidence from field research over the past two decades. *Ecological research, 31*(3), 307-319.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science, 349*(6249), 734-738.

Tannenbaum, B. R. (1975). *Reproductive strategies in the white-lined bat.* (PhD). Cornell University Ithaca NY USA,

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science, 308*(5730), 1934-1937. doi:10.1126/science.1108841.

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*: Blackwell Publishing.

Vihman, M. M. (2014). *Phonological development: The first two years*: Wiley-Blackwell Boston, MA.

Vihman, M. M., Ferguson, C. A., & Elbert, M. (1986). Phonological Development from Babbling to Speech - Common Tendencies and Individual-Differences. *Applied Psycholinguistics, 7*(1), 3-40. doi: 10.1017/S0142716400007165.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat,*Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

Voigt, C. C., & Streich, W. J. (2003). Queuing for harem access in colonies of the greater sac-winged bat. *Animal Behaviour, 65*, 149-156. doi:10.1006/anbe.2002.2031.

West, M. J., & King, A. P. (1988). Vocalizations of juvenile cowbirds (*Molothrus ater ater*) evoke copulatory responses from females. *Developmental Psychobiology, 21*(6), 543-552.

Whitham, J. C., Gerald, M. S., & Maestripieri, D. (2007). Intended receivers and functional significance of grunt and girney vocalizations in free-ranging female rhesus Macaques. *Ethology, 113*(9), 862-874.

Wilkinson, G. S. (2003). Social and vocal complexity in bats. In *Animal social complexity: Intelligence, culture, and individualized societies,* (322-341), de Waal & Tyack (Eds.)

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., . . . Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron, 104*(1), 87-99.

Babbling in bat pups and human infants: common ontogenetic features of vocal production learners

Authors: Ahana A. Fernandez, Lara S. Burchardt & Mirjam Knörnschild

ABSTRACT

Canonical babbling in infants is a conspicuous speech behavior and, together with vocal production learning (VPL), a key element in spoken language acquisition. Infants acquire their phonological repertoire by imitation of speech sound subunits, usually consonant-vowel pairs which are produced in reduplicated sequences. To shed light on the mechanisms necessary for the evolution of spoken language, comparative studies on VPL species, which babble during ontogeny, are especially promising. Nevertheless, current comparisons are restricted to songbirds (plastic song is analogous to babbling) and one non-VPL mammal, the pygmy marmoset. We introduce a new mammal as a promising candidate for comparative research on the ontogeny of VPL. The bat *Saccopteryx bilineata* exhibits VPL and a conspicuous ontogenetic phase of intense vocal practice highly reminiscent of infant canonical babbling, i.e. a daily production of several babbling bouts with durations of up to 43 minutes each. We investigated the entire 10-week vocal ontogeny of 20 pups from two different regions in Central America. We found that pup babbling is characterized by the same nine features that define infant canonical babbling; early ontogenetic onset, babbling sequences composed of non-adult elements and simultaneous emergence of adult-like syllable types, vocal overproduction, subset acquisition of the adult repertoire, non-linear acquisition pattern of syllable types, repetitiveness, rhythmicity, meaninglessness, and universality. These findings suggest that a distinctive vocal practice phase is characterized by similar features across VPL species, probably arising from similar underlying mechanisms necessary to acquire large adult vocal repertoires.

SIGNIFICANCE STATEMENT

Canonical babbling is a milestone in speech development and mandatory for spoken language acquisition. Many studies have found commonalities in the vocal development of human infants and songbirds but detailed comparisons between human infants and other mammals capable of vocal production learning have been lacking so far. In this study, we report a conspicuous vocal practice phase, i.e. babbling, of the bat *Saccopteryx bilineata*, which is a vocal production learner. We found that features defining canonical babbling in human infants are likewise characteristic of pup babbling in bats. The obvious parallels in vocal ontogeny between different species exhibiting vocal production learning offer exciting future possibilities to unravel shared mechanisms and cognitive skills mandatory for the acquisition of large vocal repertoires.

Keywords: pup babbling, infant canonical babbling, vocal ontogeny, *Saccopteryx bilineata,* bats, spoken language acquisition, vocal practice

INTRODUCTION

Language, the cornerstone of human culture, is a complex system composed of key components like vocal production learning (VPL), syntactical composition, and semantic reference (Fitch, 2010; Hauser, Chomsky, & Fitch, 2002; Margoliash & Nusbaum, 2009). 'Speech' is the vocal motor output aspect of language and its realization requires a defined and complex control of the jaw, orofacial, and laryngeal muscles. Every child is challenged with the acquisition of speech sounds forming the basis for spoken language. Considering the complexity of this task, it remains puzzling how children acquire speech and spoken language in early infancy seemingly without effort (Oller & Eilers, 1988; Stoll, Lieven, Winskel, & Padakannaya, 2013). Human speech is constituted of language specific phonetic repertoires, which are used to create syllables that are subsequently combined into words (Oller & Eilers, 1988; Vihman, 1996; Stoll, Lieven, Winskel, & Padakannaya, 2013). Consequently, the first milestone in speech acquisition, the canonical babbling phase, concerns the acquisition of the language-specific phonetic repertoire (Vihman, 1996). During canonical babbling (CB), children acquire speech subunits (usually consonant-vowel pairs) of syllables, which are the basic unit of most languages (Oller, 1980; Vihman, 1996; Vihman, Ferguson, & Elbert, 1986). Children pay attention to the different frequencies of occurrence of certain language specific vowels, consonants and pairings thereof, thus enabling them to acquire the relevant phonetic units through statistical learning (Aslin & Newport, 2012; Aslin, Saffran, & Newport, 1998; Kuhl, 2004; Misyak, Goldstein, & Christiansen, 2012; Saffran, Aslin, & Newport, 1996). While babbling, children imitate and rehearse the relevant phonetic units of their native language by matching their own vocal production to the engram of the perceptual input (Doupe & Kuhl, 1999; Oller & Eilers, 1988). Consequently, perception of tutors' and one's own vocal production is crucial for laying the foundation for the next functional level, i.e. speech (Vihman, 1996). Canonical babbling is stringently required for speech acquisition, irrespective of the language to be learned. The few studies reporting cases where children were tragically deprived of social contact as well as studies with deaf or tracheotomized children agree that babbling is mandatory for language acquisition (Fagan, 2015; Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974; Locke & Pearson, 2009).

To investigate the biological foundations of the human language capacity and its evolution, biolinguistic research plays an important role (Boeckx, Leivada, & Martins, 2013; Martins & Boeckx, 2016). The key assumption of the field of biolinguistics is that the possibility of acquiring and using language is based on biological factors and thus can be studied from different biological perspectives such as neurobiology, genetics, and cognitive psychology (Boeckx & Piattelli-Palmarini, 2005; Fitch, 2009). By using a comparative approach it is possible to investigate if and to what extend key features of human language are present in other taxa, thus increasing our understanding of human language evolution (Fitch, 2000, 2018; Fitch, Hauser, & Chomsky, 2005). So far, songbirds have been the preferred model taxon for comparative biolinguistic analyses on human speech development (Doupe & Kuhl, 1999; Goldstein & Schwade, 2010; Kuhl, 2003). The suitability of comparing speech development in infants and song acquisition in songbirds is essentially based on their ability of VPL and their conspicuous periods of vocal practice (Beecher & Brenowitz, 2005; Doupe & Kuhl, 1999; Konishi, 1985; Kuhl, 2003; Marler, 1970). Songbirds also share with humans anatomical and functional similarities in the organization of neuronal pathways for perception and production (Jarvis, 2004; Prather, Okanoya, & Bolhuis, 2017). Moreover, research in songbirds has demonstrated that the social environment (i.e. tutors – parents, siblings or unrelated conspecifics) influences the vocal development of the young bird, as is the case in infant speech development (Albert, Schwade, & Goldstein, 2018; Beecher & Burt, 2004; Carouso-Peck & Goldstein, 2019; Goldstein, King, & West, 2003; Goldstein & Schwade, 2008, 2010; Goldstein & West, 1999; West & King, 1988). Additionally, the plastic song produced by juvenile male songbirds is reminiscent of CB in infants (Bolhuis, Okanoya, & Scharff, 2010; Brainard & Doupe, 2002; Marler, 1970). However, despite those advantages, there are limitations of the songbird model, the most obvious being the phylogenetic distance between songbirds and humans (Table 1). So far, babbling outside the taxon of songbirds has been reported to occur in three species; in the parrot *Forpus passerinus,* the pygmy marmoset *Cebullea pygmaea* and the greater sac-winged bat *Saccopteryx bilineata* (Elowson, Snowdon, & Lazaro-Perea, 1998; Knörnschild, Behr, & von Helversen, 2006; Tatiana, 2017). The major difference to human infants is the fact that pygmy marmosets are not capable of VPL (Table 1), despite being very vocal (e.g. diverse and large adult vocal repertoire, distinctive vocal ontogeny phase). Otherwise, pygmy marmoset babbling shares several features with infant CB; in contrast to most songbirds, both sexes engage in babbling (Elowson, Snowdon, & Lazaro-Perea, 1998). However, it should be stressed here that the present songbird literature is still dominated by studies from the Northern Hemisphere where song is predominantly produced by males (Garamszegi, Pavlova, Eens, & Møller, 2006; Mahr, Seifert, & Hoi, 2016). An increasing number in studies about female song (Evans & Kleindorfer, 2016; Geberzahn & Gahr, 2013; Odom, Hall, Riebel, Omland, & Langmore, 2014; Riebel, 2016; Roper, Harmer, & Brunton,

2018) will most likely show that song development in female birds follows the same developmental stages described for males.

Table 1. Differences in vocal ontogeny between babbling species

Traits	Humans	Songbirds	Pygmy marmosets	*Saccopteryx bilineata*
Vocal production learner	Yes	Yes	No	Yes
Babbling				
Present in both sexes	Yes	No (only males*)	Yes	Yes
Acquisition of all or most of the entire adult syllable repertoire	Yes	No, only song syllables	Yes	Yes
Onset of babbling	Infancy (around 7 months of age)	Puberty (Subsong: ~3-4 months of age. Plastic song: 5months of age until crystallization of song, often during spring of the following year)**	Infancy (around 3 weeks after birth)	Infancy (around 2-3 weeks after birth)

* But see references about female song development mentioned in the text. **Depends on whether the birds are open or closed ended learners and whether sensory and sensorimotor phases overlap (Doupe & Kuhl, 1999; Konishi, 1985). But see for zebra finch: subsong production at 30 days of age, plastic song around 50-60 days of age (Pytte & Suthers, 2000).

With this study, we want to put forward a new species for comparative biolinguistic studies on vocal ontogeny. The bat *S. bilineata* not only exhibits VPL (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010) but also a conspicuous phase of vocal practice during ontogeny (Fig. S1*A*). Pups' vocal practice is reminiscent of infant CB (Knörnschild, Behr, & von Helversen, 2006); it is organized in bouts composed of long multisyllabic sequences that are interspersed by short silent intervals (Fig. 1). Babbling bouts have an average duration of several minutes and a single bout can last up to 43 minutes. This extraordinarily long bout duration is astonishing considering that the longest vocalizations of adult bats have durations of several seconds to less than one minute (Fig. S2) (Behr & von Helversen, 2004). Babbling is a diurnal behavior (Strauss, von Helversen, & Knörnschild, 2010). Pups spend about 30% of their active time during the day with vocal practice (Fig. S1*B*). Both male and female pups repeatedly produce babbling bouts throughout the entire day (Knörnschild, Behr, & von Helversen, 2006) which are audible to human observers over a distance of several meters. It is conceivable that such an energetically demanding and potentially costly behavior (e.g. increased predation risk at the day-roost) serves an important function. Indeed, pups learn to produce song by imitating adult tutors while babbling; pups' renditions of adult song are often found in babbling bouts (Knörnschild, Behr, & von Helversen, 2006; Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010).

A prerequisite for investigating a species' vocal ontogeny is detailed knowledge about the adult vocal repertoire. For decades the social behavior and communication of *S. bilineata* have been studied in detail (Bradbury & Emmons, 1974; Bradbury & Vehrencamp, 1977; Knörnschild & Nagy, 2016; Tannenbaum, 1975; Voigt et al., 2008). The adult vocal repertoire is delineated completely and supplemented by the corresponding behavioral context (Behr & von Helversen, 2004; Davidson & Wilkinson, 2002; Knörnschild & von Helversen, 2008). Both adult females and males utter several different vocalization types but only adult males produce songs, i.e. territorial and courtship song (Behr & von Helversen, 2004). Similar to songbirds, adult males engage in daily territorial chorusing at dusk and dawn (Behr, Knörnschild, & von Helversen, 2009; Eckenweber & Knörnschild, 2013; Knörnschild, Blüml, Steidl, Eckenweber, & Nagy, 2017). Pups learn the territorial song during ontogeny by imitating adult males' songs (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Adult males also produce courtship songs every day to interact with females (Behr & von Helversen, 2004) but it is currently unknown whether the courtship song is learned as well. This level of detailed knowledge on *S. bilineata*'s vocal communication is the indispensable basis for in-depth studies on vocal ontogeny.

We aim to provide the first comparison of babbling features across vocal learning mammals. The scope of this study is to analyze whether babbling behavior in *S. bilineata pups* is characterized by the same set of features used to define infant CB. Research about infant speech acquisition demonstrated that despite the differences between language-specific phonetic repertoires, CB is characterized by several common features (in Latin numbers below). CB is defined by an early onset in infancy (I) (Oller, 1980). This transitional stage of development offers the child the possibility of vocal exploration through babbling, thus providing the vocal basis for the next functional level, i.e. speech (Vihman, 1996). Therefore, babbling is characterized by both the emergence of canonical syllables (i.e. syllables with adult-like resonance properties and mature temporal vowel-consonant transitions) and the presence of non-speech sounds, called protophones (Oller, 1980; Oller, Griebel, & Warlaumont, 2016) (II), the latter leading to vocal overproduction (III) (Doupe & Kuhl, 1999; Oller & Eilers, 1988; Oller, Wieman, Doyle, & Ross, 2008; Vihman, 1996; Vihman, Ferguson, & Elbert, 1986). Vocal overproduction is defined as the oral output which is not classifiable as adult speech, respectively adult vocalizations, and includes very variable sounds including for example quasi-resonant vowels, raspberries, squeals and yells (Oller, 1980). The most conspicuous feature of CB is the reduplication of syllables (IV) concatenated to vocal sequences (Oller, 1980). During CB, infants acquire only a subset of the adult phonetic repertoire (V), and the emergence and production of new speech sounds is non-linear (VI) (Vihman, 1996). The cyclic alternation of consonant-vowel production leads to temporal regularity, i.e. rhythmicity (VII) in CB (Vihman, 1996). Furthermore, CB is meaningless (VIII), i.e. lacks semantic information and is universal (IX), irrespective of the sex, culture, and language to be learned (Kent, Mitchell, & Sancier, 1991; Oller & Eilers, 1982; Vihman, 2014). We hypothesized that babbling in *S. bilineata* pups is characterized by the aforementioned nine features described for infant CB. Furthermore, we hypothesized that babbling bouts become more diverse during vocal ontogeny, i.e. innate vocalizations become less frequent and newly acquired syllables become more frequent.

RESULTS

We investigated the vocal ontogeny (from birth until weaning) of 20 pups from two populations in Costa Rica and Panama (10 pups per population). In general, vocal ontogeny was composed of two distinct phases: the *pre-babbling phase* and the *babbling phase* during which pups engaged in *babbling behavior* (Fig. S1; for detailed information about definitions of terms marked in *italics,* see *SI Appendix*). Babbling was organized in three hierarchical levels (Fig. 1). The basic unit were *syllables* (level 1) which were concatenated to sequences, called *syllable trains* (level 2), which – interspaced with silent intervals – were composed into *babbling bouts* (level 3; Fig. 1.). *Babbling bouts* had a mean duration of 7 minutes (max. duration: 43 min.).

The entire adult vocal repertoire is delineated (Fig. S9) and was used for visual classification of *syllable types* in babbling bouts. Pup syllable types had a high spectro-temporal similarity with adult syllable types (Fig. 2), thus making their manual classification feasible. The resemblance to adult syllable types was obvious when a new syllable type was first produced in a babbling bout allowing identifying their occurrence and first appearance. Our visual classification was statistically verified by measuring a subset of our data (*SI, Appendix*). The entire *adult vocal repertoire* is composed of 25 different syllable types which can be combined into ten different, often multisyllabic vocalization types. For a general overview of sample sizes of the different analyses performed on levels 1-3 see supporting information (Table S7). The data analyses for each of the three levels were performed independent from each other, thus resulting in different sample sizes for each level (e.g. number of babbling bouts analyzed). Furthermore, to analyze features III and VI (both within level 1) we analyzed a data subset which resulted in different sample sizes compared to the ones used for the analysis of the other babbling features within level 1 (*SI Appendix,* Table S7).

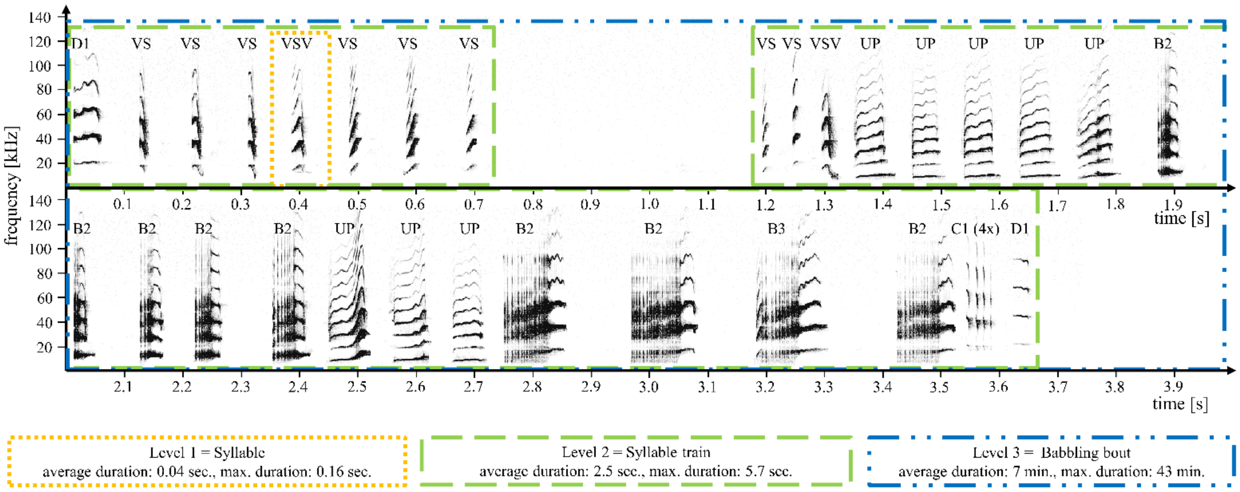


Fig. 1. Excerpt of a babbling bout.

The spectrogram shows the three-level organization of our analyses: syllables (Level 1: yellow color) are embedded in syllable trains (Level 2: green color) which are embedded in babbling bouts (Level 3: blue color). Definitions of syllables, syllable trains and babbling bouts are provided in the supplementary information. Labels depict different syllable types. UP nominates the highly variable undifferentiated proto-syllable type which is characteristic for babbling bouts. Other labels nominate different syllable types found in the adult vocal repertoire (for explanations of the syllable type labels see *SI Appendix)*. Average and maximum durations of syllables, syllable trains and babbling bouts are given. This spectrogram, and all others shown in this publication, depicts frequency [kHz] as a function of time [s] and was generated using a 1024 point fast Fourier transform and a Hamming window with 85% overlap.

Level 1: Syllable level

On the syllable level, we analyzed at least one babbling bout per week and pup (216 babbling bouts from 20 pups) and investigated the following babbling features: individual age at babbling onset (I), the syllabic composition of babbling bouts (II), vocal overproduction (III), the size of the *pups' final syllable type repertoire* to find out whether pups acquired the entire or a subset of the adult vocal repertoire (IV), the acquisition pattern of syllable types (V), repetitiveness (VI), meaninglessness (VIII) and universality (IX). To analyze whether babbling was characterized by vocal overproduction (III) and repetitiveness (VI), we classified 56,611 syllables of 10 pups (Panama population) into syllable types. This allowed us to calculate the frequency of occurrence of syllable types and the transition probabilities between syllable types. Rhythmicity (VII) was investigated in syllable trains (level 2; see below).

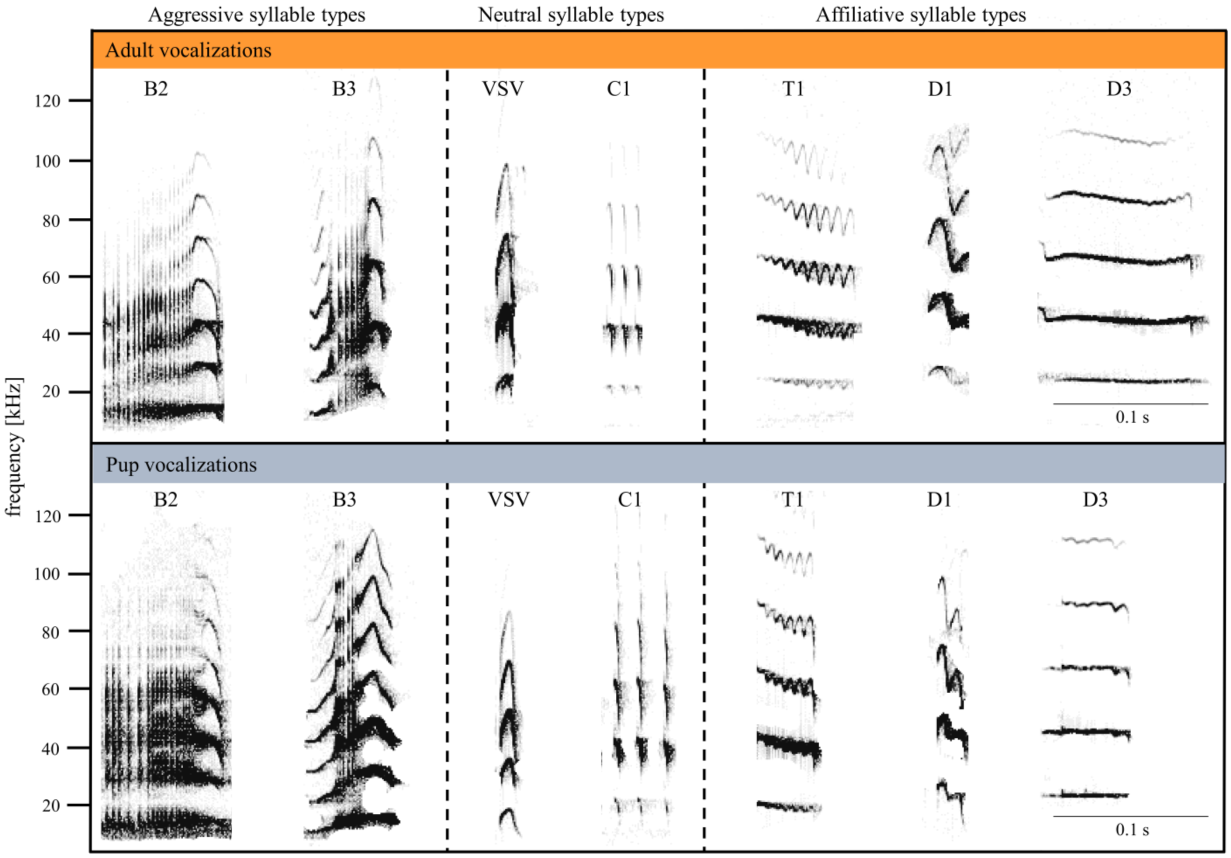


Fig. 2. Similarity of syllable types from the adult vocal repertoire and adult-like syllables produced by pups.

The upper panel depicts a selection of syllable types of the adult vocal repertoire. Labels above syllables indicate the syllable type. The complete syllabic repertoire of adult *S. bilineata* (25 syllable types) can be found in the *SI Appendix*. The lower panel depicts the corresponding syllable types produced by pups. Syllables in babbling bouts were described based on spectro-temporal parameters and categorized based on their similarity to adult syllable types.

Babbling features

(I) Early babbling onset

Babbling onset occurred early in pup development, namely within the first three weeks after birth, which is approximately at one third of the entire vocal ontogeny (on average 9.5 weeks, Fig. 3*A*). The babbling phase lasted about three times longer than the pre-babbling phase (mean duration 51 vs. 17.2 days; N=20 pups, Mann-Whitney-U test: W=2, p < 0.001, for separate calculation per population see *SI Appendix*). The vocalizations during the pre-babbling phase were confined to multisyllabic isolation calls (*SI Appendix,* Fig. S8*A*) and echolocation calls.

The babbling phase began as soon as pups started perching without direct body contact to their mothers in the day-roost.

(II) Babbling bout composition and (III) vocal overproduction

Babbling bouts were composed of so called *undifferentiated proto-syllables* and adult-like syllable types (Fig. 3*B*). Undifferentiated proto-syllables were multiharmonic tonal syllables present during the entire babbling phase; their high variability caused considerable vocal overproduction (*SI Appendix,* Fig. S8*B*). Undifferentiated proto-syllables were the most frequent and thus the most characteristic syllable type in babbling bouts (39% of 56,611 syllables; *SI Appendix,* Fig. S8*B*) because pups produced them when switching from one adult-like syllable type to the next. Undifferentiated proto-syllables were exclusively produced by pups. Since the vocal repertoire of adult *S. bilineata* is completely described (own descriptions of syllable types and former studies (Behr & von Helversen, 2004; Davidson & Wilkinson, 2002; Knörnschild, Behr, & von Helversen, 2006; Knörnschild & Nagy, 2016; Knörnschild & von Helversen, 2008) we are certain that undifferentiated proto-syllables are not part of the adult vocal repertoire. An analysis of acoustic parameters from a set of undifferentiated proto-syllables demonstrated that they are clearly distinguishable from the adult-like syllable types (*SI Appendix).* Furthermore, they were exclusively present in babbling bouts; therefore, babbling behavior was unambiguously recognizable and also clearly distinguishable from the former infrequent isolation call production in the pre-babbling phase.

(IV) Subset acquisition of the adult syllable type repertoire

Not a single pup acquired the entire adult syllable type repertoire. In all cases a subset of the 25 adult syllable types was reproduced (Fig. 3*C*). Final pup repertoire sizes varied, ranging from 12 to 20 acquired adult syllable types (*SI Appendix,* Table S1). Final pup syllable type repertoires had universal characteristics concerning presence or absence of particular syllable types: First, all pups acquired the same ten syllable types, amongst them the syllable types belonging to the later adult territorial song which is learned during ontogeny(Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Second, for all pups the same five adult syllable types had not emerged yet at the time of weaning; three of those five absent syllable types are part of the adult courtship song, the other two syllable types belong to the neutral respectively aggressive syllable type category (*SI Appendix,* Table S1 (Behr & von Helversen, 2004). Variability in the final size and composition of pups' syllable type repertoires resulted mainly from differences in the acquisition of affiliative syllable types (*SI Appendix,* Table S1, column 6).

(V) Non-linear acquisition pattern of adult-like syllable types

The ontogenetic acquisition of new adult-like syllable types followed a non-linear pattern (Fig. 3*D*). An initial "acquisition spurt" could be observed during the early babbling phase; a short time period (within 5 days after babbling started) in which the acquisition of new syllable types was at its maximum. During the remaining babbling phase, syllable acquisition was much slower. Most syllable types acquired after the initial acquisition spurt were affiliative (belonging to the later adult male courtship song). When the acquisition process stopped, babbling pups were rehearsing the syllable types acquired so far. Therefore, this time period from acquiring the last syllable type to weaning was termed *repetition phase* (*SI Appendix,* Fig. S5). We documented the repetition phase for 15 out of 20 pups. For the remaining five pups, we did not detect a repetition phase since they still acquired a new syllable type in our last recording session. Those five pups babbled for approximately two more weeks after our last recording session but, unfortunately, we were not able to collect more recordings of sufficient quality for analysis.

(VI) Repetitiveness

Babbling bouts were dominated by the reduplication of syllable types (Fig. 3*E*). In 76% of cases (100%=56,611 syllables), syllable types were succeeded by the same syllable type (Dataset S1). Repetition of the same syllable type was noted in four of the five *syllable type categories* (i.e. aggressive type, undifferentiated proto-syllable type, isolation call type, neutral type category). On the contrary, the majority of affiliative syllable types were not self-repeating (Dataset S1).

(VIII) Meaninglessness

Neither the syllabic composition nor the succession of syllables within syllabic trains was correspondent to adult vocalizations described for S. *bilineata*. Vocalizations produced by babbling pups did not elicit the reaction they would normally cause when they were produced by adult bats. For instance, territorial song normally elicits counter-singing in adult males but, this was never the case (personal observation A. A. F. & M. K.) when babbling pups produced syllables belonging to the territorial song or even whole renditions of territorial song (Fig. 3*G*). Furthermore, babbling was not correlated with a single specific behavioral context. Pup babbling was either accompanied by and restricted to interactions of mother-pup dyads or it occurred when a pup perched in the day-roost without interacting with conspecifics. When interactions occurred, both mothers and pups crawled towards or hovered in front of each other (*SI Appendix,* Fig. S7). While the pups babbled the whole time during interactions, mothers only rarely vocalized during these interactions and, if so, produced so-called directive calls (Knörnschild & von Helversen, 2008).

(IX) Universality

Irrespective of sex or regional origin, all 20 pups engaged in babbling during ontogeny (Fig. 3*H*). Similarly, all pups produced undifferentiated proto-syllables throughout the entire babbling phase. Moreover, neither age at babbling onset nor the duration of the babbling phase was significantly different between sexes or the different regions (*SI Appendix,* Table S2). Also, the size of the pups' final syllable type repertoire was not significantly different between regions or sexes (Mann-Whitney-U test, regions: W=67.5, p=0.19; pup sex: W=34, p=0.06, Bonferroni corr. α=0.05/2). Pups of both sexes acquired the syllable types constituting the adult male territorial song, even though only males sing as adults (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010).

Level 2: Syllable trains

The second level (N=8,448 syllable trains from 20 pups) included analyses on the level of *syllable type trains* (i.e. a train of at least five consecutive syllables; trains were interspaced by silent intervals). Five different *syllable type train categories* were established based on the function of the syllable types in the adult vocal repertoire, namely the *neutral -, aggressive -, affiliative -, undifferentiated proto-syllable-* and *isolation call* – category and trains were classified based on visual scanning of each bout. To analyze whether babbling is rhythmic (VII), we analyzed the inter-onset-intervals (IOI) between syllables in trains from each train category (N=30 trains, 712 syllables). Furthermore, we used multivariate models to investigate how the train composition of babbling bouts changed during ontogeny.

(VII) Rhythmicity

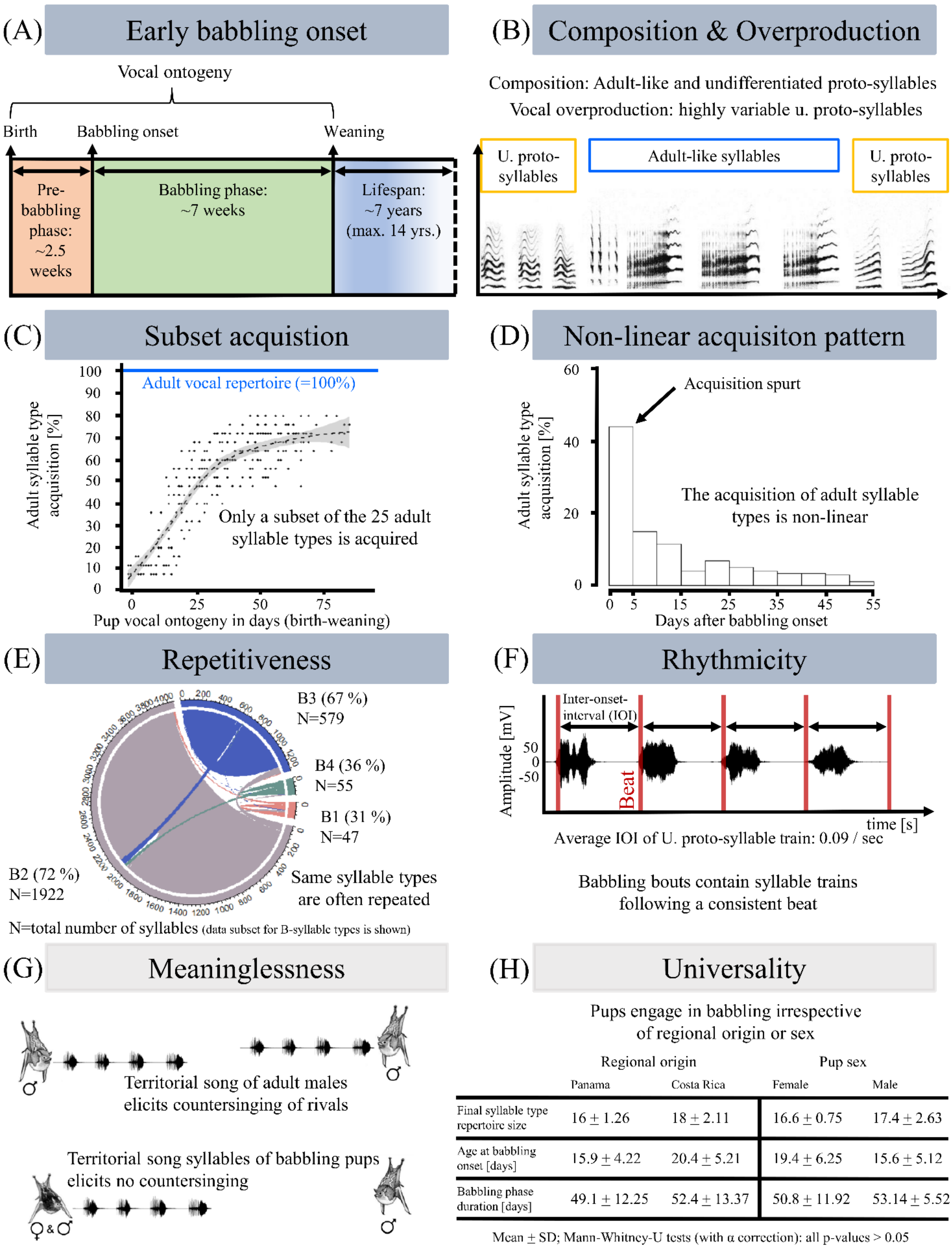
To assess whether trains had a regular beat based on the measured inter-onset-intervals, we calculated a normalized pairwise variability index (*nPVI)* for each train category. Syllable trains belonging to the neutral, undifferentiated proto-syllable and aggressive train category were the most isochronous trains in babbling bouts (Fig. 3*F*). The isolation call train category was less isochronous than the aforementioned train categories. The least isochronous trains belonged to the affiliative train category. The average nPVI and the mean inter-onset-intervals of the five syllable train categories are shown in the supporting information (*SI Appendix,* Table S3).

Babbling bouts get more diverse during ontogeny

Separate general linear mixed models (GLMMs) for each syllable train category showed that both the neutral and isolation call train categories were produced significantly less often with ongoing ontogeny (neutral train category: z=-5.489, p < 0.001; isolation train category: z=-15.93, p < 0.001). On the contrary, the undifferentiated proto-syllable, aggressive and affiliative train categories were produced significantly more often with ongoing ontogeny (undifferentiated proto-syllable train category: z=11.76, p < 0.001; aggressive train category: z=5.299, p < 0.001; affiliative train category: z=10.480, p < 0.001; *SI Appendix,* Fig. S4, Table S4). Results concerning the total number of trains per babbling bout and the train rate (i.e. the ratio of total number of trains divided by the respective bout duration) are reported in the supporting information. Both the total number of syllable trains and the syllable train rate significantly increased during ontogeny (*SI Appendix,* Table S5).

Level 3: General babbling bout features

On the bout level (N=221 babbling bouts of 20 pups) we investigated the influence of babbling bout duration on the potential outcome of the mother-pup interaction (i.e. nursing or not nursing). The duration of a babbling bout did not have an influence on whether the pup was subsequently nursed or not (*SI Appendix,* Fig. S6, Mann-Whitney-U test: W=4950, p=0.7). Moreover, we found that bout duration significantly increased with pup age, irrespective of sex or time of day in which babbling occurred (*SI Appendix,* Table S6).



	Regional origin		Pup sex	
	Panama	Costa Rica	Female	Male
Final syllable type repertoire size	16 ± 1.26	18 ± 2.11	16.6 ± 0.75	17.4 ± 2.63
Age at babbling onset [days]	15.9 ± 4.22	20.4 ± 5.21	19.4 ± 6.25	15.6 ± 5.12
Babbling phase duration [days]	49.1 ± 12.25	52.4 ± 13.37	50.8 ± 11.92	53.14 ± 5.52

Mean ± SD; Mann-Whitney-U tests (with α correction): all p-values > 0.05

Fig. 3. Babbling features of *S. bilineata*.

The same features are used to define human canonical babbling.

(*A*) Early babbling onset: The figure shows the entire vocal ontogeny of pups which is composed of the pre-babbling and the babbling phase. The babbling phase is on average three times longer than the pre-babbling phase. Thus, vocal practice occupies a considerable amount of time during pup development**.** The duration of the respective phases is displayed in weeks (averaged over 20 pups).

(*B*) Syllabic composition and vocal overproduction in babbling bouts. Babbling pups produce vocal sequences containing adult-like syllable types and highly variable undifferentiated proto- syllables (UP). The adult-like syllable types correspond well to the syllable types found in the adult vocal repertoire whereas the UP syllables are uniquely produced by pups. Undifferentiated proto-syllables are multiharmonic tonal syllables of variable duration and not corresponding to any syllable type of the adult vocal repertoire. Theyare the most common syllable type in babbling bouts (39 %). The high variability of UPsyllables and their graded nature leads to considerable vocal overproduction. The figure depicts an excerpt of a babbling bout composed of different adult-like and UPsyllables.

(*C*) Subset acquisition of the adult vocal repertoire. The figure depicts the rate of syllable type acquisition [%] as a function of pup age in days (from birth until weaning). The blue line represents the entire adult vocal repertoire (N=25 syllable types). Pups only acquired a subset of the adult vocal repertoire during ontogeny (max. 80% or 20 syllable types).

(*D*) Non-linear syllable type acquisition pattern. The figure depicts the rate of syllable type acquisition [%] as a function of pup age in days (depicted in 5-day units) during the babbling phase. Day 0 corresponds to the individual point in time at which each pup commences babbling (N=216 babbling bouts of 20 pups), which is at day 15 to 20 of age. Most syllable types are acquired during the first five days of babbling.

(*E*) Repetitiveness of the same syllable types. The figure depicts the transition probabilities of four adult-like syllable types belonging to the aggressive syllable category (B1-B4). These syllable types were chosen as an example to illustrate the repetitive nature of babbling bouts. Syllable types are represented by different colors; numbers (N) depict the total amount of syllables from each syllable type drawn from the original dataset. Each color has to be interpreted in a clockwise manner; starting from 0 to the middle, the color represents the transition probabilities to the succeeding syllable. Different colors pointing inwards (middle to max. number) represent the preceding syllable types. The thickness of the color arrays corresponds to the proportion of the respective transition probability within or between syllable types. The most dominant pattern depicted here is the self-repetition of syllable types (e.g. B1 is mainly succeeded by B1, B2 by B2, etc.).

(*F*) Rhythmicity. Syllable trains in babbling bouts have a regular rhythm. This isochronous beat (indicated by red bars in the oscillogram) was analyzed by measuring the inter-onset-intervals (black arrows) of syllables and comparing their lengths (N=30 syllable trains from 14 pups). The oscillogram depicts amplitude [mV] as a function of time [s].

(*G*) Meaninglessness. Babbling pups produce adult-like syllable types irrespective of the social context of a given situation, while adults produce specific syllable types for different social contexts (aggression, territoriality, courtship, etc.). Moreover, adults do not react to pups' adult-like syllable types, but they show a clear behavioral reaction when these syllable types are produced by adult conspecifics instead of pups. The figure exemplifies this scenario for territorial songs (composed of 3-6 syllable types). While territorial songs produced by adult males (drawings with light fur) elicit counter-singing from rivals, the same syllable type produced by babbling pups (drawing with dark fur) does not elicit any reaction. This feature cannot be weighted equally in comparison with features (*A*)-(*G*) (see discussion). Drawings by Dr. C.A.S. Mumm.

(*H*) Universality. Babbling is found in all *S. bilineata* pups, irrespective of their sex and regional origin. Final repertoire size, age at babbling onset and duration of the babbling phase are comparable for pups from Costa Rica and Panama as well as for male and female pups. The table gives means and SD for each population (N=10 pups each) and for male and female pups (7 males, 6 females, 7 pups were not sexed). This feature cannot be weighted equally in comparison with features (*A*)-(*G*) (see discussion).

DISCUSSION

The babbling behavior of *S. bilineata* pups shows the same features characteristic of infant CB (Table 2, Fig. 3). Similar to human infants, *S. bilineata* pups have a very distinct babbling phase during which they allocate a fair amount of their active time for vocal practice. CB in infants is universal irrespective of language or culture. Likewise, babbling in *S. bilineata* is present irrespective of sexand was observed in both populations from Panama and Costa Rica.

Babbling phase

Saccopteryx bilineata's babbling phase starts early and spans on average 74% of the entire pup development. Similarly, CB onset in infants is early during the child's development, around 7 months of age (Eilers et al., 1993; Oller, 1980). In human infants, language acquisition is constrained to a sensitive period, which is critical for learning (Kuhl, 2004; Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Newport, 1990, 2002). Studies showed that the learning capability of phonology, grammar and prosody decrease with age, probably in line with decreasing brain plasticity thereby explaining early CB onset in infants (Doupe & Kuhl, 1999). Cases in which children were deprived of language input and social care are thankfully rare but, those cases document that fully spoken language was not achieved when children had reached a certain age (Fromkin, Krashen, Curtiss, Rigler & Rigler, 1974; Newport, 1990, 2002). In human infants, CB onset is fairly sudden; this might be also due to the fact that marginal babbling (the stage preceding CB) lacks the characteristic feature of repetitiveness of later CB and thus might not be as easily recognized as babbling onset by parents (Oller, 1980). Likewise, babbling onset in *S. bilineata* is sudden and clearly detectable, both by the behavioral and vocal output of the pups.

The early babbling onset in pups might be explained by the same reason as in human infants. Brain plasticity might decrease with age, which could indicate a sensitive period for VPL. Furthermore, the adult vocal repertoire of *S. bilineata* is fairly complex, i.e. consists of many different syllable types that are combined in various ways into distinct vocalization types (Behr & von Helversen, 2004; Voigt et al., 2008). Rehearsing the different syllable types during a pronounced babbling phase might be necessary to gain control over the vocal apparatus to produce them correctly. Furthermore, the long babbling phase probably also increases the reinforcement of the pups' own neuronal template matching the sensory template through repeated perceptual input. In human infants, examples of hearing impaired children demonstrate how crucial auditory input is, both from hearing the own speech production and, speech from adult caregivers, in order to produce well-formed syllables. Without auditory input during vocal development infants do not develop speech (Oller & Eilers, 1988).

Composition of babbling bouts

During a babbling bout, on average 13 adult-like syllable types and the highly variable undifferentiated proto-syllables were produced (*SI Appendix,* Fig. S2, Fig. S8*B*). This composition of adult-like syllable types and exclusively juvenile undifferentiated proto-syllables is comparable to what can be observed in babbling infants. Likewise, CB in infants is composed of canonical syllables with adult phonetic properties (i.e. fully resonant vowels and adult-like temporal transition between vowel and consonant) (Oller, Griebel & Warlaumont, 2016) but also non-adult speech forms, so called protophones lacking adult syllable properties (Oller, 1980; Oller, Griebel & Warlaumont, 2016; Vihman, 2014). Protophones in infant speech include sounds which are not even speech like, such as yells, sequals, growls or raspberries. The production of non-speech sounds are probably related to refine control over the vocal apparatus by a playful manner (Oller et al., 2019). For example, yelling is related to learning how to control voice pitch and amplitude. Hence, the protophones probably result from general sound exploration (i.e. vocal play) and not yet fully developed control over the articulatory musculature (Oller, Griebel & Warlaumont, 2016). In *S. bilineata*, undifferentiated proto-syllables are highly variable and clearly distinct from the adult-like syllable types *(SI Appendix* Table S8), thus leading to vocal overproduction (i.e. syllables not present in the adult repertoire). Comparable to human infants, vocal overproduction might be caused by pups exploring their sound production. Undifferentiated proto-syllables were mainly located between different adult syllable types. Therefore, they might facilitate the production switch between two distinct syllable types. Furthermore, undifferentiated proto-syllables may have a labelling function for adult conspecifics, especially adult males. The mixture of adult-like and juvenile syllables constituting babbling bouts may identify them as ’harmless’ vocal practice although bouts contain syllables which would normally elicit aggressive reactions from adult males. Vocal overproduction is a phenomenon also known to occur in songbirds (Marler & Peters, 1982). Subsong is composed of a mixture of variants of adult syllables and non-adult vocalizations, which can be seen as a form of vocal overproduction (Doupe & Kuhl, 1999). Vocal overproduction in juvenile birds can also occur on the level of plastic song when juvenile songbirds overproduce entire songs of which only a few are retained after song crystallization, usually selected by social influences such as female preference or neighbor matching (Marler & Peters, 1982; Nelson, 2000; Nelson & Marler, 1994). In songbirds, overproduction might be a possible mechanism for innovation or serving as a potential reservoir for the invention of new songs in open-ended learners blackbird (Marler & Peters, 1982).

The composition of babbling bouts changed during ontogeny: isolation call syllable trains and neutral syllable trains were less frequently produced, whereas aggressive and affiliative syllable trains were more frequently produced (*SI Appendix,* Fig. S4, Table S4). Since the total number of trains per bout increased significantly with pup age (*SI Appendix,* Table S5), pups are not only changing babbling bout composition but, presumably, babbling gets more complex. Isolation calls are innate and neutral syllable types - including 5 syllable types from which 2 belong to the isolation call and thus are present from birth on – are short simple frequency modulated types (but see C2, *SI Appendix,* Fig. S10). The aggressive syllable types belonging to the later territorial song are acquired through VPL (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Whether the affiliative syllable types belonging to the later courtship song are also acquired through VPL still remains to be answered. Nevertheless, affiliative syllable types are mainly characterized by frequency modulations and long duration (*SI Appendix,* Fig. S10) which both probably involve increased vocal motor control compared to the simpler neutral syllable types. Therefore, the observed change of production of syllable trains involving more complex syllable types may be associated with increased motor control. Similarly, in human infants, the observed increase in phonetic complexity during the babbling phase is hypothesized to be associated with increased articulatory control of the vocal apparatus (Smith, Brown-Sweeney, & Stoel-Gammon, 1989). This finding further supports the importance of the babbling period as a vocal practice phase to acquire the adult vocal repertoire.

Acquisition process of the adult syllable type repertoire

Pups only acquired a subset of the adult syllable type repertoire (*SI Appendix,* Table S1). Similarly, CB in children does not contain the full set of phonemes yet. Some phonemes appear only at later stages during speech development (Vihman, 2014). In *S. bilineata* the syllable type acquisition pattern was non-linear. Within the first 10 days after babbling onset, pups acquired more than 50% of their individual final syllable type repertoire size (Fig. 3*D*). The pups' acquisition spurt of syllables has no parallels in infants' CB phase but may be comparable to the 'vocabulary spurt' which occurs at the lexical, i.e. word, level (Goldstein & Schwade, 2010; Lenneberg, 1967). Within a short time period the children's vocabulary increases exponentially from a former 50-word lexicon to a repertoire consisting of approximately 200 words.

The acquisition spurt was observable in each pup and, in most cases, the syllables of the territorial song were among the first to appear after babbling onset. Also, those were the syllables which all pups, irrespective of sex, acquired by imitating adult males which they heard singing on a daily basis (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). This is especially noteworthy since adult females do not produce territorial songs (nor courtship song); thus, many syllables acquired during babbling are not produced after females are weaned. What could be the possible advantage of this behavior for female pups? Females freely select their mating partner among the colony males (Nagy, Heckel, Voigt, & Mayer, 2007). Their mating decision is based on multiple sensory modalities, amongst which the auditory modality plays an important role (Behr & von Helversen, 2004; Voigt et al., 2008). The experience of producing the corresponding syllable types during babbling might influence the later mating choice, since females might be better able to assess the production costs or imitation quality of songs. Furthermore, female pups possibly use their neural template established during their babbling phase to compare it later to the acoustic output of adult males in their colonies.

Many affiliative syllable types appeared only later in the acquisition process or were completely missing (Table S1). A possible reason for this result could be related to the lack of templates. Most affiliative syllable types (12 out of 16) are part of the male courtship song. If affiliative syllable types belonging to the male courtship song are acquired through VPL, pups are probably not provided with abundant courtship song examples during ontogeny. Even though adult males court females on a daily basis they produce complete courtship songs exclusively during the mating season and thus use only a subset of the courtship song syllable types during the pups' ontogeny (Behr & von Helversen, 2004). Correspondingly, pups produce only a few affiliative syllable types which correspond to the most frequently used syllable types by courting males. It remains to be investigated when the missing affiliative syllable types appear in the young bats' repertoires; for instance, subadult males may acquire them during their first breeding season, either through VPL or a hormonally induced change in vocal production or both. Similar mechanisms have been described for songbirds (Bottjer & Johnson, 1997; Brainard & Doupe, 2002).

Coupling of repetitiveness and rhythmicity in babbling

Babbling bouts were clearly dominated by repetition of the same syllable type (Fig. 3*E*). In human infants the benefits of reduplication are to support vocal practice by repetition and possibly enhance the neuronal sensory-motor template by rehearsing oral production of syllables (Fagan, 2015). Those benefits might also be possible reasons for repetitiveness in pup babbling. Additionally, the repetition of syllables might also reflect, to a certain degree, the general perceptual impression pups gain during ontogeny. Many adult vocalization types contain repetitions of the same syllable type (Behr & von Helversen, 2004; Eckenweber & Knörnschild, 2013). Therefore, pups might also copy the repetitive composition of adult vocalizations during babbling. This repetitive character is coupled with the impression of rhythmicity when listening to babbling pups. Together, repetitiveness and rhythmicity are most characteristic for babbling in both human infants and *S. bilineata* pups. Rhythmicity in infant CB is usually ascribed from the impression the person gains when listening to a babbling infant or described qualitatively in relation to another rhythmic behavior such as hand-banging (Ejiri, 1998; Oller, 1980). The impression of rhythmicity in infant CB results from the strictly organized production of repetitive speech sounds (Oller, 1980). The syllables produced are a result of the infant's successive mandibular oscillations, which are alternated closed and open phases of the mouth, and function as an independent rhythm generator in babbling (Ejiri, 1998). Infant CB sequences are composed of repetitive vowel-consonant pairs that also facilitate a strict temporal pattern provoking an impression of rhythmicity in the listeners.

Our data showed that some syllabic sequences in babbling bouts that possess a regular beat (Fig. 3 *E)*. Likewise, we assume that the different rhythmicity of the syllable train categories in pups' babbling bouts is partially caused by the composition of syllable trains. Train categories composed of syllable types that are very similar in their spectro-temporal features should lead to more isochronous patterns. Correspondingly, the neutral and undifferentiated proto-syllable train categories had the most consistent isochronous rhythm (Fig. 3*F*), whereas the other syllable train categories had a less consistent temporal pattern (*SI Appendix,* Table S3). The more variable temporal structure of aggressive and affiliative syllable train categories is most comparable to variegated babbling (i.e. multisyllabic utterances with changing vowel or consonant or both).

Babbling is universal and meaningless

The feature universality is assigned to infant CB based on the fact that all infants studied so far show 1) an early CB onset, 2) a similar duration of the CB phase, 3) the occurrence of CB irrespective of sex or regional origin and 4) similarities in syllable type production during CB (Oller, 1980; Vihman, 1996). Our study demonstrates that all sampled pups engaged in babbling and that the aforementioned five parameters are very similar among babbling pups, irrespective of sex or regional origin (Fig. 3*H)*. However, our data is restricted to pups from two populations and further investigations of populations from other regions are necessary to fully corroborate the feature universality.

An interesting finding of our study is that both sexes babble. This is also the case in human infants and pygmy marmosets (Elowson, Snowdon, & Lazaro-Perea, 1998a,b; Vihman, 1996). In the latter two cases, it seems intuitive that male and female infants of both species babble since both produce the same adult vocal repertoire after infancy. It also seems intuitive that subsong and plastic song production is male-biased in songbirds, since only adult males are producing song in many species (but see increasing literature about female song, (Odom, Hall, Riebel, Omland & Langmore, 2014; Riebel, 2016; Riebel, Hall, & Langmore, 2005). In *S. bilineata*, however, male and female pups did not show any differences in babbling behavior, even though their vocal repertoire differs remarkably as adults (Behr & von Helversen, 2004; Knörnschild & Nagy, 2016). For male pups the reason of babbling is obvious (i.e. to practice the syllable types which are later produced during territorial defense and courtship). Why female pups practice syllable types which they never produce as adults remains an open question, as discussed before.

A characteristic feature of infant CB is its meaninglessness (Vihman, 1996). The phonetic units which are produced during infant CB lack semantics and only later in development, around 12 months of age, children connect word meaning with production (Vihman, 2014). It seems that in *S. bilineata,* babbling is meaningless as well. The production of distinct syllable types by babbling pups did not elicit the reaction it would elicit when produced by adults (e.g. adult males did not react to the aggressive syllable types occurring in babbling bouts, Fig. 3*G*). Furthermore, the succession of syllable types and production of the different syllable types in syllable trains was not consistent with any vocalization types produced by adults. Whether the lack of response by adult bats is explained by meaninglessness of babbling bouts - maybe related to the different succession and composition of syllable types within bouts – or whether it is explained by the fact that the sender is a pup (or a combination of the aforementioned points) remains a conundrum we cannot solve. Hence, like the feature universality, meaninglessness in pup babbling can neither be fully approved nor declined.

Table 2. Shared features between canonical babbling of human infants and babbling of bat pups.

Babbling features	Infant canonical babbling (CB)	Pup babbling in *S. bilineata*
(I) Early start of babbling	CB starts at an age of 6-7 months, irrespective of gender or culture; the onset is fairly sudden (Eilers et al., 1993).	Babbling starts at a mean age of two to three weeks.
(II) Emergence of adult-like syllables and non-speech sound	CB is composed of resonant, well-formed syllables (consonant-vowel pairs) and non-identifiable speech forms (Oller, 1980; Vihman, 1996).	Babbling is composed of adult-like syllable types and transition syllables exclusively produced by babbling pups.
(III) Vocal overproduction	CB contains non-adult speech forms which probably are produced while exploring the vocal apparatus (Oller et al., 2008; Vihman, 1996).	Undifferentiated proto-syllables are highly variable within and between pups; they are mainly produced when switching from one syllable type to another.
(IV) Subset acquisition	Only a subset of the speech sounds are present during the canonical babbling (Vihman, 2014; Vihman et al., 1986).	Only a subset of the adult syllable repertoire is acquired.
(V) Non-linear acquisition pattern	Neither the acquisition of consonants nor vowels follow a gradual acquisition pattern during the CB period (Vihman, 2014).	The acquisition of adult-like syllable types does not follow a linear pattern; most syllable types are acquired in a 5-day period after babbling onset.
(VI) Repetitiveness	CB is characterized by repetitions of the same syllable type, one of the main characteristics making CB recognizable also by non-parents (Oller, 1980; Vihman, 1996).	Babbling bouts are dominated by repetitions of the same syllable type.

(VII) Rhythmicity	Reduplicated syllable series show a relatively rigid timing. This rhythmicity is a product of the cyclic alternation of consonant and vowel like gestures which effect articulatory closure and opening (Kent et al., 1991).	Some syllable trains are highly rhythmic while others are not.
(VIII) Meaninglessness VIII in pup babbling cannot be fully compared with infant CB (see discussion)	CB is composed of meaningless speech units (Oller & Eilers, 1982).	Neither syllabic composition nor sequence of syllables in babbling bouts was consistent with adult vocalization types. The context in which syllable types were produced during babbling were not alike the context in which the same syllable types were produced by adults.
(IX) Universality IX in pup babbling cannot be fully compared with infant CB (see discussion)	CB is a mandatory phase to acquire speech and shows universal features irrespective of the language to be learned (Eilers et al., 1993; Oller & Eilers, 1982).	All listed features are present in each pup, irrespective of sex or regional origin.

Prolonged babbling is not a conflict over nursing

Females are perfectly able to refuse nursing attempts of pups and occasionally do so but babbling bout duration did not predict whether a pup was subsequently being nursed or not (*SI Appendix,* Fig. S6). Refusing pups' solicitation attempts, e.g. by hitting or biting the pup, is distinctly different from the non-aggressive, elusive or even inviting behaviors of mothers towards babbling pups, e.g. crawling around the pup (*SI Appendix,* Fig. S7). Moreover, mothers regularly initiated babbling bouts by hovering in front of their pups (*SI Appendix,* Fig. S7). When pups' need for nursing was high, i.e. after returning to the day-roost at dawn, they produced isolation calls instead of babbling bouts which caused mothers to nurse them immediately. Within a few days after detaching for the first time from the mother in the day-roost, pups are able to fly on their own (own observation and former studies, e.g. Bradbury &

Emmons, 1974; Tannenbaum, 1975). Fecal samples from pups demonstrated that even at this early age, insect carcasses are already present, indicating that pups started to hunt (Tannenbaum, 1975). Nevertheless, insect remains are not abundant, indicating that the hunting is not yet very successful (Tannenbaum, 1975). Pups only started to engage in long babbling bouts after returning at dawn during the last 2 to 3 weeks before they were weaned. This is probably an indication that their foraging success had increased. Therefore, we argue that solicitation of maternal care is, if anything, solely a minor purpose of babbling in *S. bilineata*.

The connection between vocal learning and babbling - two related traits?

Like human infants and songbirds, *S. bilineata* exhibits VPL (Knörnschild, 2014; Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010; Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012). This might elicit - at first glance - the conclusion that babbling is a necessary prerequisite for VPL. In human language development, the function of babbling is to learn how to produce the foundation of each spoken language – the phonetic repertoire (Oller, 1980; Vihman, 1996, 2014). In songbirds, the vocal ontogeny also includes a pronounced vocal practice phase (i.e. sensorimotor phase) to acquire the species-specific song. Songbirds first produce subsong, vocal practice composed of unstructured vocalizations which is suggested to map the motor commands for particular sound production to the memorized template (Marler, 1970; Marler & Peters, 1982). The next step includes the production of plastic song during which juvenile birds rehearse the learned material (Marler, 1970). Plastic song is characterized by features also present in infant CB, such as repetitiveness, well-formed adult like vocalizations, and meaninglessness (Doupe & Kuhl, 1999; Marler, 1970).

Yet, babbling and VPL are not necessarily connected. Pygmy marmosets are not capable of VPL but, nevertheless, engage in babbling behavior during infancy which shares many features with infant CB (Elowson, Snowdon, & Lazaro-Perea, 1998a, b). However, pygmy marmosets mainly seem to babble to solicit social care, e.g. being picked up or approached by a caregiver or grooming, and not to practice their vocal repertoire (Elowson, Snowdon, & Lazaro-Perea, 1998a,b). The association between social interactions and babbling remains unchanged throughout ontogeny and babbling infants are most likely to elicit a response from caregivers (Elowson, Snowdon, & Lazaro-Perea, 1998a,b).

Furthermore, the existence of VPL in a species does not imply that babbling has to occur during its ontogeny. VPL is also present in other bat species (e.g. Boughman, 1998), in elephants (e.g. Poole, Tyack, Stoeger-Horwath, & Watwood, 2005), seals (e.g. Reichmuth & Casey, 2014), cetaceans (Janik, 2014), parrots (e.g. Berg, Delgado, Cortopassi, Beissinger, & Bradbury, 2011) and hummingbirds (e.g. Baptista & Schuchmann, 1990). But, to our knowledge, descriptions of vocal practice behavior reminiscent of babbling are lacking for all but one of the aforementioned taxa (see Master thesis of T. Dolgushina for parrots; Tatiana, 2017). This lack of evidence may partly be caused by the challenges of describing the vocal ontogeny in some of those species, especially cetaceans.

An alternative explanation for the scarcity of babbling reports across taxa is that different circumstances have to occur for babbling to evolve. Firstly, if babbling is used as vocal practice to imitate a tutor, tutors must be present and producing the required vocalizations during juvenile ontogeny, which may not be the case in all VPL species. Secondly, babbling might be necessary if aggressive and/or courtship signals, which normally elicit aggressive responses in adults, need to be learned. Mixing adult vocalizations with juvenile ones during babbling may be less offensive, thus rendering vocal practice safe for juveniles. Thirdly, babbling might be necessary when adult vocal repertoires are large and thus difficult to acquire (i.e., large number of syllables, structured vocal sequences, etc.).

Conclusions

This study introduces bats as a highly promising taxon for comparative research on the ontogeny of VPL and provides the first qualitative comparison for quantitatively described babbling features across mammals, thus setting the stage for future quantitative comparisons. Studying the vocal ontogeny of *S. bilineata*, with its pronounced babbling phase reminiscent of infants' CB phase, will enable us to investigate the genetic requirements, neural substrates and behavioral adaptations involved in mammalian VPL in unprecedented ways. Future neurogenetic and genomic studies can explore the role of genes, such as FoxP2 and its associated molecular network, in babbling *S. bilineata* pups and thus make long pending comparisons to the molecular mechanisms of VPL in humans. Moreover, behavioral experiments can disentangle the importance of different factors influencing the speed and precision of VPL, such as maternal feedback and social interactions with peers and tutors. Together, these future approaches will allow us to decipher commonalities among mammals with VPL and to understand what may be unique to humans.

METHODS

Study sites and study species

Data were collected during two consecutive field seasons from two populations (Panama: May-August 2015, Costa Rica: May-September 2016; for details see *SI Appendix*). Acoustic recordings came from 20 pups (7 males, 6 females, 7 unsexed pups) belonging to 8 different colonies. Both populations are part of long-term projects and used to human presence. Therefore, after a habituation period bats were not noticeably disturbed by human presence in close vicinity to their roosts. This species occupies day-roosts in well-lit tree cavities and on man-made structures which facilitates behavioral observations and sound recordings. To ensure recordings of focal individuals, pups were first identified via their banded mothers (mothers only nurse and interact with their own pup) and subsequently banded themselves with colored-plastic rings on their forearms. Within the day-roost, adult individuals maintain an interindividual distance of 5 to 8 centimeters, thus facilitating data collection of mother-pup pairs (*SI Appendix,* Fig. S8).

Data collection

Acoustic recordings and behavioral observations were performed from birth until weaning. At least twice a week, focal pups were recorded ad libitum *sensu* Altmann (Altmann, 1974) during alternating morning (i.e. 6 am-noon) or afternoon (i.e. noon-6 pm) recording sessions. During a recording session, it was possible to record several babbling bouts per pup which allowed us to select babbling bouts with very good signal-to-noise ratio for subsequent analyses. Recordings were performed in close distance to the day-roost (2-4 meters). Please see the supporting information for more details on recordings and study animals.

Analyses of babbling bouts (levels 1-3)

Level 1

The day of birth was used to calculate pup age in general, age at babbling onset and the duration of the babbling phase (in days). Based on pup age we calculated if and at which point in time new adult-like syllable types were acquired. We determined the first and last day of babbling for each pup individually and calculated the *babbling phase duration* (in days). The point in time (i.e. number of days since babbling onset) during which new syllable types appeared during the babbling phase was used to calculate the syllable type acquisition pattern. Babbling bouts were analyzed weekly (including the very first and last babbling bout produced during the babbling phase) to investigate the composition of babbling bouts and pups' final syllable type repertoire sizes. Each bout was manually analysed and a) new syllable types and b) overall occurrence of syllable types was noted based on spectro-temporal similarities to the adult vocal syllable types (Fig. 3 (for detailed acoustic characteristics of each syllable type, see *SI Appendix*). The visual classification of syllables was performed with the software Avisoft SasLab Pro (Version 5.2.07, R. Specht, Avisoft Bioacoustics, Berlin, Germany) and verified by measuring acoustic parameters of a syllable subset (N=528 syllables) which were subsequently used to calculate discriminant function analyses (*SI Appendix,* Tables S9-S11). Automatic classification was not feasible because the recordings included background noise such as bird calls and sometimes overlapping vocalizations from several bats. To investigate repetitiveness in babbling bouts, we analyzed at least five babbling bout excerpts (between one and four minutes of duration) for 10 pups (Panama population) from the entire babbling phase (N=56,611 syllables). For each pup, we classified between 2,500 and 8,900 syllables based on the adult vocal repertoire. We then calculated transition probability matrices for each pup to obtain individual syllable transition probabilities. From those 56,611 classified syllable types, the proportion of undifferentiated proto-syllables was calculated to investigate their occurrence in babbling. The aforementioned analyses were used to test whether there were significant differences between regions and sexes.

Level 2

A total of 8,448 syllable trains from 180 babbling bouts of 20 pups (at least one bout per week and pup) were classified into one of five different syllable train categories based on syllable type composition; neutral -, aggressive -, affiliative -, undifferentiated proto-, and isolation call category (see supporting information).

Rhythmicity in babbling

For each train, the inter-onset intervals (IOIs) of its syllables were defined manually based on oscillograms. We analyzed 30 syllable trains (three per category and per population). Each train was composed of at least five syllables. IOIs are defined as the duration between the start of a syllable and the start of the subsequent one. Then, the mean IOI and the standard deviation for each train category were calculated to assess rhythmic consistency within a syllable train. Furthermore, the normalized pairwise variability index (nPVI) was calculated to assess temporal variance in trains and whether a train was isochronous. The nPVI indicates the goodness of predictability from one IOI to the subsequent. The lower the nPVI the more isochronous is the syllabic train (0 = metronomic isochrony, all IOIs are equal).

Change of babbling bout composition during the babbling phase

To investigate the change in the occurrence of syllable train categories (i.e. babbling bout composition), babbling phases of each pup were split into 5-day units and the proportions of each different train category was calculated and compared (data was pooled over all pups). Furthermore, we investigated the change of total number of syllable trains in babbling bouts and syllable train rate (babbling bout duration divided by the number of syllable trains) during the babbling phase.

Level 3

To investigate if and how babbling bout duration was influenced by pup age, pup sex and daytime at which babbling occurred we analyzed 221 babbling bouts from 20 pups. Furthermore, we tested if there was a significant difference in bout duration between babbling bouts which led to nursing and which did not (66 bouts not succeeded by nursing, 155 bouts succeeded by nursing).

Statistical analyses

Multiple testing was accounted for with sequential Bonferroni corrections. Statistical differences were considered significant for $p < 0.05$ (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). All statistical analyses were performed in R (RStudio 2018, version 3.5.2 (Team, 2015)).

Level 1: To calculate differences between populations or pup sex (feature IX), we used non-parametric Mann-Whitney-U tests.

Level 2: To investigate whether and how the babbling bout composition (i.e. syllable train category production) changed during the babbling phase, separate GLMMs were performed for each syllable train category (babbling phase duration in days as a covariate, pup ID as random factor, Poisson family with log link function). To analyze the change of syllable train number and syllable train rate, two separate GLMMs were performed (babbling phase duration in days as covariate, pup sex as fixed factor, pup ID as random factor, Gamma family with log link).

Level 3: A GLMM was performed to analyze the influence of pup age, pup sex and daytime on babbling bout duration (pup age as covariate, pup sex and daytime of babbling as fixed factors, pup ID as random factor, Gamma family with log link). A Mann-Whitney-U test was performed to test whether there was a difference in babbling bout duration depending on whether the pup was subsequently nursed or not (*SI Appendix,* Fig. S6).

AUTHOR CONTRIBUTIONS

A.A.F. and M.K. designed the study, A.A.F. collected the data in the field, A.A.F, analyzed the data, L.S.B. performed the rhythm analysis, A.A.F. and M.K. wrote the paper.

ACKNOWLEDGEMENTS

We thank the Smithsonian Tropical Research Institute and the field station Barro Colorado Island for providing excellent research conditions. Furthermore, we want to thank the Natural Reserve Curú in Costa Rica for providing us with the possibility to use their accommodation during our field stay. We thank J. Benichov, M. Nagy, M. Metz, L. Günther and S. Ripperger for valuable comments on the manuscript. Many thanks to S. Hiller for support and advice on the coding of the circlize plot and the realization of the supplementary video. Funding: This study was financed by a stipend from the Elsa-Neumann Foundation (A.A.F. and L.S.B.) and a Heisenberg Fellowship (DFG KN935 3-1) and a research grant (DFG KN935 4-1) from the German Research Foundation (M.K.).

REFERENCES

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018). The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science, 21*(5), e12641. doi:10.1111/desc.12641.

Altmann, J. (1974). Obsrevational study of behaviour: sampling methods. *Behaviour, 49*, 227-267.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current directions in psychological science, 21*(3), 170-176.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological science, 9*(4), 321-324.

Baptista, L. F., & Schuchmann, K. L. (1990). Song learning in the Anna hummingbird (*Calypte anna). Ethology, 84*(1), 15-26.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in ecology & evolution, 20*(3), 143-149.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song learning. *Current Directions in Psychological Science, 13*(6), 224-228.

Behr, O., Knörnschild, M., & von Helversen, O. (2009). Territorial counter-singing in male sac-winged bats (*Saccopteryx bilineata*): low-frequency songs trigger a stronger response. *Behavioral Ecology and Sociobiology, 63*(3), 433-442.

Behr, O., & von Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology, 56*(2), 106-115. doi:10.1007/s00265-004-0768-7.

Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R., & Bradbury, J. W. (2011). Vertical transmission of learned signatures in a wild parrot. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 585-591.

Boeckx, C., Leivada, E., & Martins, P. T. (2013). Language and complexity considerations: A biolinguistic perspective. *LSC-Llengua, Societat i Comunicació, 2013, vol. 11, p. 20-26.*

Boeckx, C., & Piattelli-Palmarini, M. (2005). Language as a natural object–linguistics as a natural science. *The linguistic review, 22*(2-4), 447-466.

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

Bottjer, S. W., & Johnson, F. (1997). Circuits, hormones, and learning: vocal behavior in songbirds. *Journal of neurobiology, 33*(5), 602-618.

Boughman, J. W. (1998). Vocal learning by greater spear–nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 265*(1392), 227-233.

Bradbury, J. W., & Emmons, L. H. (1974). Social organization of some Trinidad bats: I. Emballonuridae. *Zeitschrift für Tierpsychologie, 36*(1-5), 137-183.

Bradbury, J. W., & Vehrencamp, S. (1977). Social organization and foraging in emballonurid bats. *Behavioral Ecology and Sociobiology, 2*(1), 1-17.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature, 417*(6886), 351.

Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology, 29*(4), 631-636. e633.

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. doi: 10.1644/1545-1542(2002)083<0526:Gaiviv>2.0.Co;2.

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

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

Eilers, R. E., Oller, D. K., Levine, S., Basinger, D., Lynch, M. P., & Urbano, R. (1993). The role of prematurity and socioeconomic status in the onset of canonical babbling in infants. *Infant Behavior and Development, 16*(3), 297-315.

Ejiri, K. (1998). Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica, 55*(4), 226-237.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998)a. 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends in Cognitive Sciences, 2*(1), 31-37. Retrieved from doi: 10.1016/S1364-6613(97)01115-7.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998)b. Infant'babbling'in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour, 135*(5), 643-664.

Evans, C., & Kleindorfer, S. (2016). Superb fairy-wren (*Malurus cyaneus*) sons and daughters acquire song elements of mothers and social fathers. *Frontiers in Ecology and Evolution, 4*, 9.

Fagan, M. K. (2015). Why repetition? Repetitive babbling, auditory feedback, and cochlear implantation. *Journal of Experimental Child Psychology, 137*, 125-136. doi:10.1016/j.jecp.2015.04.005.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences, 4*(7), 258-267.

Fitch, W. T. (2009). The evolution of language: a comparative perspective. In *The Oxford handbook of psycholinguistics*.

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

Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annual Review of Linguistics, 4*, 255-279.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi:10.1016/j.cognition.2005.02.005

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the “critical period”. *Brain and language, 1*(1), 81-107.

Garamszegi, L. Z., Pavlova, D. Z., Eens, M., & Møller, A. P. (2006). The evolution of song in female birds in Europe. *Behavioral Ecology, 18*(1), 86-96.

Geberzahn, N., & Gahr, M. (2013). Song learning in male and female Uraeginthus cyanocephalus, a tropical songbird species. *Journal of Comparative Psychology, 127*(4), 352.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological science, 19*(5), 515-523.

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. *The Oxford handbook of developmental and comparative neuroscience*, 708-729.

Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*(1), 52.

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), 1569-1579. doi: 10.1126/science.298.5598.1569.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology, 28*, 60-65.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016*(1), 749-777.

Kent, R. D., Mitchell, P. R., & Sancier, M. (1991). Evidence and role of rhythmic organization in early vocal development in human infants. In *Advances in psychology* (Vol. 81, pp. 135-149): Elsevier.

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

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*(9), 451-454. doi:10.1007/s00114-006-0127-9.

Knörnschild, M., Blüml, S., Steidl, P., Eckenweber, M., & Nagy, M. (2017). Bat songs as acoustic beacons-male territorial songs attract dispersing females. *Scientific reports, 7*(1), 13918.

Knörnschild, M., & Nagy, M. (2016). Sex-biased dispersal and social systems of Neotropical Emballonurids. In J. Ortega (Ed.), *Sociality in bats* (pp. 47-63): Springer International Publishing Switzerland.

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. doi:10.1098/rsbl.2009.0685.

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*(4), 761-769. doi:10.1016/j.anbehav.2012.06.029.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Konishi, M. (1985). Birdsong: from behavior to neuron. *Annual Review of Neuroscience, 8*(1), 125-170.

Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(17), 9645-9646. doi:10.1073/pnas.1733998100.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831-843. doi:10.1038/nrn1533.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early Speech Perception and Later Language Development: Implications for the "Critical Period". *Language Learning and Development, 1*(3-4), 237-264. doi:10.1080/15475441.2005.9671948.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2*(12), 59-67.

Locke, J. L., & Pearson, D. M. (2009). Linguistic significance of babbling: evidence from a tracheostomized infant. *Journal of Child Language, 17*(01). doi:10.1017/s0305000900013076

Mahr, K., Seifert, C. L., & Hoi, H. (2016). Female and male Blue Tits (*Cyanistes caeruleus*) sing in response to experimental predator exposition. *Journal of Ornithology, 157*(3), 907-911.

Margoliash, D., & Nusbaum, H. C. (2009). Language: the perspective from a organismal biology. *Trends in Cognitive Sciences, 13*(12), 505-510. doi:10.1016/j.tics.2009.10.003.

Marler, P. (1970). Birdsong and Speech Development - Could There Be Parallels. *American Scientist, 58*(6), 669-673.

Marler, P., & Peters, S. (1982). Developmental overproduction and selective attrition: new processes in the epigenesis of birdsong. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 15*(4), 369-378.

Martins, P. T., & Boeckx, C. (2016). What we talk about when we talk about biolinguistics. *Linguistics Vanguard, 2*(1).

Misyak, J. B., Goldstein, M. H., & Christiansen, M. H. (2012). Statistical-sequential learning in development. *Statistical learning and language acquisition*, 13-54.

Nagy, M., Heckel, G., Voigt, C. C., & Mayer, F. (2007). Female-biased dispersal and patrilocal kin groups in a mammal with resource-defence polygyny. *Proceedings of the Royal Society B: Biological Sciences, 274*(1628), 3019-3025. doi:10.1098/rspb.2007.1008.

Nelson, D. A. (2000). Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour, 60*(6), 887-898.

Nelson, D. A., & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences, 91*(22), 10498-10501.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive science, 14*(1), 11-28.

Newport, E. L. (2002). Critical periods in language development.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nature Communications, 5*, 3379.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In J. F. K. a. C. A. F. G. H. Yeni-Komshian (Ed.), *Child phonology* (Vol. 1, pp. 93-112). New York: Academic Press, Inc.

Oller, D. K., & Eilers, R. E. (1982). Similarity of babbling in Spanish-and English-learning babies. *Journal of Child Language, 9*(3), 565-577.

Oller, D. K., & Eilers, R. E. (1988). The Role of Audition in Infant Babbling. *Child Development, 59*(2), 441-449. doi:Doi 10.2307/1130323.

Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., & Call, J. (2019). Language origins viewed in spontaneous and interactive vocal rates of human and bonobo infants. *Frontiers in psychology, 10*, 729.

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal Development as a Guide to Modeling the Evolution of Language. *Top Cogn Sci, 8*(2), 382-392. doi:10.1111/tops.12198.

Oller, D. K., Wieman, L. A., Doyle, W. J., & Ross, C. (2008). Infant babbling and speech. *Journal of Child Language, 3*(01). doi:10.1017/s0305000900001276.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature, 434*(7032), 455.

Prather, J. F., Okanoya, K., & Bolhuis, J. J. (2017). Brains for birds and babies: Neural parallels between birdsong and speech acquisition. *Neuroscience & Biobehavioral Reviews, 81*, 225-237.

Pytte, C. L., & Suthers, R. A. (2000). Sensitive period for sensorimotor integration during vocal motor learning. *Journal of neurobiology, 42*(2), 172-189.

Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology, 28*, 66-71.

Riebel, K. (2016). Understanding sex differences in form and function of bird song: The importance of studying song learning processes. *Frontiers in Ecology and Evolution, 4*, 62.

Riebel, K., Hall, M. L., & Langmore, N. E. (2005). Female songbirds still struggling to be heard. *Trends in ecology & evolution, 20*(8), 419-420.

Roper, M. M., Harmer, A. M., & Brunton, D. H. (2018). Developmental changes in song production in free-living male and female New Zealand bellbirds. *Animal Behaviour, 140*, 57-71.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Smith, B. L., Brown-Sweeney, S., & Stoel-Gammon, C. (1989). A quantitative analysis of reduplicated and variegated babbling. *First Language, 9*(6), 175-189.

Stoll, S., Lieven, E., Winskel, H., & Padakannaya, P. (2013). Studying language acquisition cross-linguistically. In *South and Southeast Asian Psycholinguistics* (pp. 19-35).

Strauss, M., von Helversen, O., & Knörnschild, M. (2010). The ontogeny of courtship behaviours in bat pups (*Saccopteryx bilineata*). *Behaviour, 147*(5), 661-676.

Tannenbaum, B. R. (1975). *Reproductive strategies in the white-lined bat.* (PhD). Cornell University Ithaca NY USA,

Tatiana, D. (2017). *The vocal babbling behavior and its sibling effects in a wild parrot.* (Masterthesis). University of Texas Rio Grande Valley, ProQuest LLC.

Team, R. (2015). RStudio: Integrated Development for R. (Version 3.5.2): RStudio, Inc., Boston Retrieved from MA URL

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*: Blackwell Publishing.

Vihman, M. M. (2014). *Phonological development: The first two years*: Wiley-Blackwell Boston, MA.

Vihman, M. M., Ferguson, C. A., & Elbert, M. (1986). Phonological Development from Babbling to Speech - Common Tendencies and Individual-Differences. *Applied Psycholinguistics, 7*(1), 3-40. doi:10.1017/S0142716400007165.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat, *Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

West, M. J., & King, A. P. (1988). Vocalizations of juvenile cowbirds (*Molothrus ater ater*) evoke copulatory responses from females. *Developmental Psychobiology, 21*(6), 543-552.

Supplementary Information for

Babbling in bat pups and human infants: common ontogenetic features of vocal production learners

Authors: Ahana A. Fernandez, Lara S. Burchardt & Mirjam Knörnschild

Table of Contents:

The video excerpt shows a babbling pup (including a spectrogram depicting the syllables as the pup is vocalizing).

Index of Figures

Index of Tables

Results

Methods

Supporting Results

Level 1: Syllable level

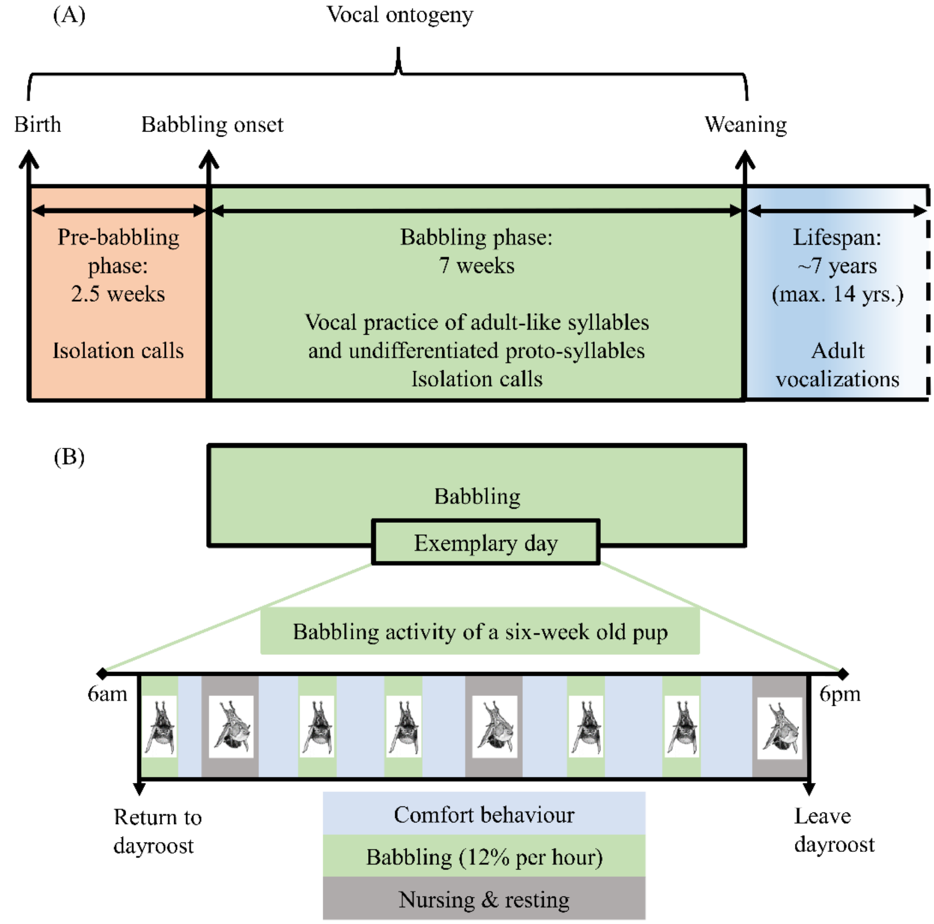


Fig. S1. Schematic vocal development of *S. bilineata* pups.

(***A***) The vocal ontogeny of *S. bilineata* pups is composed of two main phases, the pre-babbling phase (orange color) and the babbling phase (green color; mean duration in weeks of each phase in brackets, averaged over 20 pups). The pronounced babbling phase is around three times longer than the pre-babbling phase. During the pre-babbling phase, pups produce multisyllabic isolation calls, which are also present later in life. During the babbling phase, pups produce long vocal sequences composed of adult-like syllables and highly variable transition syllables. After weaning, many subadult males remain in their natal colony, whereas all subadult females disperse to new colonies. Adults have an average lifespan of 7 years (blue color). (***B***) An exemplary day of a six week-old pup during the babbling phase is depicted. Adult bats and pups return at dawn to their day-roost. During the day, pups spend a fair amount of their active time with babbling (green color): 12% per hour (i.e. average babbling bout duration within 60 minutes). On average, each pup produces five babbling bouts per day, which can last up to 43 minutes (mean: 7 min). When not babbling, pups engage in comfort behaviors such as grooming or roost without body contact to their mothers (light blue color). For approximately three hours per day, pups rest and nurse attached to their mothers' bellies (green color). At dusk, adult bats and pups leave the day-roost for foraging.

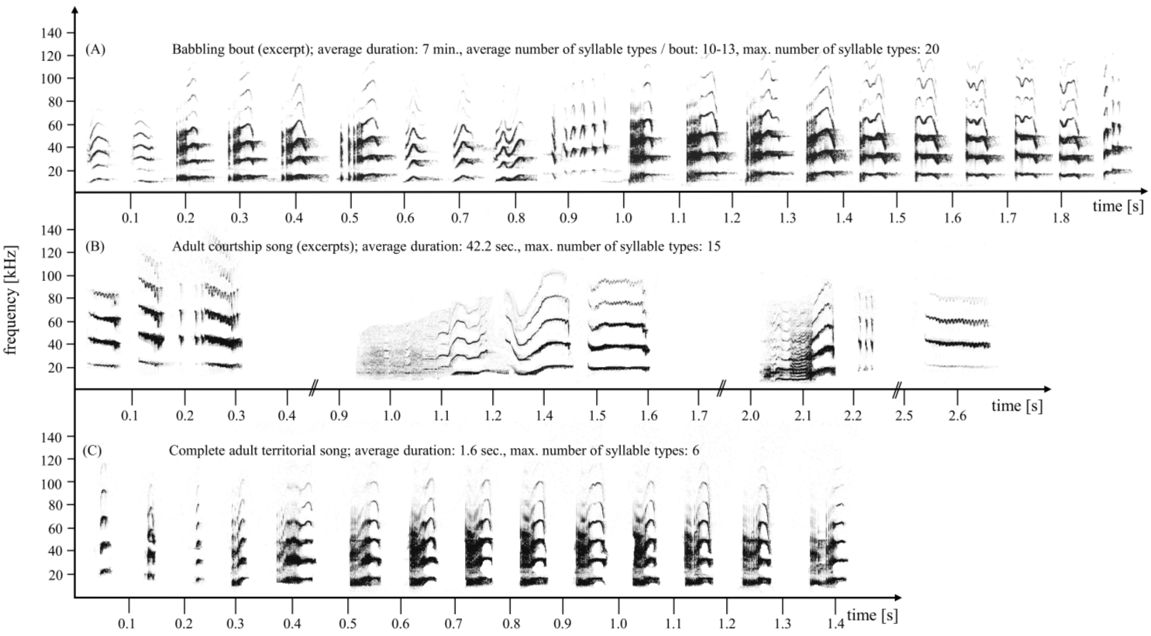


Fig. S2. Syllabic composition of babbling bouts in comparison with adult song of *S. bilineata.*

Two extraordinary characteristics of babbling are the long duration of babbling bouts (max. duration 43 minutes) and the multisyllabic composition. This stands in strong contrast to adult bat vocalizations which are in general only a few seconds long and composed of fewer syllable types. (***A***) The upmost panel shows an excerpt of a babbling bout: Babbling bouts are produced by pups (data from 20 pups, N: 221 babbling bouts). (***B***) The mid panel depicts four excerpts of courtship song. Courtship songs are produced by adult males, composed of up to 15 different syllable types and on average less than one minute long (Behr & von Helversen, 2004). (***C***) The lowest panel depicts a complete territorial song. Territorial songs are produced by adult males and on average 1.6 seconds long. With maximal six different syllable types, they have a comparatively simple syllabic composition (Behr & von Helversen, 2004).

Here we report further results of the analyses concerning the babbling features (I), (IV), and (IX).

(I) Babbling onset:

In both populations, the pre-babbling phase was significantly shorter than the babbling phase (Mann-Whitney-U test; Panama: W=0, $p<0.001$, N=10 pups; Costa Rica: W=2, $p<0.0001$, N=10 pups).

(IV) Subset acquisition of the adult syllable type repertoire

All pups acquired only a subset of the adults' vocal repertoire (column 5, Table S1). The highest number of missing (i.e. not acquired) syllable types in all pups' final syllable type repertoires were always within the affiliative syllable type category (column 6, Table S1). This table includes only the 20 syllable types which were found in the pups' final syllable type repertoires. The five syllable types (Screech: aggressive syllable type; Bark: neutral syllable type; T4, T5, W3: affiliative syllable types) which were missing in all pups at weaning are not included. The transition syllable is not included either since it is an exclusive pup vocalization only present during the babbling phase.

Table S1. Sizes of pups' individual syllable type repertoires at weaning.

ID	Sex	Population	Colony	Final repertoire size	Missing syllable types		
					Affiliative (N=14)	Aggressive (N=6)	Neutral (N=5)
1	UK	Panama	SL1	16	6	2	1
2	M	Panama	SL1	20	3	1	1
3	UK	Panama	SL2	18	5	1	1
4	F	Panama	SL2	17	6	1	1
5	UK	Panama	E1	18	5	1	1
6	M	Panama	E1	19	4	1	1
7	F	Panama	E3	17	6	1	1
8	M	Panama	E3	18	5	1	1
9	F	Panama	F1	16	7	1	1
10	UK	Panama	F1	17	5	2	1
11	UK	Costa Rica	B	18	4	2	1
12	M	Costa Rica	B	17	6	1	1
13	F	Costa Rica	B	16	6	1	2
14	UK	Costa Rica	TO	15	7	2	1
15	M	Costa Rica	TO	12	10	1	2
16	UK	Costa Rica	TR	19	4	1	1
17	F	Costa Rica	TR	15	7	1	2
18	F	Costa Rica	TR	16	6	2	1
19	M	Costa Rica	TR	19	3	2	1
20	M	Costa Rica	TR	17	5	2	1

Individual sizes of syllable type repertoires from 20 pups (N=216 babbling bouts). Columns (C) 1-4 contain focal pups' information including pup ID (C1), the pups' sex (UK=unknown pup sex; F=female pup; M=male pup, C2), the regional origin, (C3) and the pups' colony affiliation (C4). C5 represents the final repertoire size whereas C6-8 depict the number of missing syllable types in the pups' final repertoire according to the adult syllable type category to which they belong. The numbers in brackets represent the total number of adult syllable types in the corresponding category (affiliative, aggressive and neutral).

(IX) Universality

Irrespective of regional origin or sex, all pups produced babbling bouts. No significant differences could be detected concerning the final repertoire size, the age at babbling onset or the babbling phase duration (Table S2).

Table S2. Universality of babbling characteristics.

	Regional origin		**Pup sex**	
	Panama	Costa Rica	Female	Male
Final syllable type repertoire size	16 ± 1.26 (16-20)	18 ± 2.11 (12-19)	16.6 ± 0.75 (15-17)	17.4 ± 2.63 (12-20)
	W=32.5, p=0.19 (ns)		W=34, p=0.06 (with α corr.: ns)	
Age at babbling onset [days]	15.9 ± 4.22 (11-24)	20.4 ± 5.211 (10-28)	19.4 ± 6.25 (10-25)	15.6 ± 5.12 (14-28)
	W=24, p=0.053 (ns)		W=31, p=0.17 (with α corr.: ns)	
Babbling phase duration [days]	49.1 ± 12.25 (29-68)	52.4 ± 13.37 (24-1)	50.8 ± 11.92 (35-71)	53.14 ± 5.52 (42-59)
	W=42, p=0.56 (ns)		W=14, p=0.34 (with α corr.: ns)	

Mean ± SD and range (in brackets) of the pups final syllable type repertoire size, babbling phase duration and age at onset of babbling are given (Panama: N=127 babbling bouts of 10 pups, Costa Rica: N=89 babbling bouts of 10 pups). Mann-Whitney-U tests were all non-significant (with α corr.); Regions: N=10 pups each, female pups: N=6, male pups: N=7 pups.

Level 2: Syllable train level

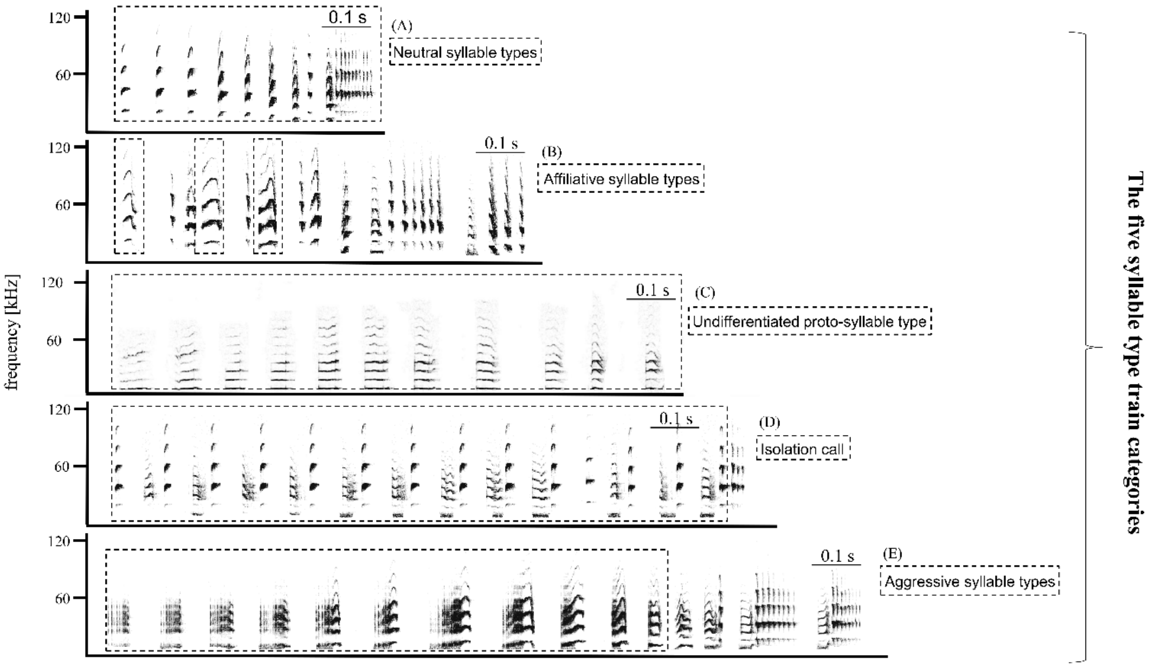


Fig. S3. Syllable type train categories.

Syllable trains are composed of at least five syllables. Trains are separated from each other by silent intervals, which are ≤ 300 ms. We established five different train categories, namely the *neutral* -, *aggressive* -, *affiliative* -, *undifferentiated proto-syllable*- and *isolation call* – syllable type train category. The name-giving functions of train categories were inferred from the use of syllable types in the adult vocal repertoire. Each bout was scanned visually and classification of train categories was obtained manually. Syllable trains were categorized according to the most abundant syllable type occurring. There was one exception: since affiliative syllable types were produced much more seldom than aggressive or transition syllables, syllable trains containing at least two affiliative syllable types were classified as *affiliative trains*, even though affiliative syllables were not the most abundant syllables in them. This was done because affiliative syllable types were produced much more rarely than other syllable types and we wanted to give their occurrence its due credit.The respective syllable types after which the train category is nominated are marked with a dashed box. The first panel (***A***) shows a neutral syllable type train, which is composed of three neutral syllable types (VS, VSV, C1). The second panel (***B***) shows an affiliative syllable type train, which is composed of neutral and affiliative syllable types (D1, VS, C1, ES). The third panel (***C***) depicts a undifferentiated proto-syllable type train, which is entirely composed of the variable undifferentiated proto-syllable. The fourth panel (***D***) depicts an isolation call train composed of end syllables of the isolation call (dashed box) and chatter. The lowest panel (***E***) depicts an aggressive syllable type train mainly composed of buzz syllables (and additionally containing a transition syllable, a few end syllables (ES) and chatter (C1). The spectrograms depict frequency [kHz] as a function of time [s] and were generated using a 1024 point fast Fourier transform and a Hamming window with 85% overlap.

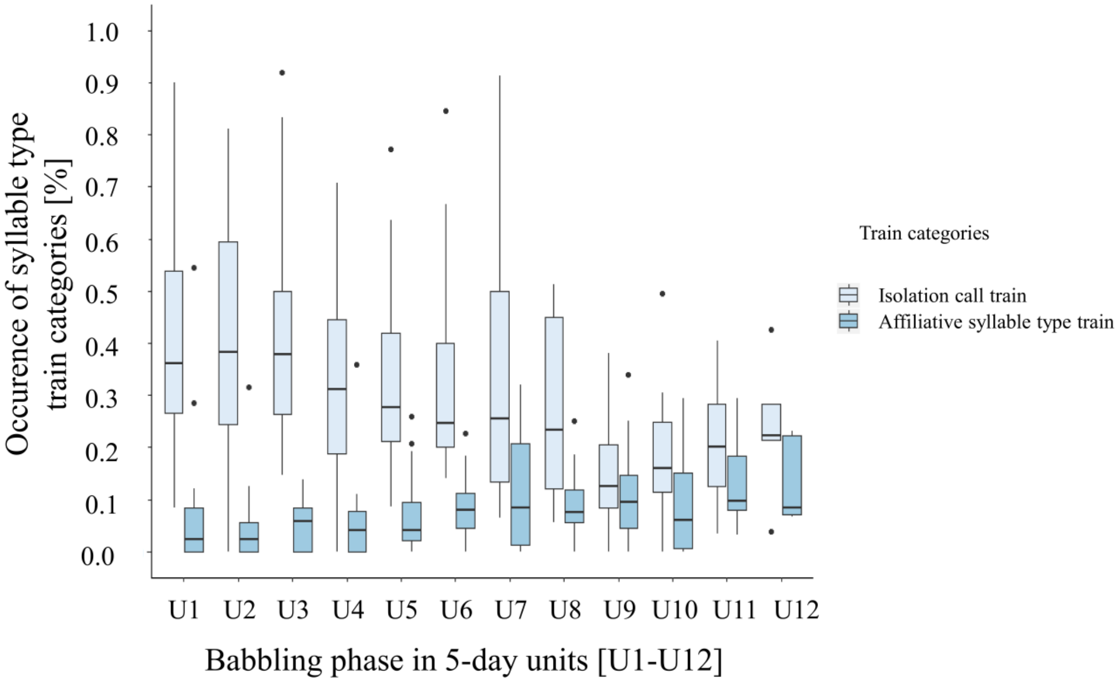


Fig. S4. Syllabic train composition change during ontogeny.

Frequency of occurrence (%) of two syllable type train categories (affiliative syllable type train category=dark blue, isolation call train category=light blue) changes during ontogeny (Results Level 2, see Table S4). While syllable trains mainly composed of isolation calls become less frequent during ontogeny, syllable trains mainly composed of affiliative syllables become more frequent. The entire babbling phase was divided into 5-day units (U1-U12). U1 starts at the first day of babbling, the last day of U12 is the day at weaning (N=8,448 trains from 20 pups).

Here we report further results on rhythmicity (feature VII), ontogenetic changes of babbling bouts and changes in syllable train production during ontogeny.

(VII) Rhythmicity

Table S3. Rhythmicity of syllable type trains in babbling bouts.

Syllable Type Train category	Mean IOI [sec]	SD	Variance	Range [sec]	Mean nPVI	Range nPVI
Isolation call syllable type	0.05	0.021	0.0004	0.008-0.108	36.7	10.8-51.12
Neutral syllable type	0.07	0.027	0.0007	0.011-0.229	21.5	4.8-46.4
U. proto-syllable type	0.09	0.046	0.0021	0.009-0.279	22.3	8.03-33.9
Aggressive syllable type	0.09	0.059	0.0035	0.007-0.262	35.3	9.8-81.9
Affiliative syllable type	0.07	0.045	0.0020	0.009-0.279	58.9	22.1-111.9

Inter-onset-intervals (IOI) of the syllable train categories were measured to assess rhythmicity within a syllable train. We calculated the normalized pairwise variability index (nPVI) to assess whether a syllable train had a regular beat. A sequence with a perfectly regular beat (i.e. isochrony) would have an nPVI of zero (N=30 syllable trains from 14 pups: 7 from each region). U=abbreviation for “undifferentiated”.

Babbling gets more diverse during ontogeny

The production rate of the five syllable type train categories changed during the babbling phase (Table S4, Fig. S4).

Table S4. Babbling bouts get more diverse during ontogeny.

Dependent variable (syllable train category)	Estimate	z-value	p-value	Trajectory
Isolation syllable type	-0.069332	-15.93	< 0.001	Decreasing ***
Neutral syllable type	-0.033933	-5.489	< 0.001	Decreasing ***
U. proto-syllable type	0.060772	11.76	< 0.001	Increasing ***
Aggressive syllable type	0.025378	5.299	< 0.001	Increasing ***
Affiliative syllables type	0.082352	10.480	< 0.001	Increasing ***

Changes in production rate of the five syllable train categories indicated an ontogenetic change of babbling bout composition. We calculated GLMMs (Poisson family with log link) with babbling phase duration in days as covariate and pup ID as random factor. Data was collected from N=8,448 syllable trains of 20 pups. U=abbreviation for “undifferentiated”.

Train production becomes more frequent during ontogeny

Table S5. Syllable train characteristics change during the babbling phase.

Dependent variable	Covariates Fixed factors	Estimate	t-value	p-value	Trajectory
Total number of trains / bout	Babbling phase duration [days]	0.016868	5.282	<0.001	Increasing ***
	Female	-0.051808	-0.368	0.713	NS
	Male	-0.021384	-0.162	0.872	NS
Train quotient (trains per babbling bout duration)	Babbling phase duration [days]	0.006558	3.089	0.002	Increasing**
	Female	-0.113233	-0.965	0.334	NS
	Male	0.031976	0.294	0.768	NS

With increasing age, pups produced more syllable trains and with a higher rate. We calculated GLMMs (Gamma family with log link) with babbling phase duration in days as covariate, pup sex as fixed factor and pup ID as random factor. Data was collected from N=8,448 syllable trains of 20 pups. The *syllable train* quotient (rate of trains per bout) increased significantly during the babbling phase (GLMM with Gamma distribution and log link, t=3.089, p=0.0020, N=8,448 trains) but without any difference between male and female pups (GLMM with Gamma distribution and log link, females: t=-0.965, p=0.33; males: t=-0.294, p=0.76). Correspondingly, the number of trains per babbling bout also increased during ontogeny (GLMM with Gamma distribution and log link, t=5.282, p<0.001, N=8448 trains), with no difference between male and female pups (GLMM with Gamma distribution and log link, females: t=-0.368, p=0.71, males: t=-0.162, p=0.87, Table S5).

Level 3: Babbling bout level

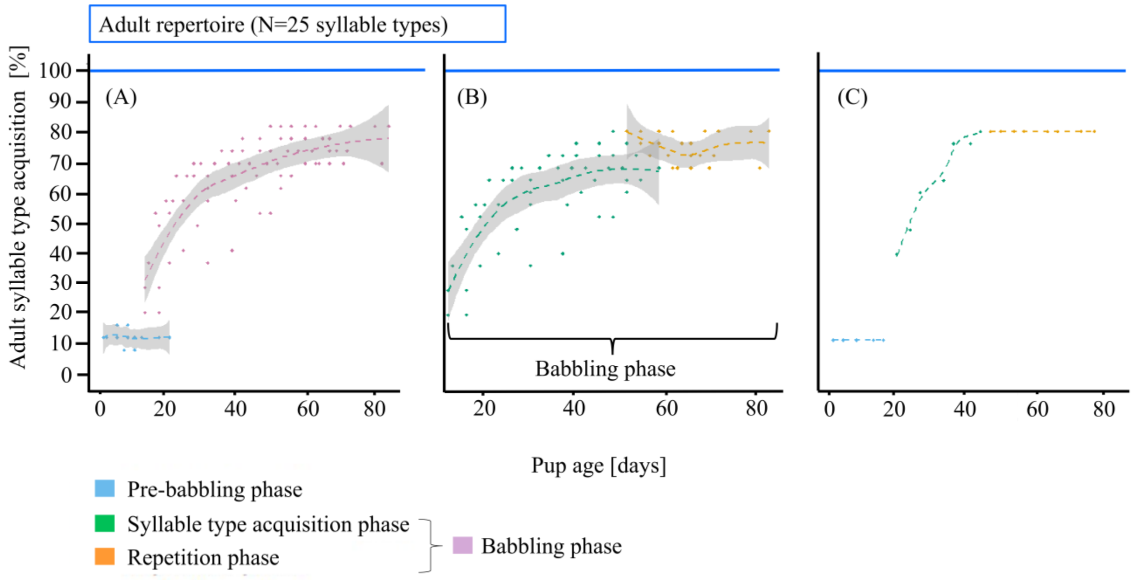


Fig. S5. Vocal ontogeny phases.

The figure depicts the acquisition of syllable types of the adult repertoire as a function of pup age. The sudden increase in syllable types right at the start of the syllable acquisition phase is due to the fact that the pups started to produce 10 new syllable types already during the first few days of the babbling phase (i.e. syllable spurt). (***A***) Pups' vocal ontogeny is composed of two consecutive phases. The pre-babbling phase lasts from birth until babbling onset (light blue color). It is followed by the babbling phase (light purple color) which lasts from babbling onset until weaning. For illustrative purposes, data for only five pups is pooled and depicted here. Please note that the x-axis starts at the pups' respective day of birth (0). (***B***) The babbling phase can be further separated into two phases, the syllable type acquisition phase (green color) and the repetition phase (orange color). Again, data for only five pups is pooled and depicted here. Please note that the x-axis starts at 18.2 days of age, corresponding to the mean age at babbling onset. (***C***) The syllable spurt is very well recognizable when data for only one pup is depicted. The blue line at 100% shows the whole syllable type repertoire of adults (25 different types), demonstrating that none of the pups acquired the entire repertoire. (*A*, *B*: N=216 babbling bouts from 20 pups; *C*: N=20 babbling bouts from one pup).

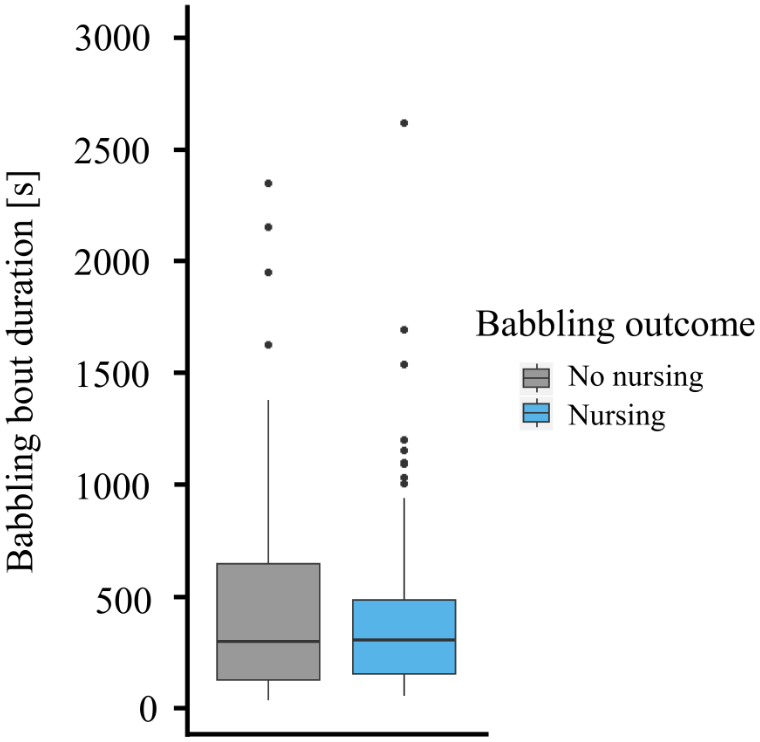


Fig. S6. Babbling bout duration was not influenced by a conflict over nursing.

Whether a pup is allowed to nurse after babbling does not depend on the duration of the babbling bout (N=221 babbling bouts from 20 pups). Long babbling bouts are not the result of a conflict over nursing between mother and pup because mothers terminate solicitation attempts of pups aggressively if they do not want to nurse the pup.

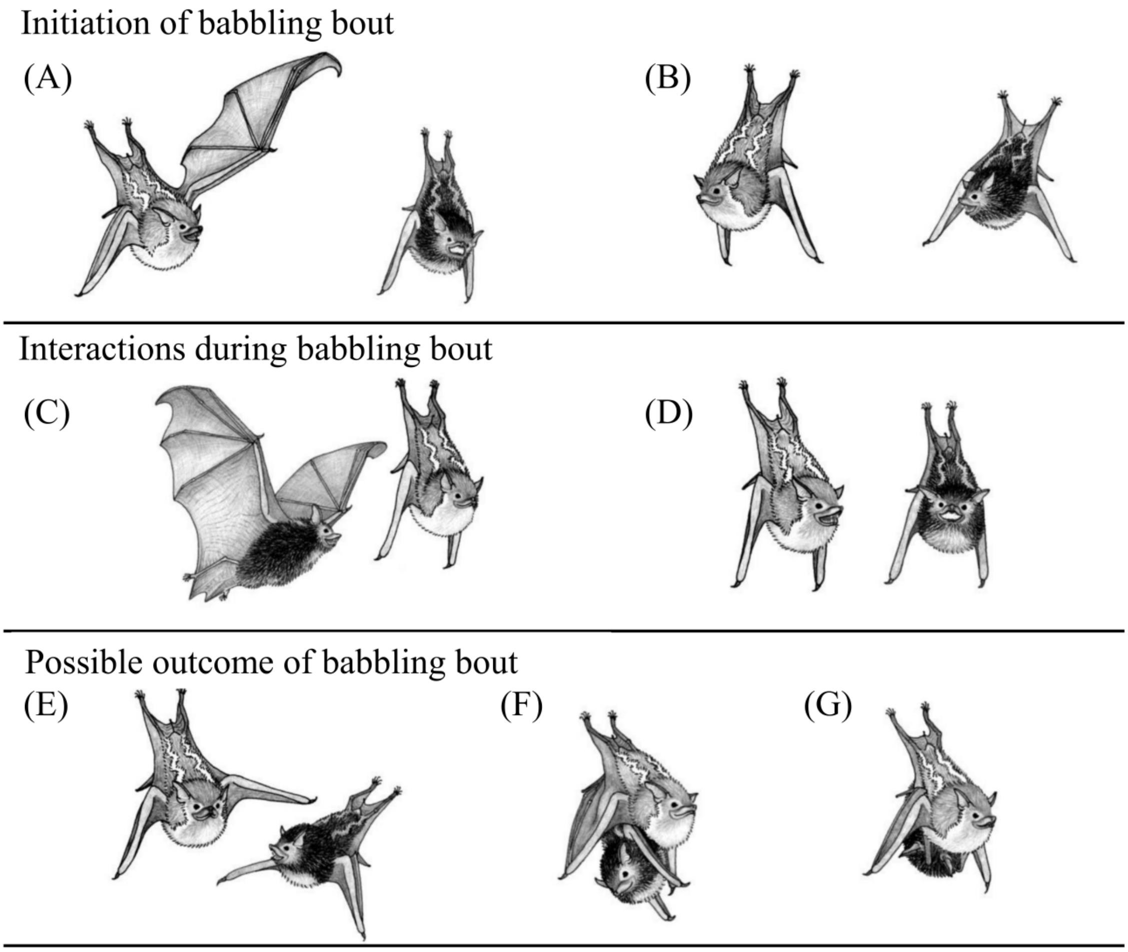


Fig. S7. Mother-pup interactions during babbling.

During babbling, mothers and pups frequently interacted with each other. The interactions comprised different behaviors such as crawling towards each other, hovering in front of each other or making short flights in the day-roost. Those behaviors were repeatedly produced by both, mothers and pups while the pup was constantly babbling. A mother only interacted with its own pup, never with alien pups. (***A***) A babbling bout can either be initiated by the mother or (***B***) by the pup. If the pup initiates it usually starts vocalizing while the mother is perching and not paying attention to the pup (illustrated here by the mothers' head directed away from the vocalizing pup). After a while, the pup crawls towards the mother while vocalizing to start interacting with her. Typical initiating behaviors are wing strokes or crawling towards the mother (***B***). The mother initiates babbling bouts by either hovering in front of, flying next to or crawling towards the pup (**A**). During the babbling bout, a pup frequently hovers in front of its mother (***C***). The mother sometimes produces maternal directive calls (***D***) while her pup is vocalizing. A possible outcome of a babbling bout is nursing. This is initiated when the mother lifts one wing (***E***) and allows the pup to attach to her belly (***F***). After attachment, the pup nurses and rests (***G***). Drawings by C.A.S. Mumm.

General babbling bout characteristics

On the babbling bout level, we demonstrated that the babbling bout duration increased during ontogeny. Neither pup sex nor the daytime of recording had any influence on the babbling bout duration (Table S6).

Table S6. Babbling bout duration during ontogeny.

Dependent variable	Covariates Fixed factors	Estimate	t-value	p-value	Trajectory
Babbling bout duration [s]	Age	0.0151352	4.570	<0.001	Increase***
	Sex: Female	0.0009278	0.005	0.996	NS
	Sex: Male	0.0096686	0.056	0.955	NS
	Daytime: Morning	0.0345356	0.306	0.760	NS
	Daytime: Afternoon	-0.0345369	-0.306	0.760	NS

With increasing age, pups produced longer babbling bouts (daytime and sex did not have a significant effect on bout duration). We calculated a GLMM (Gamma family with log link) with pup age as covariate, pup sex and daytime of babbling (e.g. morning respectively afternoon) as fixed factors and pup ID as random factor. Data was collected from N=221 babbling bouts of 20 pups.

Methods

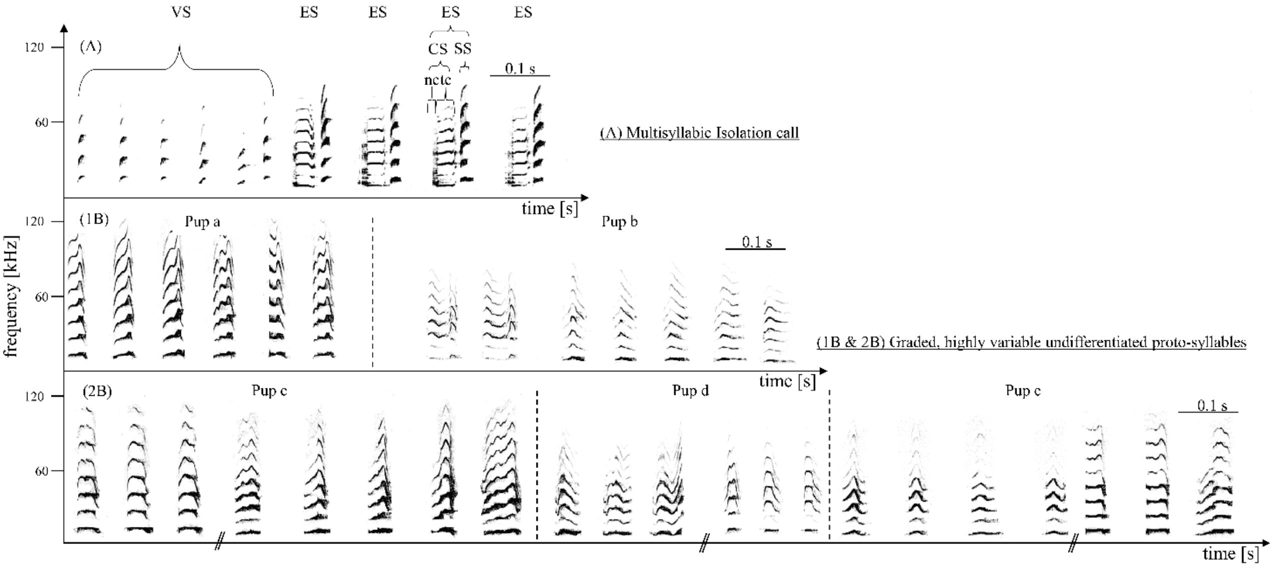


Fig. S8. Characteristic pup vocalizations.

Two characteristic pup vocalizations which were both produced during the pup ontogeny are the isolation call and the undifferentiated proto-syllable. (***A***): Multisyllabic isolation call. This vocalization type is used by pups to solicit maternal care and is produced directly after birth. The most prominent component of an isolation call is the so-called end syllable (ES) which is composed of two different syllables, the composite (CS) and the stereotyped (SS) syllable. CS comprises a noisy part (nc) connected to a tonal part (tc). (***1B***) & (***2B***): Excerpts of undifferetiated proto-syllable sequences. The undifferentiated proto-syllable type is only produced by babbling pups and characterized by multiharmonic and highly variable tonal syllables. (***1B***) depicts undifferentiated proto-syllable train excerpts from two different pups. This middle panel illustrates that this syllable type is graded; the pup mainly uses this type to transit from one adult syllable type to another. (***2B***) shows two undifferentiated proto-syllable type train excerpts from three pups each. The syllable train type excerpts of each pup are separated by // on the x-axis. This panel demonstrates the variability of this syllable type not only between but also within pups. The spectrograms depict frequency [kHz] as a function of time [s] and were generated using a 1024 point fast Fourier transform and a Hamming window with 85% overlap.

Fig. S9. Day-roost of *Saccopteryx bilineata.*

Two adult bats (leftmost bats, both banded with colored plastic rings on their forearms, female: light-blue & white plastic ring on the right forearm, male: orange plastic ring on right forearm) and two pups (about eight weeks old, darker fur than the adult bats, one not banded, the other one banded, both plastic rings are visible). Characteristic of the colony structure is the inter-individual distance which is vigorously maintained (note the distance between the two adult individuals). Only pups are allowed to ignore this inter-individual distance (see mother-pup pair in the top of the picture). Mothers only interact with the own pup (note the distance between the alien pup in the right corner to the female). The mothers' discrimination between own and alien pups and the inter-individual distance outside of the mother-pup bond allowed for directed focal acoustic recordings even when pups were not marked.

The method section includes information about sample sizes from the different analyses (Table S7), background about the study species and about the data collection. Furthermore, this paragraph includes information about the acoustic analyses.

General information about performed analyses

Table S7: Sample sizes of analyses levels 1-3.

	Variables	Sample size (N)			
		Pups	Babbling bouts	Syllable trains	Syllables
Level I					
	Babbling onset (I)	20	216	NA	NA
	Composition (II)	20	216	NA	NA
	Overproduction (III)	10	85	NA	56,611
	Subset acquisition (IV)	20	216	NA	NA
	Non-linear acquisition pattern (V)	20	216	NA	NA
	Repetitiveness (VI)	10	85	NA	56,611
	Meaninglessness (VII)	20	216	NA	NA
	Universality (IX)	20/13*	216/140*	NA	NA
Level II					
	Rhythmicity (VII)	14	NA	30	712
	Babbling bout composition	20	180	8,448	NA
Level III	Babbling bout duration	20	221	NA	NA

The Table S7 shows the sample sizes (i.e. number of pups, babbling bouts, syllable trains and syllables) of all analyses performed. For investigating the different babbling features and, the additional babbling characteristics (i.e. bout composition and bout duration) different analyses were necessary and performed independent from each other. Therefore, sample sizes differ and are sometimes not applicable (NA).

Level 1

We visually scanned 216 babbling bouts of 20 pups (10 from each population in Panama and Costa Rica) for the features: (I) babbling onset, (II) the composition of babbling bouts, (IV) the subset acquisition, (V) the non-linear acquisition pattern and (VIII) meaninglessness. The feature (IX) universality is based on two datasets: for the differences between pups of different regional origin, we examined the data from 20 pups (216 babbling bouts): for the differences between pup sex, we used the data from the 13 sexed pups* (140 babbling bouts). For the features (III) overproduction and (VI) repetitiveness we analyzed a data set from 10 pups (regional origin Panama). We randomly selected excerpts of up to 4 minutes duration from 85 babbling bouts (5-11 babbling bouts per pup) from each pup and classified each syllable (N=56,611 syllables). Based on this classification, we calculated a transition probability matrix to investigate the occurrence of syllable types and the repetition of syllable types (see Dataset 2).

Level 2

We selected 30 syllable trains (6 trains from each of the 5 syllable type train categories) from 14 pups (7 from each population) with best signal-to-noise ratio for labelling each syllable onset in the oscillogram (N=712 syllables). To analyze the babbling bout composition change during ontogeny we visually classified each syllable train (N=4,884) from 180 babbling bouts of 20 pups to the five syllable train type categories which we established.

Level 3

We analyzed the bout duration of 221 babbling bouts from 20 pups.

Background information on the study species and data collection

Study sites and study species

One field season (May-August 2015) was conducted on Barro Colorado Island, a field station of the Smithsonian Tropical Research Institute. This island is located in the artificial Gatun lake, Panamá (9°9’17’’ N, 79°51’53’’ W; 25-165 m above sea level). The second field season (May-September 2016) was realized in the natural reserve Curú, Costa Rica (9°47’0’’ N, 84° 56’ 0’’ W). During the first field season, five different colonies were recorded ad libitum. All of them were located on the outside of buildings; recording distance was within 2-4 meters. During the second field season, three different colonies were recorded ad libitum, two being located at walls of buildings, one in a hollow tree in the forest. Similarly, recordings could be performed in close vicinity to the colonies, within 2-4 meters. The adult individuals maintain an inter-individual distance of five up to eight centimeters which facilitates focal recordings (Fig. S9).

Identification of focal pups

In both study populations, adult bats were individually banded with plastic rings on their forearms (A.C. Hughes Ltd. UK, size XCL, one band per forearm). When pups’ forearm length was almost similar to adult body size, we captured them with mist nets (Ultrathin Mist Nets M-14; Ecotone, Gdynia, Poland) in the vicinity of their natal roosts. Pups were sexed, individually banded with plastic rings and released again at the site of capture. This banding method is well established for *S. bilineat*a (Von Helversen, Voigt, Mayer, & Heckel, 1999) and did not cause any noticeable negative effects on the pups' behavior or welfare. Since capturing attempts were kept to a minimum to avoid disturbing the bats, some pups were not banded and remained unsexed.

Additional information on the social structure of *S. bilineata*

S. bilineata is a highly social bat species roosting in perennial stable social groups (i.e. colonies). The smallest social units are harems, composed of a territorial harem male and up to eight females roosting in the territory (Voigt et al., 2008). The mating system is a resource-defense polygyny; successful males defend a roosting territory of 1-2 m^2 within a day-roost. A colony can constitute of a single harem or several harem territories in close vicinity. Harem males are very faithful to their territory, whereas females tend to occasionally switch between harem territories of a colony. Apart from harem switching, females normally stay within the same colony they chose after dispersal as subadults. In contrast to most other mammals, female *S. bilineata* disperse after weaning (Nagy, Heckel, Voigt, & Mayer, 2007). Colonies are thus patrilines of related males (Heckel & von Helversen, 2003). Yearly, females synchronize the birth of a single pup around the beginning of the rainy season (Voigt et al., 2008). Pups remain with their mothers for about 9.5 weeks until weaning. Within a few days after weaning, female juvenile bats start dispersing to new colonies.

Acoustic recording of focal pups

During the *pre-babbling phase*, acoustic recordings were restricted to the main vocal activity period of *S. bilineata*, one hour at dawn and dusk (Behr & von Helversen, 2004; Bradbury & Emmons, 1974; Eckenweber & Knörnschild, 2013). At dusk, adult males produce territorial songs for approximately half an hour before adult bats leave the day-roost for foraging. Similarly, at dawn, adult males produce territorial songs. Furthermore, harem males produce courtship song (containing affiliative syllables) when greeting each female returning to their territory. During the pre-babbling phase, females return with their pups attached to their belly. After mothers reoccupied their perches in the roost, pups usually produced isolation call bouts while remaining attached to their mothers until they started nursing. Pups who were still attached to the mothers' belly often produced several sequential isolation calls when being awake.

While vocalizing, pups remain attached to their mothers, arching their head away from them while vocalizing. Subsequently, whenever pups changed teats during the day, they produced one or more isolation calls. During the pre-babbling phase, at least one vocal bout (i.e. vocalizations uttered during teat change and not interrupted by silence longer than 3 minutes) per week and pup was analyzed. This was done to investigate which syllable types were already present shortly after birth and before pups started to engage in babbling behavior.

During the babbling phase, acoustic recording sessions were extended to entire mornings (between 6am – 1 pm) and afternoons (1pm - 6pm) where focal pups were recorded ad libitum. For each colony, we alternated between morning and afternoon sessions to avoid potential daytime biases in the vocal activity of pups. Morning sessions started when bats returned to the day-roost. During the first three weeks after birth (on average), pups reunited very quickly with their mothers to nurse and only produced several isolation calls. With increasing age, pups produced babbling bouts directly after arrival at the day-roost. For the remaining babbling phase, pups engaged on average 4-5 times in babbling during the day, as shown in Figure 1B. Upon returning to the day-roost, pups produced the first babbling bout of the day and subsequently nursed. After approximately one hour of nursing, pups usually detached from their mothers and perched independently in their vicinity for a few hours. During this time period, pups repeatedly engaged in babbling behavior. They nursed again around mid-day and right before leaving the day-roost at dusk; the intermediate time was again spent with perching independently and babbling (Fig. 1B). With increasing age, pups became more independent, thus increasing the time periods when perching independently and prolonging their vocal practice (Fig. 1B: 12% corresponds to the average babbling bout duration (7 minutes) a pup spent babbling per hour).

It was possible to monitor the complete pup vocal ontogeny from birth until weaning for 15 out of 20 pups. Those five pups babbled for approximately two more weeks after our last recording session but, unfortunately, we could not obtain recordings of sufficient quality for analysis. Vocal activity of pups and the duration of their babbling bouts and babbling phase varied among individuals, resulting in varying numbers of acoustic recordings.

Babbling excludes the production of single or multiple vocalizations not according to the *babbling bout* definition (i.e. excluding echolocation calls, isolation calls, emission of single vocalizations e.g. production of a few simple syllables or a single noise burst). To investigate which other vocalizations pups produced when not engaging in babbling behavior, we randomly choose a 5 min period of a recording session per week for ten pups (Panama population) during the babbling phase. We identified the vocalizations uttered during his period. The analysis revealed that pups uttered isolation calls in 95% of the cases, singly or in repetitive manner. The other 5% of vocalizations were either discomfort sounds such a single noise bursts or a single chatter (repeated C1 calls).

Pup echolocation call behavior

Saccopteryx bilineata pups produce echolocation calls within hours after birth (personal observation A.A.F). The pup echolocation calls already possess an adult-like spectral form, a similar length (adults: 0.008s; pups: 0.01s) and similar peak frequencies (mean peak frequency adults: 46.8 kHz, pups: 45.9 kHz; mean maximal frequency adults: 49.4 kHz; pups: 48.6 kHz). This is in strong contrast to other bat species which do not produce echolocation calls for several days, sometimes weeks after birth and undergo dramatic acoustic parameter changes during ontogeny (Matsumura, 1981; Monroy, Carter, Miller, & Covey, 2011; Sterbing, 2002).

Sound recording equipment

Vocalizations were recorded using a high-quality ultrasonic sound equipment (500kHz sampling rate, 16-bit depth resolution). The set-up consisted of a microphone (Avisoft UltraSoundGate 116Hm, with condenser microphone CM16, frequency range 1-200kHz ± 3dB) connected to a laptop (Lenovo S21e) running the software Avisoft RECORDER (v4.2.05 R. Specht, Avisoft Bioacoustics, Glienicke, Germany). The microphone was mounted on a tripod which was installed in close distance to the day-roosts (2-4 meters) and directed to the respective focal pup.

Detailed information about acoustic analyses

Pre-babbling phase

For 20 pups, we analyzed 50 vocal bouts during the pre-babbling phase to investigate the vocalizations produced during this phase.

Statistical verification of syllable type classification

To verify our visual classification of syllable types in babbling bouts, which was based on their spectral similarity to adult syllable types, we analyzed a subset of our classified syllables statistically. We measured 528 syllables from 10 different pups (Panama population) belonging to 14 different adult-like syllable types found in babbling bouts of pups. Three syllable types (T2, T3, W1) were excluded from our analysis; they were produced rarely and, therefore, not available in adequate numbers. Furthermore, we excluded the composite end syllable (ES), the noise burst (N) and the pure buzz syllable (B1) since these syllable types were unambiguously distinguishable from all other syllable types due to their unique shape. In addition, we measured 80 undifferentiated proto-syllables from the same 10 pup (Panama population) to investigate their distinctiveness from adult-like syllable types.

Prior to acoustic measurements, syllables were bandpass filtered (5-90 kHz) and normalized to 100% (Cool Edit 2000 Inc., Syntrillium Software Corporation P.O. Box 62255, Phoenix, AZ, USA). For our acoustic analysis, we used the software Avisoft SASLab Pro (v.5.2.09; R. Specht, Avisoft Bioacoustics, Glienicke, Germany). Start and end of syllables were determined manually based on the oscillogram. Subsequently, automatic parameter measurements provided by the software were applied. Spectrograms for measuring syllables were created using a Hamming window with 1024-point fast Fourier transform and 87.5 % overlap (frequency resolution: 488 Hz, time resolution: 0.256 ms). Syllables were multiharmonic but we restricted our measurements to the fundamental frequency (first harmonic) because it often contained most of the sound energy.

Acoustic measurements for syllable type classification

We categorized our acoustic syllable data set into subsets according to syllabic spectral similarities potentially bearing the greatest risk for misclassification. This led to three separate syllable sets. Set 1: the simple frequency modulated syllable types (i.e. tonal syllables which included only a single frequency modulation: VS, VSV, Md, C1, and C2), Set 2: the composite syllable types set (i.e. syllables composed of distinct acoustic units: B2, B3, B4, and W2), and Set 3: the multiple frequency modulated syllable types set (i.e. tonal syllables which included multiple frequency modulations: D1, D2, D3, D4, and T1). If possible, we included syllables from all ten pups from different recording sessions into the three data subsets to minimize potential classification bias due to individual differences or temporal dependencies.

We measured the same acoustic parameters for Set 1 and Set 2, and included an additional count parameter for Set 3. For each syllable we measured 1 temporal parameter (duration) and 6 spectrum-based parameters (peak frequency, minimum and maximum frequencies, bandwidth, entropy and the harmonics-to-noise ratio). The spectrum based parameters were measured at the center, the minimum and maximum point of the syllable, averaged over the entire syllable and at 10 points evenly distributed over the syllable (start, end, and 8 intermediate locations of the syllable). The spectrum-based parameters were used to estimate the frequency and entropy curvatures of syllable types. For Set 3, we additionally counted the number of frequency modulations.

The same acoustic parameters were measured for the undifferentiated proto-syllables. To investigate whether they were distinct from the other adult-like syllable types, the undifferentiated proto-syllables were included in each of the aforementioned sets.

Principal component analyses

For each set (including the undifferentiated proto-syllables), we separately obtained call curvature parameters by performing separate principal component analyses (PCA) with varimax rotation on the above mentioned parameters (one PCA on all 40 frequency parameters and another PCA on all 20 entropy and harmonics-to-noise ratio parameters) which reduced multicollinearity between original acoustic parameters. All 6 (2 PCAs per subset) PCA's fulfilled Kaiser-Meyer-Olkin (KMO) and Bartlett's test criteria, ensuring the appropriateness of our data for PCAs.

Discriminant function analyses

Subsequently, we conducted three separate discriminant function analyses (DFAs) for each of the sets. The DFAs were adjusted to the unequal number of analyzed syllables per syllable type by computing group sizes based on prior probabilities. We used a cross-validation procedure to estimate the correct classification success (n-1 cross-validation procedure) which classified each syllable based on discriminant functions established with all syllables except the syllable being classified. The DFAs were performed in a step-wise manner based on the Wilks-Lambda method. For each DFA, we included the same 9 original parameters, namely, duration, four parameters averaged over the entire syllable (mean minimal, maximal and peak frequency, mean entropy), the minimum and maximum frequencies of the minimum values of the entire syllable and the minimum and maximum frequencies of the maximum values of the entire syllable. For each DFA, we further included eight derived acoustic parameters describing call and entropy harmonics-to-noise ratio call curvatures. In total 17 variables were included in all three separate DFAs (n=18 was the lowest number of cases per group in Set 2).

Results of syllable type classification

Set 1: Simple frequency modulated syllable types

The data set comprised 342 syllables which belonged to 5 different adult-like syllable types (C1, C2, MD, VS, VSV) and undifferentiated proto-syllables (UP). For the frequency curvatures, we extracted 5 principal components (with eigenvalues > 1) which explained 93.58 % of the total variance. For the entropy and harmonics-to-noise ratio curvatures, we extracted 4 principal components (with eigenvalues > 1), which explained 63.95 % of the total variance. We obtained 9 derived acoustic parameters for subsequent statistical analysis.

We analyzed at least 32 syllables per type (32-70 syllables per type, 262 syllables in total). We included 17 acoustic parameters, the aforementioned 9 original parameters and 8 derived acoustic parameters (we included the PC1-4 for the frequency curvature, not including PC5, and PC1-4 for the entropy curvature). Variables remaining in the analysis after the Wilks-Lambda step-wise selection process were: duration, mean maximum frequency, mean peak frequency, the maximum frequency of the minimum values of the entire syllable, frequency call curvature components 1-4, entropy and harmonic-noise ratio call curvature components 1-3. Most syllables were correctly classified to the respective syllable type. A DFA with 342 syllables of 6 syllable types classified 83.6 % of all syllables to the correct syllable type (cross-validation), which was significantly higher than expected by chance (16.66 %). Classification success is shown in Table S8.

Set 2: Multiple frequency modulated tonal syllables

The data set comprised 211 syllables which belonged to five different adult-like syllable types (D1, D2, D3, D4, T1) and undifferentiated proto-syllables (UP). For the frequency curvatures, we extracted 3 principal components (with eigenvalues > 1) which explained 88.80 % of the total variance. For the entropy curvatures, we extracted 6 principal components (with eigenvalues > 1) which explained 59.68 % of the total variance. We obtained 9 derived acoustic parameters for subsequent statistical analysis. We analysed at least 18 syllables per type (18-80 syllable per type, 211 syllables in total). We included 17 acoustic parameters, the aforementioned 9 original parameters and 7 derived acoustic parameters. For this data set, we additionally included the number of frequency modulations since this was an important parameter for our visual classification. Variables remaining in the analysis after the Wilks-lambda step-wise elimination process were: duration, number of frequency modulations, the frequency call curvature components 1-2 and entropy, the harmonics-to-noise-ratio call curvature 4 and the mean maximal frequency. Most syllables were correctly classified to the respective syllable type. A DFA with 211 syllables of 6 syllable types classified 81.5 % of all syllables to the correct syllable type (cross-validation), which was significantly higher than expected by chance (16.66 %). Classification success is shown in Table S8.

Set 3: Composite syllables

The data set comprised 215 syllables which belonged to four different adult-like syllable types (B2, B3, B4, W2) and undifferentiated proto-syllables (UP). For the frequency curvatures, we extracted 6 principal components (with eigenvalues > 1) which explained 91.06 % of the total variance. For the entropy and harmonics-to-noise ratio curvatures, we extracted 5 principal components (with eigenvalues > 1) which explained 60.93 % of the total variance. We obtained 11 derived acoustic parameters for subsequent statistical analysis.

We analyzed at least 21 syllables per type (21-80, 215 syllables in total). We included 17 acoustic parameters, the aforementioned nine original parameters and 8 derived acoustic parameters (we selected PC1-4 for the frequency curvature not including PC5-7, and PC1-4 for the entropy curvature not including PC5-8). Variables remaining in the analysis after the Wilks-Lambda step-wise elimination process were: the minimum frequency of the maximum values of the entire syllable, frequency call curvature components 1-2 and 4, and entropy call curvature component 2. Most syllables were correctly classified to the respective syllable type. A DFA with 215 syllables of 5 syllable types classified 81.5 % of all syllables to the correct syllable type, which was significantly higher than expected by chance (20%). Classification success is shown in Table S8.

Table S8. **Classification success of DFAs (Set 1-3).**

Syllable types

Syllable types

Simple frequency modulated tonal syllables (Set 1)

	C1	C2	MD	VS	VSV	UP
C1	**95.7**	0.0	0.0	0.0	4.3	0.0
C2	0.0	**68.8**	3.1	0.0	28.1	0.0
MD	2.3	0.0	**83.7**	0.0	7.0	7.0
VS	3.4	1.7	1.7	**66.1**	21.1	0.0
VSV	1.7	5.2	0.0	3.4	**89.7**	0.0
UP	0.0	5.0	5.0	0.0	2.5	**87.5**

Multiple frequency modulated tonal syllables (Set 2)

	D1	D2	D3	D4	T1	UP
D1	**75.0**	2.3	0.0	4.5	0.0	18.2
D2	5.0	**80.0**	5.0	0.0	5.0	5.0
D3	0.0	0.0	**52.6**	5.3	31.6	10.5
D4	27.8	5.6	11.1	**38.9**	0.0	16.7
T1	0.0	3.3	3.3	0.0	**93.3**	0.0
UP	0.0	0.0	0.0	2.5	0.0	**97.5**

Composite syllables (Set 3)

	B2	B3	B4	W2	UP
B2	**83.3**	8.3	6.3	2.1	0.0
B3	7.7	**82.1**	2.6	0.0	7.7
B4	11.1	18.5	**33.3**	11.1	25.9
W2	0.0	0.0	23.8	**61.9**	14.3
UP	0.0	0.0	1.3	2.5	**96.3**

DFA classification success for Set 1 (simple frequency modulated tonal syllables and UP syllables; N=342 syllables), Set 2 (multiple frequency modulated tonal syllables and UP syllables; N=211 syllables) and Set 3 (composite syllables and UP syllables; N=215 syllables). Numbers in bold indicate the classification success for the respective syllable type. UP=abbreviation for undifferentiated proto-syllables.

Statistics level 2

To assess whether the production rate of syllable trains per babbling bout changed with increasing babbling phase duration, a GLMM was performed (babbling phase duration as a covariate, sex as fixed factor, pup ID and colony membership as random factors, Gamma family with log link function). A separate GLMM was conducted to assess how the number of trains changed during ontogeny (babbling phase duration as covariate, sex as fixed effect, pup ID and colony membership as random factors, Gamma family with log link function).

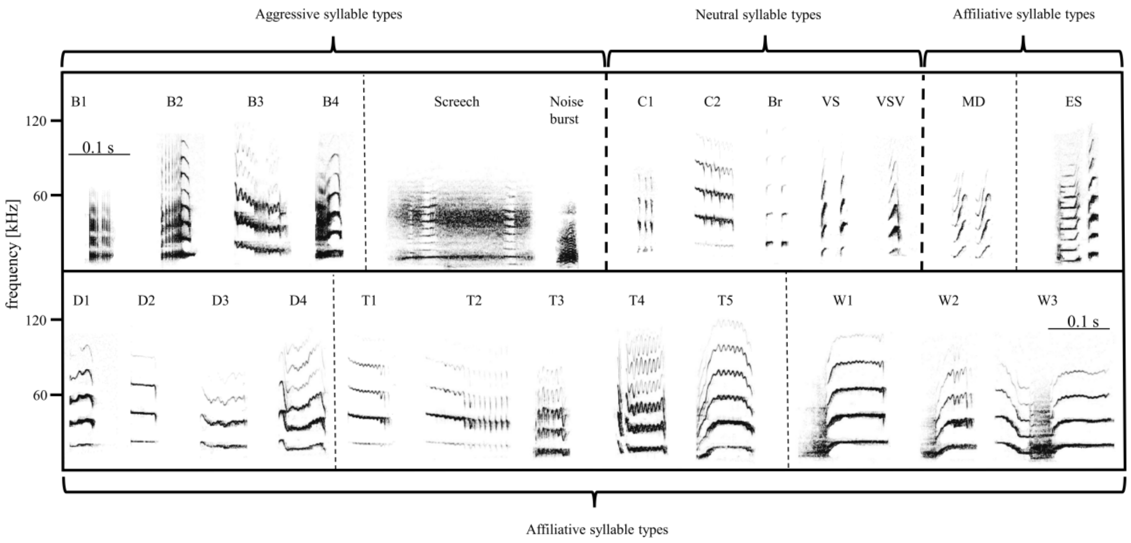


Fig. S10. Adult vocal repertoire.

The adult vocal repertoire is composed of 25 distinct syllable types which can be combined into 10 different and often multisyllabic vocalization types that are produced in distinct social contexts (Behr & von Helversen, 2004; Davidson & Wilkinson, 2002; Knörnschild & von Helversen, 2008). Syllables can be concatenated to vocalizations, either by repeating the same syllable type or by combining multiple syllable types of the adult repertoire (e.g. territorial or courtship song). The adult syllable type repertoire was the basis for our classification of syllable types present in pups' babbling bouts. Since pup versions of adult syllable types (termed 'adult-like syllable types') were so similar to adult syllable types they were clearly recognizable. Thus, we visually classified syllables in babbling bouts based on their spectro-temporal appearance in spectrograms (Fig. 3). While adult males produce the entire set of 25 adult syllable types, adult females only produce seven syllable types (Screech, Noise burst, C1, C2, MD, VSV, VS) after weaning, even though they produce 15-17 syllable types during babbling. Both panels depict exemplary syllable types of the adult vocal repertoire. The syllable types are categorized according to the context in which they are emitted. The upper panel contains the aggressive, neutral and affiliative syllable types (separated by bold dashed lines). Aggressive syllable types are further categorized into the syllable types belonging to the territorial song (B1-B4) and syllable types produced in various agonistic contexts (Screech, Noise burst). The affiliative syllable types are further categorized into maternal directive calls (MD) which are produced exclusively by females and the ES syllables (part of the pup isolation call), which are produced by adult males to signal submissiveness. The lower panel contains the affiliative syllable types belonging to the adult courtship song (categorized according to their spectral appearance: tonals: D1-D4; trills: T1-T5; whistles: W1-W3). A detailed description of the 25 syllable types can be found on the subsequent pages. The spectrograms depict frequency [kHz] as a function of time [s] and were generated using a 1024 point fast Fourier transform and a Hamming window with 85% overlap.

Description of the entire adult syllable repertoire

Even though the complete adult vocal repertoire, consisting of 10 often multi-syllabic vocalization types, had already been described (Behr & von Helversen, 2004; Behr, Knörnschild & von Helversen, 2006; Davidson & Wilkinson, 2002; Knörnschild & Nagy, 2016; Knörnschild & von Helversen, 2008), the different syllable types occurring in adult vocalization types have received less attention. Therefore, it was necessary to describe the entire adult syllable repertoire for this study. We complemented already described syllable types from former studies (Behr & von Helversen, 2004; Eckenweber & Knörnschild, 2013; Knörnschild & von Helversen, 2008) with further syllable type descriptions. A complete list and definition of adult syllable types can be found below.

B1 (pure buzz syllable): aggressive syllable type. The syllable type belongs to the territorial song. It consists entirely of very short pulses. The energy is distributed over the entire syllable. The duration is variable (from 0.02 s to 0.09 s). Produced by adult males. Present in babbling bouts of male and female pups.

B2 (buzz-tonal syllable): aggressive syllable type. The syllable type belongs to the territorial song. It is composed of a buzzed part connected to a subsequent tonal part. The tonal part can have several frequency modulations. The energy is distributed over the entire syllable. The duration is variable, ranging from short (~ 0.6 s) to long (~ 1.6 s). Produced by adult males. Present in babbling bouts of male and female pups.

B3 (mixed buzz-tonal syllable): aggressive syllable type. The syllable type belongs to the territorial song. It is composed of at least two tonal parts with a buzzed part connecting them. The tonal part can include several frequency modulations. The energy is distributed over the entire syllable. This is the buzz syllable type with the most apparent harmonics. The duration is variable, ranging from short (~ 0.8 s) to long (~ 1.7 s). The bandwidth of the second harmonic is located on average between 30 and 47 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

B4 (blurred buzz syllable): aggressive syllable type. The syllable type belongs to the territorial song. It is composed of a blurred buzz part (detectable only in the oscillogramm) followed by a tonal part. The energy is distributed over the entire syllable. The duration is variable (on average 0.03 s). Produced by adult males. Present in babbling bouts of male and female pups.

Screech syllable: aggressive syllable type. This syllable type is produced during agonistic encounters. It is often emitted by adult females as a reaction to a courting male. The syllable type is composed of noisy and tonal parts with predominant noisy parts including nonlinear phenomena. Screeches have a long duration of > 0.1 s. Produced by adult females and males.

Noise burst syllable: aggressive syllable type. This syllable type is produced in agonistic contexts. It is noisy with tonal parts including nonlinear phenomena such as subharmonics. The duration is variable (from 0.03 s to 0.05 s). Compared to the screech syllable, it is much shorter and the tonal parts are not clearly separated from the noisy parts. Produced by adult females and males. Present in babbling bouts of male and female pups.

C1 (chatter): neutral syllable type. This syllable type is produced in the day-roost, mainly in repetitive series. It is tonal and is characterized by a steep frequency down-sweep. A chatter syllable is very short: > 0.004 s. The main energy is located in the second harmonic. The bandwidth of the second harmonic is located between 36 and 46 kHz. Produced by adult females and males. Present in babbling bouts of male and female pups.

C2 (connected chatter): neutral syllable type. This syllable type is produced in the day-roost. It is very similar to C1 but the single chatter syllables are connected (duration is thus variable, on average 0.01 s). The main energy is located in the second and third harmonic. The bandwidth of the second harmonic is located between 40 and 44 kHz. Produced by adult females and males. Present in babbling bouts of male and female pups.

Bark: neutral syllable type: This syllable type was described to mainly be produced by adult males that were not evidently interacting with other colony members, thus we categorized it as neutral type. Nevertheless, in some cases, barks were also emitted in agonistic interactions(see Behr & von Helversen, 2004). The syllable is tonal and quasi constant in frequency. The duration is around 0.01 s. The main spectral energy is located in the fundamental frequency around 22 kHz. Produced by adult males.

VS (very short syllable): neutral syllable type. This syllable type is present in the territorial song, the pup isolation call (in both cases as introductory syllables) but also appears in series on its own. It is tonal and characterized by a steep frequency up- or downsweep (see Fig. 2). The duration is on average 0.009 s. The main energy is located in the second and third harmonic. The bandwidth of the second harmonic is located between 29 and 37 kHz. Produced by adult males and females. Present in babbling bouts of male and female pups.

VSV (inverted V-shaped syllable): neutral syllable type. This syllable type is present in the territorial song and the pup isolation call (in both cases as introductory syllables). It is tonal and contains always one single frequency modulation; the frequency up- and down sweep are of equal length. The duration is around 0.01 s. The main energy is located in the second and third harmonic. The bandwidth of the second harmonic is located between 27 and 37 kHz. Produced by adult females and males. Present in babbling bouts of male and female pups.

Directive maternal syllable: affiliative syllable type. This syllable type is exclusively produced by adult females and directed at their respective pups. It is tonal and characterized by a steep and long frequency up-sweep followed by a very short downwards modulated part, leading to the characteristic "hook" appearance. The duration is around 0.02 s. The energy is equally distributed over the harmonics. The bandwidth of the fundamental frequency is located between 10 and 20 kHz. Produced by adult females. Present in babbling bouts of male and female pups.

ES (end syllable of the pup isolation call): affiliative syllable type. Pups produce this syllable type as part of the isolation call. Furthermore, it is produced by adult males as submissive signal during agonistic encounters (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012). The end syllable is always composed of two separate syllables, the composite and stereotyped syllables (Fig. S8*A*). The composite syllable is characterized by a noisy part, which is preceding a tonal part. The fundamental frequency is located at 8 kHz. The composite syllable is followed by the stereotyped syllable, which is very short (0.01 s). The stereotyped syllable is characterized by a short frequency up-sweep. The fundamental frequency of the stereotyped syllable is at 14 kHz. Produced by adult males and pups of both sexes. Present in babbling bouts of male and female pups.

D1 (short tonal syllable): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and contains at least one and up to several frequency modulations. The duration is short (on average less than 0.03 s). The main energy is located in the second and third harmonic. The second harmonic is on average located around 35 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

D2 (quasi-constant frequency tonal syllable): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and characterized by a quasi-constant frequency with a very short frequency down-sweep at the end of the call. The syllable duration is on average longer than 0.04 s. The main energy is mainly located in the second and third harmonic. The second harmonic contains most energy and is located around 45 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

D3 (irregular frequency modulated tonal syllable): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and characterized by small and irregular frequency modulations. The average duration is longer than 0.3 s. The main energy is located in the second harmonic, which lies around 37 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

D4 (long tonal syllable): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal with very small frequency modulations. Furthermore, it is characterized by a very steep frequency down-sweep at the start of the syllable followed by a shallow and long frequency up-sweep. The duration is longer than 0.4 s. Most energy is located in second harmonic (ranging between 32 – 47 kHz). Produced by adult males. Present in babbling bouts of male and female pups.

T1 (regular trill): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is a tonal syllable, which is entirely defined by frequency modulations. The frequency modulations transit from very small at the start of the vocalization to large modulations at the end. This gives the syllable a trill-like appearance. The duration lies around 0.05 s. The main energy is located in the second harmonic, which is located around 41 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

T2 (chatter-trill): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and characterized by the same trill-like appearance as in T1. But in contrast to T1, T2 is always either preceded or followed by at least two chatter syllables (C1). The duration is variable (includes also the chatter, on average 0.14 s). The main energy is located in the second harmonic, which is located around 41 kHz. Produced by adult males. Present in babbling bouts of male and female pups.

T3 (interrupted trill): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is a tonal syllable characterized by large frequency modulations, which are often not connected to each other. T3 has a certain resemblance to the syllable type C2 but the difference is observable in the size of the frequency modulations (T2 has much larger frequency modulations). The duration is variable (on average 0.06 s). The main energy is distributed evenly over the three first harmonics. Produced by adult males. Present in babbling bouts of male and female pups.

T4 (large trill): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and characterized by large mostly regularly patterned frequency modulations. The trill is preceded by a short tonal frequency down-sweep syllable. The duration is variable (on average 0.08 s). The main energy is evenly distributed over the three first harmonics. Produced by adult males.

T5 (M-shaped trill): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is tonal and characterized by a very steep frequency up-sweep followed by a long part, which is characterized by frequency modulations of variable size. The duration is on average around 0.09 s. The main energy is located in the second harmonic. Produced by adult males.

W1 (regular whistle): affiliative syllable type. The syllable type is present in the adult courtship song. The syllable is composed of two acoustic units, the first entirely noisy part is connected to a subsequent tonal part. The tonal part is very long and is characterized by very small regular frequency modulations (trill-like appearance, see T1). The duration is variable (on average 0.05 s). Most energy is evenly distributed over the first three harmonics. Produced by adult males. Present in babbling bouts of males and female pups.

W2 (tonal modulated whistle): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is composed of two distinct acoustic units. The noisy part is connected to a tonal part, which is much shorter than the tonal part of W1. The tonal part is characterized by large and irregularly patterned frequency modulations. The duration is variable (on average 0.06 s). The main energy is located in the second harmonic. Produced by adult males. Present in babbling bouts of males and female pups.

W3 (composed modified whistle): affiliative syllable type. This syllable type is present in the courtship song of adult males. It is composed of several acoustic units, mostly of two tonal parts connected to a noisy part. The tonal parts are characterized by very small frequency modulations (trill-like appearance) and can also include large up-and down sweeps. The duration is variable, but of rather long duration (on average 0.19 s). The main energy is located in the first harmonic. Produced by adult males.

Glossary

Level 1: Syllable level.

Pre-babbling phase: the duration in days from the day the pup is born until the day it perches temporarily without body contact to its mother and starts with *babbling* (Fig. *1A*). During the *pre-babbling phase* pups produce isolation calls and echolocation calls (both are innate since they are produced within the first hour after birth) and remain constantly attached to the mothers' belly.

Babbling phase: starts when pups detach from their mothers' belly and perch independently for a certain time in the day-roost. As soon as they perch independently, pups start to babble. The babbling phase started when a pup emitted the first babbling bout (= day 1 of babbling phase). The end of the babbling phase was the day pups were weaned and stopped babbling (Fig. *1A*).

Babbling phase duration: time span in days during which the pups were observed to engage in babbling. Babbling onset corresponds to the day during which pups were babbling for the first time (day 1 of babbling phase duration). Babbling phase duration is not to be confounded with pup age, which is also counted in days.

Babbling behavior: reminiscent of the so-called canonical babbling in human infants and characterized by concatenation of different syllable types into repetitive sequences (Oller & Eilers, 1988). Contains adult-like syllable types and the transition syllable type which is exclusively produced by pups. Also called vocal practice or vocal play.

Syllable(s): the smallest acoustic unit i.e. a continuous sound emission with no silent gaps longer than 0.001 ms (Behr & von Helversen, 2004; Catchpole, 1980).

Syllable type: the nomination of each syllable type of the adult vocal repertoire (which is composed of 25 different syllable types). The syllable types have been named after their spectral appearance and / or behavioral context: Buzzes (B1-B4), Chatter (C1-C2), simple single frequency modulated syllables (V1-V2, Bark), pure tonal syllables (D1-D4), Trills (T1-T5), Whistles (W1-W4), noisy syllables (NB=noise burst, Screech), and maternal directive call (MD). See supporting methods "Spectro-temporal description of the adult vocal repertoire" for further description. Further, syllable types are categorized into aggressive, neutral and affiliative syllable types corresponding to the context in which they are produced by adults. For the description of the syllable types and their categorization into aggressive, affiliative, neutral, isolation call and transitional syllable types see supporting results (pages 24-26).

Undifferentiated proto-syllable: a syllable type produced exclusively by pups and characteristic of babbling bouts (Fig. *S8B*). Multi-harmonic, frequency modulated tonal syllables, characterized by high variability, both between and within pups. This syllable type was often graded within a syllable train, changing in frequency modulations as well as duration when pups moved from the production of one adult-like syllable type to the next. They were present during the entire babbling phase. The term undifferentiated proto-syllable is related to the term “protophone” known from the literature about infant speech acquisition (Oller, 1980; Oller, Griebel, & Warlaumont, 2016). Protophones are non-canonical speech sounds which do not have the adult-like speech properties, i.e. they are not fully resonant yet and the temporal transition between consonant and vowel are much slower. While adult-like syllables in pup babbling bouts can be best compared to infant canonical syllables, the undifferentiated proto-syllables are best compared to infant speech protophones.

Pups’ final syllable type repertoire size: the number of syllable types which were present in babbling bouts, from the first to the last babbling bout analyzed. We excluded undifferentiated proto-syllables from our count of the final syllable type repertoire of pups, since this syllable type is exclusively produced by pups and our aim was to determine how many of the adult syllable types the pups could already produce at the time of weaning.

Adult vocal repertoire: 25 different syllable types which are used in social interactions (Fig. S10). Echolocation calls, which are predominantly used for orientation, navigation and foraging, are excluded.

Acquisition: The usage of the term “acquisition” is not implying that pups obtain each adult syllable type by vocal production learning (VPL). A former study showed that the syllable types belonging to the later male territorial song (B1-B4) are obtained through vocal imitation of adult tutors (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). However, other syllable types (VS, VSV, ES) are innate since they are part of the multisyllabic isolation call which is produced directly after birth. For the remaining adult syllable types, it is not yet investigated whether they are learned or not.

Repetition phase: was defined as the time period during which pups did not acquire new adult-like syllable types while babbling, but only rehearsed the ones already present in their pup syllable repertoire.

Level 2: Syllable train level

Syllable train: a sequence of at least five syllables in direct succession; syllable trains were separated by silent gaps not longer than 300ms.

Syllable type train category: a sequence of syllables usually nominated after the most abundant syllable type in the *train*. We established five different train categories, namely the *neutral -, aggressive -, affiliative -, undifferentiated proto-syllable* - and *isolation call* – syllable type train category. The name-giving functions of train categories were inferred from the use of syllable types in the adult vocal repertoire. There was one exception, since affiliative syllable types were produced much more seldom compared to aggressive or transition syllables, Syllable trains containing at least two affiliative syllable types were classified as *affiliative trains*, even though affiliative syllables were not the most abundant syllables in them (Fig. S3*B*). This was done because affiliative syllable types were produced much more rarely than other syllable types and we wanted to give their occurrence its due credit.

Level 3: Babbling bout level

Babbling bout: composed of at least three syllable trains which belong to at least two different syllable train categories. Babbling bouts had a minimum duration of 30 seconds. Silent gaps between syllable trains had a maximum duration of one minute. The end of a babbling bout was defined by the pups' behavior. We considered a babbling bout to be terminated when the pup remained silent for longer than 3 minutes or when the pup was allowed to nurse.

References

Behr, O., & von Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology, 56*(2), 106-115. doi:10.1007/s00265-004-0768-7.

Behr, O., Von 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). *Behavioral Ecology, 17*(5), 810-817.

Bradbury, J. W., & Emmons, L. H. (1974). Social organization of some Trinidad bats: I. Emballonuridae. *Zeitschrift für Tierpsychologie, 36*(1-5), 137-183.

Catchpole, C. K. (1980). Sexual Selection and the Evolution of Complex Songs Among European Warblers of the Genus Acr Ocephal Us. *Behaviour, 74*(1), 149-165.

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. doi: 10.1644/1545-1542(2002)083<0526:Gaiviv>2.0.Co;2.

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

Heckel, G., & Von Helversen, O. (2003). Genetic mating system and the significance of harem associations in the bat *Saccopteryx bilineata*. *Molecular Ecology, 12*(1), 219-227.

Knörnschild, M., & Nagy, M. (2016). Sex-biased dispersal and social systems of Neotropical Emballonurids. In J. Ortega (Ed.), *Sociality in bats* (pp. 47-63): Springer International Publishing Switzerland.

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. doi:10.1098/rsbl.2009.0685.

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*(4), 761-769. doi:10.1016/j.anbehav.2012.06.029.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Matsumura, S. (1981). Mother-infant communication in a horseshoe bat (*Rhinolophus ferrumequinum nippon*): vocal communication in three-week-old infants. *Journal of Mammalogy, 62*(1), 20-28.

Monroy, J. A., Carter, M. E., Miller, K. E., & Covey, E. (2011). Development of echolocation and communication vocalizations in the big brown bat, *Eptesicus fuscus*. *Journal of Comparative Physiology A, 197*(5), 459-467.

Nagy, M., Heckel, G., Voigt, C. C., & Mayer, F. (2007). Female-biased dispersal and patrilocal kin groups in a mammal with resource-defence polygyny. *Proceedings of the Royal Society B: Biological Sciences, 274*(1628), 3019-3025. doi:10.1098/rspb.2007.1008.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In J. F. K. a. C. A. F. G. H. Yeni-Komshian (Ed.), *Child phonology* (Vol. 1, pp. 93-112). New York: Academic Press, Inc.

Oller, D. K., & Eilers, R. E. (1988). The Role of Audition in Infant Babbling. *Child Development, 59*(2), 441-449. doi:10.2307/1130323.

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Topics in cognitive science, 8*(2), 382-392.

Sterbing, S. J. (2002). Postnatal development of vocalizations and hearing in the phyllostomid bat, *Carollia perspicillata*. *Journal of Mammalogy, 83*(2), 516-525. doi:10.1644/1545-1542(2002)083<0516:Pdovah>2.0.Co;2.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat,*Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

Von Helversen, O., Voigt, C., Mayer, F., & Heckel, G. (1999). Extra-harem paternity in the white-lined bat *Saccopteryx bilineata* (Emballonuridae). *Behaviour, 136*(9), 1173-1185.

When published, these additional datasets and the movie are going to be available in a data repository connected to the publication. For now, this data is available on request to AAF.

Movie S1. Babbling bat pup

The movie shows a short excerpt of a babbling bout. The individual with the colored plastic rings on both forearms is the mother. The unbanded individual with the dark fur is her pup. At the bottom of the video, the syllables produced by the pup while babbling are depicted in real time. The video shows how the mouth of the pup is clearly open while vocalizing.

Additional dataset S1 (separate excel file)

Syllable repertoire of pups' at weaning

For each pup (N=20, column ID) the syllable types which were acquired (=1) during the babbling phase and syllable types that were not acquired (=0) are given (N=216 babbling bouts). Five syllable types from the adult repertoire are missing in all pups at weaning (B=bark, T4, T5, W3, and Screech) and are therefore not included in the dataset.

Additional dataset S2 (separate excel file)

Repetitiveness in babbling bouts

This table represents the transition probabilities (in %) between syllable types in babbling bouts (N=56,611 syllables, N=10 pups, Panama population). Numbers in bold indicate the highest transition probability for the respective syllable type. Reading the tables in series from left to right gives the transition probability between a syllable type and the subsequent one.

Social influences shape the babbling behavior of greater sac-winged bat pups, *Saccopteryx bilineata*

Authors: Ahana A. Fernandez, Nora Serve, Sarah-Cecil Fabian & Mirjam Knörnschild

ABSTRACT

Infant babbling is a crucial and mandatory steep in human speech acquisition. Social interactions with caregivers significantly influence babbling behavior, and consequently language development. Likewise, research in songbirds demonstrated that vocal learning is a social process, affected by interactions with conspecifics. During ontogeny, the vocal production learning bat *Saccopteryx bilineata* goes through a babbling phase that is characterized by the same features as infant canonical babbling. The pups´ social environment consists of acoustic input from adult males and behavioral interactions with the mother. This study investigates the influences of the social environment on three aspects of pup babbling, namely the amount of vocal practice, the syllabic composition of babbling bouts and the pups' final syllable repertoire size. Our results show that the amount of vocal practice, both on a daily basis and over the entire babbling phase, is influenced by maternal behaviors. While the pups' final syllable repertoire size is not influenced by the social environment, the babbling bout composition is shaped by social influences. We found that the number of syllable sequences containing territorial song syllables increase with the number of adult singing males (i.e. potential tutors) and with the number of maternal behaviors during pup babbling bouts. This is especially noteworthy, since the territorial song syllables are acquired during babbling through vocal production learning. In summary, our study shows that, similar to infant speech acquisition, social feedback influences the vocal ontogeny in a vocal production learner with a pronounced vocal practice phase.

Key words: social environment, vocal ontogeny, vocal practice, maternal behaviors

INTRODUCTION

Infant canonical babbling is a milestone in speech acquisition (Oller, 1980; Vihman, 2014). Canonical babbling is universal, irrespective of sex or cultural background and characterized by several key features such as rhythmicity and reduplication. During canonical babbling, infants imitate speech sound subunits by matching their own oral output to the auditory input provided by the social environment to acquire the basis of any spoken language, namely the phonological repertoire (Oller, 1980). Consequently, the ability of vocal production learning and the vocal practice itself are two mandatory factors for speech acquisition (Doupe & Kuhl, 1999; Vihman, 1996, 2014). The intense learning process during speech acquisition is influenced by social feedback. The child's social environment shapes speech development through directed and contingent feedback on the vocal output of the child (i.e. direct interaction with the infant) and through the indirect linguistic environment (e.g. talking people in the child's environment) (Goldstein & Schwade, 2010; Kuhl, 2003, 2007). Direct interactions involve different modalities, from acoustic (e.g. talking to the child) to behavioral responses (e.g. smiling at the infant, moving closer, touching the infant; Goldstein, King, & West, 2003; Goldstein & Schwade, 2008, 2010). For example, it was demonstrated that babbling rate increases when children were approached and touched while babbling (Goldstein, King, & West, 2003).

So far, songbirds have been the preferred taxon to study parallels in vocal development of humans and animals (Doupe & Kuhl, 1999; Goldstein & Schwade, 2010; Marler, 1970). Comparisons between songbirds and humans are essentially based on the shared ability of vocal production learning and parallels in certain vocal ontogeny stages, especially the plastic song production in songbirds which is reminiscent of the canonical babbling phase (Beecher & Brenowitz, 2005; Doupe & Kuhl, 1999; Konishi, 1985; Marler, 1970). Many studies investigated the role of social interactions on song learning (Beecher, 2017; Beecher & Brenowitz, 2005; Chen, Matheson, & Sakata, 2016; García, 2019). Most of these studies focused on the vocal interaction between tutor and tutee (Beecher & Brenowitz, 2005; García, 2019). While this gained crucial insights on the process of vocal learning per se, song development research was mostly carried out in restricted social environments, i.e. in artificial laboratory set-ups (García, 2019), but see (Mennill et al., 2018). This leads to a somewhat incomplete picture of vocal development, since in the wild the social environment of a juvenile songbird is not restricted to a single interactive partner but includes several tutors, siblings or other adult conspecifics (e.g. females). Furthermore, a growing body of studies shows that social processes influencing vocal development are not confined to vocal interactions but involve other modalities such as visual displays as well (Adret, 2004; Carouso-Peck & Goldstein, 2019; King, Freeberg, & West, 1996; West & King, 1988).

A suitable mammalian model for studying vocal ontogenetic processes are bats. First, the vocal modality is very distinctive in bats. Besides their fine-tuned echolocation used for foraging and navigation bats produce social vocalizations, ranging from simple calls (e.g. Chaverri, Gillam, & Kunz, 2012) to songs (e.g. Jahelková, Horáček, & Bartonička, 2008) and small (e.g. Knörnschild, Kalko, & Feifel, 2014) to large vocal repertoires (e.g. Behr & von Helversen, 2004). Furthermore, bats are social animals, they live in groups which range from small (e.g. up to eight individuals, Chaverri, Gillam, & Kunz, 2012), to medium sized to huge groups (e.g. several thousand individuals, Bohn, Schmidt-French, Schwartz, Smotherman, & Pollak, 2009). Hence, the pups´ social environment is normally constituted of many conspecifics which could potentially influence its vocal ontogeny. Studying how the social environment shapes vocal ontogenetic processes is especially interesting in bats that are vocal production learners, like the greater sac-winged bat *Saccopteryx bilineata*.

During ontogeny, *S. bilineata* pups exhibit a conspicuous phase of intense vocal practice (i.e. babbling), that is highly reminiscent of infant canonical babbling. While babbling, pups acquire a part of the large adult vocal repertoire through vocal production learning (Knörnschild, Behr, & von Helversen, 2006; Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Pups produce babbling bouts (average duration of seven minutes), which are composed of multisyllabic repetitive vocal sequences containing adult-like syllable types and the juvenile undifferentiated proto-syllables (Chapter 1). Adult-like syllable types represent syllables which constitute the later adult vocal repertoire whereas the undifferentiated proto-syllables are exclusively produced by pups and only present during babbling (Chapter 1). The daily production of babbling bouts (up to 5 bouts per day) starts at an early age (on average 2.5 weeks after birth) and lasts on average 7.5 weeks (Chapter 1). Male and female pups babble (Knörnschild, Behr, & von Helversen, 2006), which is a significant difference to the male-biased plastic song production in songbirds (Doupe & Kuhl, 1999; Marler, 1970) but see Odom et al. for female birdsong (Odom, Hall, Riebel, Omland, & Langmore, 2014). To our knowledge, *S. bilineata* is the first mammalian vocal production learner with a vocal practice phase characterized by the same features that define infant canonical babbling (Chapter 1).

During ontogeny, the social environment (i.e. the colony) of pups is constituted of the mother, other females with their respective pups and adult males. Mothers are the only adult colony members that directly interact with their own pups, mainly through non-vocal behaviors. During babbling bouts, both mother and pup produce behaviors while the pup is babbling (Table S1). Behaviors often occur in sequences, i.e. interactions between mother and pup, and can be initiated by both. Behaviors and interactions are restricted to mother-pup dyads and never involve other adult conspecifics or pups. Outside of the babbling context adult females never produce these behaviors (own observations, Bradbury & Emmons, 1974; Tannenbaum, 1975; Voigt et al., 2008). Since anecdotal evidence indicates that mothers differ in their behavior (i.e. number of behaviors and number of interactions with pups) we wanted to investigate if the maternal behavior influences the pups` babbling behavior. We investigated if the maternal behavior influences the following three aspects of pup babbling: the amount of vocal practice, the syllabic train composition of babbling bouts and the pups' final syllable repertoire size. The amount of vocal practice is defined as a) the daily practice (i.e. babbling bout duration) and b) the overall babbling phase duration (i.e. time in days from the first until the last day of babbling). A babbling bout is composed of a varying number of syllable trains (i.e. sequences of at least five syllables in direct succession) belonging to the five different syllable train categories (see material and methods). The pups' final syllable repertoire size is defined as the number of adult syllables which is present at the last day of babbling. We hypothesized that the mothers´ behavior influences the babbling behavior of pups. Specifically, we predicted that an increase of maternal behaviors occurring during babbling would result in longer babbling bouts and mothers that are more active would have pups with a longer babbling phase duration. Furthermore, we predicted that babbling bouts of pups with more active mothers would constitute of more different syllable types and would have a larger final syllable repertoire size.

Adult females do not produce many vocalizations whereas adult males provide ample vocal input by producing two song types, the territorial and courtship song (Behr & von Helversen, 2004). Adult males vocally defend their territories in the day-roost engaging every day at dusk and dawn in territorial song production (Eckenweber & Knörnschild, 2013). Pups acquire this territorial song through vocal production learning (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Furthermore, adult males daily produce courtship song when greeting females that are returning to the day-roost (Behr & von Helversen, 2004). We hypothesized that the number of singing adult males present in a colony positively influences the aforementioned three aspects of pup babbling, namely the amount of vocal practice, both the daily and the overall babbling behavior, the babbling bout composition and pups´ final syllable repertoire size.

METHODS

Study sites and animals

We studied the babbling phase of 29 pups from 11 free-living colonies of *S. bilineata* during three consecutive field seasons (May-September 2015-2017) in Panama and Costa Rica. The vocal and social behaviors of pups were recorded *at libitum* throughout the pups′ ontogeny (i.e. from birth until weaning at 10-12 weeks of age). In 2015, we conducted acoustic recordings of 10 pups belonging to five colonies at Barro Colorado Island, a field station of the Smithsonian Tropical Research Institute located in the Gatún lake of the Panamá Canal. In 2016, we conducted acoustic recordings and additional observations of accompanying behaviors from 10 mother-pup dyads belonging to three colonies in the natural reserve Curú in Costa Rica. In 2017, we likewise conducted acoustic recordings and simultaneous behavioral observations from nine mother-pup dyads belonging to three colonies in Gamboa, a field station of the Smithsonian Tropical Research Institute located at the Panamá Canal. Since only mothers interact with pups, they constitute their direct social environment. During all field seasons, we also counted the number of 1) singing males and 2) other pups per group; they represented the indirect social environment of our focal pups.

Saccopteryx bilineata roosts in perennial stable groups, i.e. colonies (Tannenbaum, 1975). Females synchronize the birth of a single pup around the beginning of the rainy season (i.e. May). Pups remain for 10-12 weeks with their mothers upon weaning (Tannenbaum, 1975). Day roosts of *S. bilineata* are located in tree cavities or outer walls of man-made structures (Yancey, Goetze, & Jones, 1998). *Saccopteryx bilineata* is very light-tolerant and most social interactions (e.g. courtship, territorial defense, babbling behavior) occur during the day within the roost (Behr & von Helversen, 2004; Knörnschild, Behr, & von Helversen, 2006). Individuals in the day-roost maintain an inter-individual distance of 5 to 8 centimeters (Voigt et al., 2008). Only pups are allowed to ignore this inter-individual distance without eliciting aggression. Since adult bats are individually banded (see supporting information) and mothers only nurse their own pup, individual identity of our focal bats could be determined exactly. After habituation, it was possible to record and observe focal bats within a distance of 2-4 meters without noticeable disturbance.

Acoustic recordings

During the babbling phase, each focal pup was recorded *ad libitum* at least twice per week in alternating morning (6am-12am) and afternoon (12am-6pm) sessions to avoid potential daytime biases in the vocal activity of pups. We used high-quality ultrasonic recording equipment (500 kHz sampling rate, 16-bit depth resolution) to record the vocalizations. The recording set up consisted of a microphone (Avisoft Ultra SoundGate 116Hm, with condenser microphone CM16, frequency range 1-200 kHz ± 3dB, connected to a laptop computer Lenovo S21e) running the software Avisoft-Recorder (v4.2.05 R. Specht, Avisoft Bioacoustics, Glienicke, Germany). The microphone was mounted on a tripod which was positioned in close distance to the day roost and directed at the respective focal pup.

Analyses of pup babbling behavior

We performed babbling analyses on three different levels; the syllable level (i.e. the basic unit), the syllable train level (i.e. syllables concatenated to sequences), and the babbling bout level (i.e. composed of syllable trains and interspersed with silent intervals (see methods Chapter 1)). We scored the size of the pups' final syllable type repertoire at the time of weaning (syllable level) for the pups from field season 2015 & 2016 (Panama 2015: N=10 pups, Costa Rica 2016: N=10 pups, N=216 babbling bouts, see methods Chapter 1). For the same pups we classified the syllable trains within babbling bouts (Panama 2015: N=4358 syllable trains; Costa Rica 2016: N=5011 syllable trains). The five different syllable train categories were established based on the context of the syllable types in the adult vocal repertoire, namely the neutral-, affiliative-, aggressive-, undifferentiated proto-, and isolation call syllable type category. The name-giving functions of train categories were inferred from the use of syllable types in the adult vocal repertoire. There was one exception, since affiliative syllable types were produced much more seldom compared to aggressive or undifferentiated proto-syllables. Syllable trains containing at least two affiliative syllable types were classified as affiliative trains, even though affiliative syllables were not the most abundant syllables in them. This was done because affiliative syllable types were produced much more rarely than other syllable types and we wanted to give their occurrence its due credit. The daily bout duration (in seconds) and babbling phase duration (in days) was obtained during all three field seasons (Panamá 2015: N=52 babbling bouts from 10 pups, Costa Rica 2016: N=120 babbling bouts from 10 pups, Panama 2017: N=63 babbling bouts from nine pups).

Behaviors during babbling bouts

During the field seasons 2016 and 2017 we collected data on behaviors of mother-pup dyads occurring simultaneously with the acoustic recordings. We established an ethogram (Table S1) based on our observations of mother-pup behaviors and took the behavioral observations of a former study into account (Strauss, von Helversen, & Knörnschild, 2010). We analyzed a total of 12´796 behaviors produced by mothers and pups during 183 babbling bouts (Costa Rica 2016: N=120 babbling bouts, N=10 mother-pup dyads, Panama 2017: 63 babbling bouts, N=9 mother-pup dyads).

Maternal activity score

To assess the mothers` behavior not only on a daily basis but over the course of the entire pup babbling phase we calculated a maternal activity score. Since mothers differed in their activity (i.e. number of behaviors during babbling bouts with special focus on the hover behavior) we used a proxy for maternal activity during babbling bouts by calculating the mean occurrence of different behaviors. To account for the different sample sizes (i.e. number of analyzed bouts differed between mothers due to different vocal activity of pups) the mean was taken for: a) the number of behaviors per bout (exclusive hover), b) the number of hovers per bout and c) the behavioral rate (i.e. number of behaviors per bout duration). For each mother we obtained a score (min: 3, max: 15, mean: 8.3), lower numbers account for less active mothers and higher numbers for more active ones. The activity score and the number of analyzed bouts did not correlate significantly (i.e. the activity score of a mother is not biased by the different number of bouts analyzed per mother-pup pair; Spearman-rank correlation test: r_s= -0.06, p=0.64).

Analyses and statistics

We investigated the influence of the pups’ direct social environment (mothers) and indirect social environment (singing males, other pups) on three aspects of pup babbling, namely the amount of vocal practice, the syllabic composition of babbling bouts and the pups’ final pup syllable repertoire size. Data on the influence of mothers and singing males is described here; data on the influence of other pups can be found in the supporting information;

Since we did not obtain enough high-quality recordings and/or behavioral observations for all 19 mother-pup dyads, we had to restrict some analyses to a data subset. Therefore, sample sizes differed for our analyses.

The amount of vocal practice

To investigate whether the number of maternal behaviors during babbling bouts and the maternal activity score influenced the babbling bout duration, we calculated a general linear mixed model (GLMM) based on the dataset of 2016 and 2017 (N=19 mother-pup dyads, N=183 babbling bouts). The first GLMM was conducted to investigate how the number of maternal behaviors influences the babbling bout duration. The model included the number of maternal behaviors during babbling bouts, maternal activity score and centered pup age as fixed effects, pup ID as random factor, Gamma family with log link function. We conducted another GLMM to investigate the maternal influence on the babbling phase duration in days (GLMM: number of maternal behaviors during babbling bouts, maternal activity score as fixed factors, pup ID as random factor, Poisson family with log link function).

To investigate the influence of the number of singing males present in a colony we likewise conducted two GLMMs based on the same dataset as aforementioned (GLMM bout duration: number of singing males and centered pup age as fixed factors, pup ID as random factor, Gamma family with log link function; GLMM babbling phase duration: number of singing males as fixed factor, pup ID as random factor, Poisson family with log link function).

Babbling bout composition

We analyzed the influence of the mothers` behavior on the syllable train production (i.e. composition of babbling bouts) based on the 2016 dataset (N=10 mother-pup dyads, N=77 babbling bouts, N=5011 syllable trains). We calculated five separate GLMMs to investigate the maternal influence on the number of syllable trains produced for the following categories: isolation call -, undifferentiated proto-syllable -, affiliative syllable -, neutral syllable - and, aggressive syllable train category (GLMMs: maternal behavior, maternal activity score and centered pup age as fixed effects, pup ID and random observer factor as random factors, Poisson family with log link function).

To assess the potential influence of the number of singing males (ranging from 1-4) present in the colony on the syllable train production we likewise calculated five separate GLMMs, based on the entire dataset of 2015 and 2016 (N= 20 pups, N= 9369 syllable trains; GLMMs: number of tutors and centered pup age as fixed factors, pup ID and random observer factor as random factors, Poisson family with log link function).

The pups´ final repertoire size

We calculated two Spearman rank correlations to investigate the relationship between the number of maternal behaviors during babbling bouts respectively the female activity score and the pups` final repertoire size (N=10 mother-pup dyads from 2016).

Likewise, we calculated a Spearman rank correlation to investigate the relation between the number of singing males in a colony and the pups´ final repertoire size (N=20 pups from 2015 and 2016).

All statistical analyses were performed in R (RStudio 2018, version 3.5.2.). Each GLMM was checked for overdispersion using the R package 'blemco'. For each GLMM the marginal R2 (i.e. how much of the variance is explained by the fixed factors alone) and the conditional R2 (i.e. how much of the variance is explained by the fixed and random factors together) were calculated as a measure of effect size for GLMMs (following Nakagawa & Schielzeth, 2013). The detailed information about the multivariate models, the corresponding dispersion factors and the effect sizes can be found in the supporting information. Statistical differences were considered significant for p <0.05 (*p<0.05, **p<0.01, ***p<0.001).

RESULTS

The amount of vocal practice

Both the total number of maternal behaviors performed during babbling bouts and pup age had a significant influence on the daily babbling bout duration, whereas the maternal activity score had none (Fig. 1a: GLMM: maternal behaviors: t-value: 2.954, p=0.003; pup age: t-value=4.782, p < 0.001; maternal activity score: t-value: -0.141, p=0.88, Fig. 1a). The overall babbling phase duration was only significantly influenced by the maternal activity score (Fig. 1b: GLMM: maternal activity score: z-value: 2.086, p=0.03; maternal behaviors: z-value: -0.697, p=0.48, Fig. 1b). In contrast, the number of singing males neither influenced the babbling bout duration (GLMM: no. of singing males: t-value: 0.463, p=0.64, centered pup age: t-value: 6.742, p < 0.001) nor the babbling phase duration (GLMM: no. of singing males: z-value: 1.008, p=0.31).

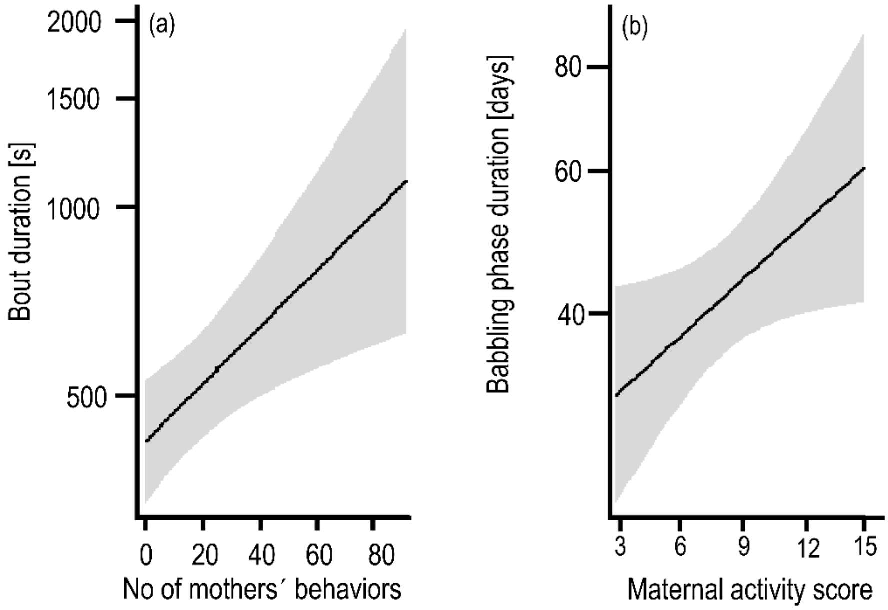


Fig. 1. The maternal behavior influences the amount of vocal practice.

The maternal behavior influences the a) daily vocal practice (i.e. babbling bout duration in seconds) and the b) overall practice behavior (i.e. babbling phase duration in days). The maternal activity score is a measure based on the daily mean number of maternal behaviors, ranging from 3 (lowest score) to 15 (highest score). Mean maternal score is 8.3 (For details see Material and Methods).

The babbling bout composition

The production of the following syllable train categories were positively influenced by maternal behavior: isolation call trains (GLMM: maternal behaviors: z-value: 2.536, p=0.01, maternal activity score: z-value: -2.438, p=0.01; centered pup age: z-value: 0.865, p=0.38), neutral syllable trains (GLMM: maternal behaviors: z-value: 3.196, p=0.009; maternal activity score: z-value: 3.196, p=0.01, centered pup age: z-value: -1.176, p=023) and aggressive syllable trains (GLMM: maternal behaviors: z-value: 2.055, p=0.03; female activity score: z-value: -0.522, p=0.6, centered pup age: z-value: 4.054, p < 0.001). The production of the other two syllable train categories was not influenced by maternal behaviors (see supporting information).

The number of singing males influenced the production of the following syllable train categories: the undifferentiated proto-syllable trains (GLMM: number of singing males: z-value: 2.163, p=0.03, centered pup age: z-value: 5.055, p < 0.001), the neutral syllable trains (GLMM: number of singing males: z-value: 0.312, p=0.002, centered pup age: z-value: 3.006, p=0.75) and the aggressive syllable trains (GLMM: number of singing males: z-value: 2.320, p=0.02, centered pup age: z-value: 4.931, p < 0.001). The production of the other two syllable train categories was not influenced by the number of singing males (see supporting information).

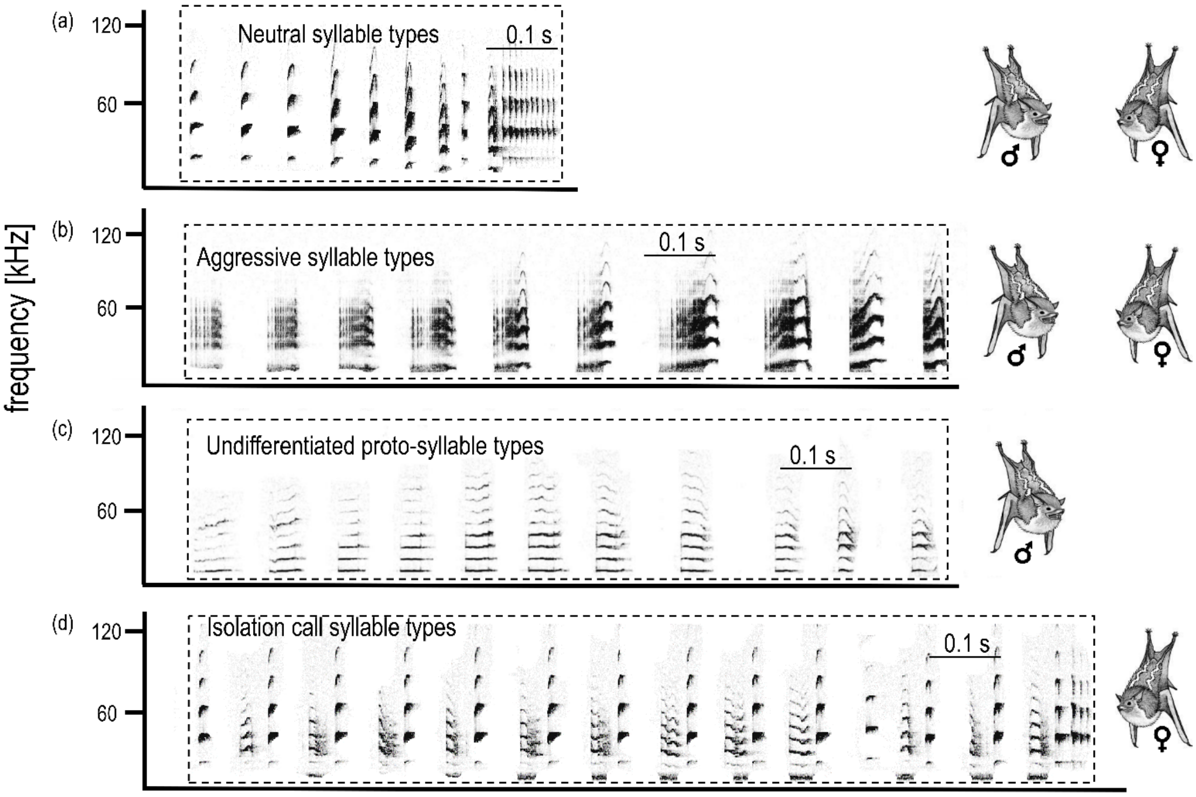


Fig. 2. The social environment shapes the babbling bout composition.

This figure illustrates which syllable train categories in babbling bouts are influenced by the maternal behavior (drawing of female bat) and/or the number of singing males (drawing of male bat). Panel (a) shows a neutral syllable train, composed of three neutral syllable types. With increasing numbers of singing males and maternal behaviors, more neutral syllable trains are produced. Panel (b) depicts an aggressive syllable train, composed of aggressive syllables which belong to the later adult male territorial song that is acquired through vocal production learning. With increasing numbers of singing males and maternal behaviors, more aggressive syllable trains are produced. Panel (c) depicts a train which is entirely composed of the undifferentiated proto-syllable type which is exclusively produced by juvenile pups and is not part of the adult vocal repertoire. With increasing numbers of singing males, more undifferentiated proto-syllable trains are produced. Panel (d) depicts the isolation call train composed of end syllables of the isolation call. With increasing number of maternal behaviors, more isolation call trains are produced. The spectrograms depict frequency [kHz] as a function of time [s] and were generated using a 1042 point fast Fourier transform and a Hamming window with 85% overlap. Drawing credit: Dr. C. A. S. Mumm.

The pups´ final repertoire size

The number of singing males did not influence the pups´ final repertoire size (Fig. 3a: Spearman rank correlation: r_s=0.02, p=0.9). Likewise, neither the maternal activity score (Spearman rank correlation: r_s=0.45, p=0.18) nor the number of maternal behaviors during babbling bouts (Spearman rank correlation: r_s=0.58, p=0.08) was significantly related to the pups´ final repertoire size (Fig. 3b & c).

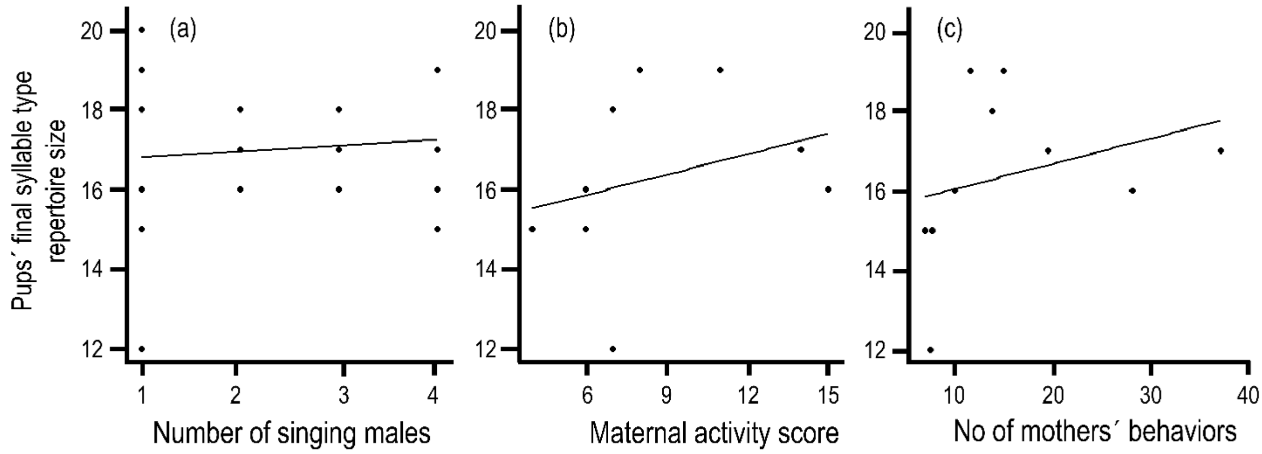


Fig. 3. The pups´ final syllable type repertoire size is not influenced by the social environment. The pups´ final syllable type repertoire size is neither influenced by the number of singing males (a) nor the maternal activity score (b) or the number of maternal behaviors (c). In a) the dataset includes 20 pups (dataset Panamá 2015 & Costa Rica 2016) whereas b) and c) include only 10 pups (dataset Costa Rica 2016) because maternal behavior was not quantified during the field season 2015.

DISCUSSION

Our study demonstrates that the both the direct and indirect social environment (i.e. mothers and singing males) shapes the pups´ vocal ontogeny. The amount of vocal practice, both on a daily basis and over the entire babbling phase, was influenced only by maternal behaviors, whereas the presence of adult males or other pups in the colony had no significant influence. More specifically, mothers who produced more behaviors during babbling bouts lead to a prolonged daily vocal practice of their pups (longer babbling bout duration) and a higher maternal activity score lead to a prolonged vocal practice period (longer babbling phase). Furthermore, the maternal behavior and the number of singing adult males influenced the production of specific syllable train categories, thus influencing which syllable types pups practiced most. Nevertheless, the pups' final repertoire size was not influenced by their direct or indirect social environment.

Social feedback shaping the vocal ontogeny can include behavioral and vocal interactions alike. In *S. bilineata,* the maternal social feedback is mainly provided through behaviors. Likewise, studies in other vocal production learners like songbirds showed that, besides vocal interactions, social feedback provided through behaviors shapes the vocal ontogeny, especially song learning. In the brown-headed cowbird, *Molothrus ater,* a specific maternal behavior impacts which acoustic elements are retained respectively discarded from the tutees song before crystallization occurs (West & King, 1988). Similarly, a recent study in zebra finches discovered that a certain behavior displayed by females in response to a singing juvenile had an impact on song learning success (Carouso-Peck & Goldstein, 2019). In human infants, social feedback during canonical babbling occurs principally through the acoustic modality (Gleitman & Newport, 2002; Goldstein, King, & West, 2003). For example, contingent social feedback coordinated with the infants` babbling led to increased production of more advanced vocalizations, such as canonical syllables, and to more resonant vowels (Goldstein, King, & West, 2003; Goldstein & Schwade, 2008). However, social interactions that shape the infant canonical babbling are not exclusively vocal but also non-verbal (Goldstein & Schwade, 2010). Behaviors such as touching the infant or offering play toys are sufficient to influence the infant canonical babbling if responses are contingent with the infants´ vocal output (Goldstein, King, & West, 2003; Goldstein & Schwade, 2008).

The social environment of *S. bilineata* pups influenced the composition of babbling bouts, namely the production of aggressive syllable trains, undifferentiated proto-syllable trains, neutral syllable trains and isolation call trains (Fig. 2). The most interesting result is that the production of aggressive syllable trains was influenced by both maternal behaviors and the number of singing males present in a colony (Fig. 2b). This is noteworthy because the aggressive syllable types constitute the later male territorial song which is acquired through vocal production learning (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Territorial songs are primarily used to deter potential rivals and are directed to other adult males (Behr & von Helversen, 2004; Behr et al., 2006). Males engage in daily 20-30 minutes sessions at dusk and dawn producing territorial songs (Behr & von Helversen, 2004; Eckenweber & Knörnschild, 2013; Tannenbaum, 1975), hence they provide a large amount of vocal input. A possible explanation why pups of colonies with many tutors produced more aggressive syllable trains might be that pups are influenced by the acoustic environment in which they are growing up. Furthermore, the number of maternal behaviors likewise lead to an increase in production of aggressive syllable trains. How exactly maternal behaviors are related to the production of aggressive syllable trains remains to be investigated. Whether an increased production of those trains, respectively increased practice of territorial song syllables leads to a higher learning success (e.g. more precise imitation of the tutor, higher song quality) remains to be investigated. If this turns out to be true, it would be beneficial for pups to grow up in colonies with many tutors and to have active mothers.

Interestingly, the number of adult males also influenced the production of undifferentiated proto-syllable trains (Fig. 2c). Undifferentiated proto-syllables are exclusive pup vocalizations and are not part of the adult vocal repertoire. They are highly variable and lead to vocal overproduction (Chapter 1). Undifferentiated proto-syllables can be understood as vocal play behavior and vocal exploration. They might play a role in gaining control over the vocal apparatus. Furthermore, they possibly have a labelling function to signal adult conspecifics, especially adult males, that babbling is a non-aggressive vocal practice behavior. This would allow pups to practice the syllable types of the territorial song which normally elicit aggressive reactions when produced by adult males (Behr, Knörnschild, & von Helversen, 2009). In big colonies, where many territorial males and mothers with their pups are present it can get confusing and noisy. In these situations, it might be necessary for pups to produce more of the undifferentiated proto-syllables to label their vocal output as harmless vocal practice behavior. The maternal behaviors also influence the production of isolation call trains (Fig. 2d). Behaviorally active mothers also have more active pups (see supporting information, additional results). Due to position changes in the day-roost while babbling (Table S1), pups are sometimes in considerable distance to their mothers. The isolation call elicits maternal care and is used for recognition (Knörnschild & von Helversen, 2008). Hence, the increased production of isolation call trains in babbling bouts might be related to increased pup movement during babbling and maternal interactive behavior in general.

In contrast to our hypothesis, the social environment did not influence the pups` final repertoire size. This might be partially explained by the fact that most pups (except for one) had very similar final syllable repertoire sizes. Pups were also very similar concerning the adult-like syllable type acquisition and production during babbling. Among the acquired adult-like syllable types were the aggressive song syllables belonging to the later territorial song. As aforementioned, this song is daily produced at dusk and dawn (Eckenweber & Knörnschild, 2013), hence provides ample auditory input. Therefore, missing syllable types are probably more interesting. The adult vocal repertoire is composed of 25 distinct syllable types (Chapter 1). The missing adult-like syllable types in all pups` final syllable repertoires mainly belonged to the affiliative syllable type category of which the adult courtship song is composed of. Even though adult males court females on a daily basis throughout the whole year, they do not produce complete courtship songs during pup ontogeny and therefore do not provide the entire set of courtship syllables to pups. It remains to be investigated how and when subadult bats acquire the courtship syllables and if the number of singing males plays a role only after the babbling phase already terminated.

Studies in human infants show that social interactions can affect speech acquisition and learning success (Goldstein, King, & West, 2003; Kuhl, 2003). Likewise, studies in songbirds demonstrate that social feedback influences song learning success (Beecher, 2017; Carouso-Peck & Goldstein, 2019; García, 2019). In children, social feedback facilitates the recognition of categorical speech perception during early speech acquisition (Kuhl, 2003, 2007) and can also influence speech production. For example, contingent interactions between mother and the babbling child lead to more fully resonant vowels and better pronounced consonants in the child’s oral output (Bloom, 1988; Goldstein, King, & West, 2003). Besides acquiring the phonetic repertoire of a language, a main function of canonical babbling is probably vocal practice to gain control over the vocal articulatory muscles. By imitating speech sound subunits infants match their own oral output to their perceptual input (Oller, 1980; Vihman, 1996, 2014). This daily rehearsal of syllables (i.e. speech sound subunits) leads to the reinforcement of the neuronal speech engram, a necessary prerequisite to acquire spoken language. Hence, the amount of vocal practice time is very likely to influence the speech acquisition success. Similarly, in *S. bilineata* pups babbling behavior probably has different functions and one might be vocal practice. During babbling pups may enhance the neuronal sensory-motor template by rehearsing production of syllable types, similar to infants.

Another function of social feedback is that social interactions are likely to evoke attention and arousal, two factors known to impact learning and memory. Especially for imitative learning, which is essential in language acquisition, joint attention plays a major role (Baldwin, 1995; Carpenter & Tomasello, 1995). For instance, attention of infants is much higher during sessions with a live tutor (i.e. person) compared to an animated tutor (i.e. video of a live person) (Kuhl, 2003). This suggests that increased attention and arousal could influence the quantity and the quality of speech information a child can process and memorize (Kuhl, 2003). Similarly, in songbirds attention during song learning was enhanced and facilitated by social interactions (Chen, Matheson & Sakata, 2016; García, 2019). In *S. bilineata*, the social interactions between mother and pup might likewise elicit increased attention and arousal in pups during babbling. Additional to the fact that the maternal behavior significantly increased the vocal practice time, the social interactions might, similar to infants, enhance the learning process, for example by enhancing memory while rehearsing.

Conclusions

Many studies demonstrate that the vocal ontogeny in non-human animals, respectively the infant speech acquisition, is strongly influenced by social feedback. Studies conducted in the wild are more realistic compared to studies in captivity because they take the influence of the entire social environment into account. Therefore, studies like ours are crucial for a better understanding of the interplay between social influences and vocal development. Furthermore, comparative research is crucial to understand the factors influencing infant speech acquisition. Investigating vocal ontogenetic processes in other vocal production learners with a pronounced practice phase could advance our understanding of underlying mechanisms required to evolve a complex communication system.

REFERENCES

Adret, P. (2004). Vocal imitation in blindfolded zebra finches (*Taeniopygia guttata*) is facilitated in the presence of a non-singing conspecific female. *Journal of ethology, 22*(1), 29-35.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. *Joint attention: Its origins and role in development*, 131-158.

Beecher, M. D. (2017). Birdsong learning as a social process. *Animal Behaviour, 124*, 233-246.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in ecology & evolution, 20*(3), 143-149.

Behr, O., Knörnschild, M., & von Helversen, O. (2009). Territorial counter-singing in male sac-winged bats (*Saccopteryx bilineata*): low-frequency songs trigger a stronger response. *Behavioral Ecology and Sociobiology, 63*(3), 433-442.

Behr, O., & von Helversen, O. (2004). Bat serenades - complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology, 56*(2), 106-115. doi:10.1007/s00265-004-0768-7

Behr, O., Von 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). *Behavioral Ecology, 17*(5), 810-817.

Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language, 15*(3), 469-480.

Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and Stereotypy of Free-Tailed Bat Songs. *PLoS One, 4*(8). doi: 10.1371/journal.pone.0006746.

Bradbury, J. W., & Emmons, L. H. (1974). Social organization of some Trinidad bats: I. Emballonuridae. *Zeitschrift für Tierpsychologie, 36*(1-5), 137-183.

Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology, 29*(4), 631-636. e633.

Carpenter, M., & Tomasello, M. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Social Development, 4*(3), 217-237.

Chaverri, G., Gillam, E. H., & Kunz, T. H. (2012). A call-and-response system facilitates group cohesion among disc-winged bats. *Behavioral Ecology, 24*(2), 481-487. doi:10.1093/beheco/ars188.

Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences, 113*(24), 6641-6646.

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

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

García, N. C. (2019). What have we recently learned about song learning and social interactions? *Behavioral Ecology*.

Gleitman, L. R., & Newport, E. L. (2002). The invention of language by children: Environmental and biological influences.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological science, 19*(5), 515-523.

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. *The Oxford handbook of developmental and comparative neuroscience*, 708-729.

Jahelková, H., Horáček, I., & Bartonička, T. (2008). The advertisement song of *Pipistrellus nathusii (*Chiroptera, Vespertilionidae): a complex message containing acoustic signatures of individuals. *Acta Chiropterologica, 10*(1), 103-126. doi:10.3161/150811008x331144.

King, A. P., Freeberg, T. M., & West, M. J. (1996). Social experience affects the process and outcome of vocal ontogeny in two populations of cowbirds (*Molothrus ater*). *Journal of Comparative Psychology, 110*(3), 276.

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*(9), 451-454. doi:10.1007/s00114-006-0127-9.

Knörnschild, M., Kalko, E. K. V., & Feifel, M. (2014). Male courtship displays and vocal communication in the polygynous bat *Carollia perspicillata*. *Behaviour, 151*(6), 781-798. doi:10.1163/1568539x-00003171.

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. doi:10.1098/rsbl.2009.0685.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Konishi, M. (1985). Birdsong: from behavior to neuron. *Annual Review of Neuroscience, 8*(1), 125-170.

Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(17), 9645-9646. doi:10.1073/pnas.1733998100.

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain? *Developmental science, 10*(1), 110-120.

Marler, P. (1970). Birdsong and Speech Development - Could There Be Parallels. *American Scientist, 58*(6), 669-673.

Mennill, D. J., Doucet, S. M., Newman, A. E., Williams, H., Moran, I. G., Thomas, I. P., . . . Norris, D. R. (2018). Wild birds learn songs from experimental vocal tutors. *Current Biology, 28*(20), 3273-3278. e3274.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution, 4*(2), 133-142.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nature Communications, 5*, 3379.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In J. F. K. a. C. A. F. G. H. Yeni-Komshian (Ed.), *Child phonology* (Vol. 1, pp. 93-112). New York: Academic Press, Inc.

Strauss, M., von Helversen, O., & Knörnschild, M. (2010). The ontogeny of courtship behaviours in bat pups (*Saccopteryx bilineata*). *Behaviour, 147*(5), 661-676.

Tannenbaum, B. R. (1975). *Reproductive strategies in the white-lined bat.* (PhD). Cornell University Ithaca NY USA,

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*: Blackwell Publishing.

Vihman, M. M. (2014). *Phonological development: The first two years*: Wiley-Blackwell Boston, MA.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat, *Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

West, M. J., & King, A. P. (1988). Vocalizations of juvenile cowbirds (*Molothrus ater ater*) evoke copulatory responses from females. *Developmental Psychobiology, 21*(6), 543-552.

Yancey, F. D., Goetze, J. R., & Jones, C. (1998). *Saccopteryx bilineata*. *Mammalian Species*(581), 1-5.

Supplementary Information for

Social influences shape the babbling behavior of the greater sac-winged bat pups, *Saccopteryx bilineata*

Authors: Ahana A. Fernandez, Nora Serve, Sarah-Cecil Fabian & Mirjam Knörnschild

Table of Contents:

Supporting results

This paragraph includes additional statistical tests not mentioned in the main manuscript. Furthermore, it includes the detailed information about the multivariate statistical models performed in this study.

GLMMs

Statistical differences were considered significant for p <0.05 (*p<0.05, **p<0.01, ***p<0.001).

Vocal practice

a) **Daily vocal practice (i.e. babbling bout duration)**

The mothers` influence on bout duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 babbling bouts (BB))

Dependent variable:

Babbling bout duration [s]

Fixed effects:

Number of maternal behaviors during babbling bout (1)

Maternal activity score (2)

Centered pup age (3)

Centered pup age: to improve the convergence of the GLMM pup age was centered by subtracting mean pup age from each pup age value.

Random effects:

pup ID

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Fixed effects	Estimate	t-value	p-value	Trajectory
(1)	0.010531	2.954	0.003	Increasing**
(2)	-0.004238	-0.141	0.888	ns
(3)	0.017953	4.782	< 0.001	Increasing***

Check for overdispersion: 0.76

Marginal R2:13.7 %;

Conditional R2: 30.0 %

The influence of the number of singing males on bout duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB)

Dependent variable:

Babbling bout duration [s]

Fixed effects:

Number of singing males (1-4)

Centered pup age

Random effects:

pup ID

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Independent variables	Estimate	t-value	p-value	Trajectory
No. of singing males	0.046189	0.463	0.64	ns
Centered pup age	0.023101	6.742	< 0.001	Increasing***

Check for overdispersion: 0.78

Marginal R2: 14.0 %;

Conditional R2: 25.6 %

The influence of other pups present on bout duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB)

Dependent variable:

Babbling bout duration [s]

Fixed effects:

Number of other pups: 1-9

Centered pup age

Random effects:

pup ID

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Independent variables	Estimate	t-value	p-value	Trajectory
No. of other pups	-0.005929	-0.028	0.97	ns
Centered pup age	0.023078	6.385	< 0.001	Increasing***

Check for overdispersion: 0.78

Marginal R2:14.0 %;

Conditional R2: 25.6 %

b) Overall vocal practice (i.e. babbling phase duration)

The mothers` influence on babbling phase duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB)

Dependent variable:

Babbling phase duration [days]

Fixed effects:

Number of maternal behaviors during babbling bout (1)

Maternal activity score (2)

Random effects:

Colony ID

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Fixed effects	Estimate	z-value	p-value	Trajectory
(1)	-0.006861	-0.697	0.48	ns
(2)	0.053042	2.086	0.037	Increasing*

Check for overdispersion: 1.30

Marginal R2:26.0 %

Conditional R2: 43.6 %

The influence of the number of singing males on babbling phase duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB)

Dependent variable:

Babbling bout duration [s]

Fixed effects:

Number of singing males (1-4)

Random effects:

Colony ID

Random observation level factor (to account for overdispersion)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Independent variables	Estimate	z-value	p-value	Trajectory
No. of singing males	0.1039	1.008	0.31	ns

Check for overdispersion: 1.01

Marginal R2: 9.5 %

Conditional R2: 35.7 %

The influence of other pups present on babbling phase duration

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB)

Dependent variable:

Babbling bout duration [s]

Fixed effects:

Number of other pups (1-9)

Random effects:

Colony ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Gamma (log link function)

Independent variables	Estimate	z-value	p-value	Trajectory
No. of other pups	-0.002295	-0.005	0.9	ns

Check for overdispersion: 1.01

Marginal R2: 0.02 %

Conditional R2: 45.1 %

Babbling bout composition

The mothers` influence on babbling bout composition

Dataset syllable trains & behavior (2016: Costa Rica): 10 mother-pup dyads, N= 77 BB, N= 5011 syllable trains)

Dependent variable:

Number of syllable trains for categories I-V (for each syllable train category we performed separate GLMMs, i.e. five independent GLMMs).

Fixed effects:

Number of maternal behaviors (1)

Maternal activity score (2)

Centered pup age (3)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Dependent variable	Fixed effects	Estimate	z-value	p-value	Trajectory
Isolation call trains	(1)	0.019541	2.536	0.011	Increasing*
	(2)	-0.095378	-2.438	0.014	Decreasing*
	(3)	0.005873	0.865	0.38	ns
Undifferentiated proto-syllable trains	(1)	0.013416	1.395	0.16	ns
	(2)	0.000167	0.004	0.99	ns
	(3)	0.022868	2.593	0.009	Increasing**
Neutral syllable trains	(1)	0.025441	3.196	0.001	Increasing**
	(2)	-0.044540	-1.176	0.23	ns
	(3)	-0.003858	-0.538	0.59	ns
Aggressive syllable trains	(1)	0.016389	2.055	0.03	Increasing*
	(2)	-0.018347	-0.522	0.60	ns
	(3)	0.029194	4.054	< 0.001	Increasing***
Affiliative syllable train	(1)	0.014859	1.402	0.16	ns
	(2)	0.038972	0.670	0.5	ns
	(3)	0.025308	2.552	0.01	Increasing*

Check for overdispersion: for each GLMM separately

GLMM isolation call trains: 1.03

Marginal R2:18.3 %

Conditional R2: 30.0 %

GLMM undifferentiated proto-syllable trains: 1.07

Marginal R2:18.0 %

Conditional R2: 32.5 %

GLMM neutral syllable trains: 1.08

Marginal R2:13.1 %

Conditional R2: 22.0 %

GLMM aggressive syllable trains: 1.09

Marginal R2:32.0 %

Conditional R2: 40.0 %

GLMM affiliative syllable trains: 1.06

Marginal R2:20.3 %

Conditional R2: 50.2 %

The influence of the number of singing males on babbling bout composition

Dataset syllable trains & indirect social environment (2015: Panama & 2016: Costa Rica): 19 pups, N=8 colonies, N=178 BB, N=9169 syllable trains).

Dependent variable:

Number of syllable trains for categories I-V (for each syllable train category we performed separate GLMMs, i.e. five independent GLMMs).

Fixed effects:

Number of singing males present (1)

Pup age (2)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Dependent variable	Fixed effects	Estimate	z-value	p-value	Trajectory
Isolation call trains	(1)	-0.067275	-0.678	0.49	ns
	(2)	0.003322	0.813	0.41	ns
Undifferentiated proto-syllable trains	(1)	0.202266	2.163	0.03	Increasing*
	(2)	0.025980	5.055	< 0.001	Increasing***
Neutral syllable trains	(1)	0.028260	0.312	0.002	Increasing**
	(2)	0.013427	3.006	0.75	ns
Aggressive syllable trains	(1)	0.383648	2.320	0.02	Increasing*
	(2)	0.025133	4.931	< 0.001	Increasing***
Affiliative syllable train	(1)	0.095885	0.615	0.53	ns
	(2)	0.029034	5.062	< 0.001	Increasing***

Check for overdispersion: for each GLMM separately

GLMM isolation call trains: 1.01

Marginal R2: 0.9 %

Conditional R2: 21.4 %

GLMM undifferentiated proto-syllable trains: 1.07

Marginal R2:19.0 %

Conditional R2: 35.0 %

GLMM neutral syllable trains: 1.06

Marginal R2: 5.1 %

Conditional R2: 17.5 %

GLMM aggressive syllable trains: 1.05

Marginal R2: 22.0 %

Conditional R2: 73.8 %

GLMM affiliative syllable trains: 1.05

Marginal R2: 16.0 %

Conditional R2: 56.0 %

The influence of the number of other pups on babbling bout composition

Dataset syllable trains & indirect social environment (2015: Panama & 2016: Costa Rica): 19 pups, N=8 colonies, N=178 BB, N=9169 syllable trains).

Dependent variable:

Number of syllable trains for categories I-V (for each syllable type train category we performed separate GLMMs, i.e. five independent GLMMs).

Fixed effects:

Number of other pups (1)

Pup age (2)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Dependent variable	Fixed effects	Estimate	z-value	p-value	Trajectory
Isolation call trains	(1)	-0.035671	-0.684	0.49	ns
	(2)	0.003290	0.806	0.42	ns
Undifferentiated proto-syllable trains	(1)	0.076027	1.469	0.14	ns
	(2)	0.026726	5.184	< 0.001	Increasing***
Neutral syllable trains	(1)	0.005342	0.113	0.91	ns
	(2)	0.013552	3.035	0.002	Increasing**
Aggressive syllable trains	(1)	0.176024	1.927	0.053	Trend (increasing)
	(2)	0.025175	4.936	< 0.001	Increasing***
Affiliative syllable train	(1)	0.02857	0.334	0.73	ns
	(2)	0.02956	5.142	< 0.001	Increasing***

Check for overdispersion:

GLMM isolation call trains: 1.01

Marginal R2: 0.9 %

Conditional R2: 21.5 %

GLMM undifferentiated proto-syllable trains: 1.07

Marginal R2: 17.4 %

Conditional R2: 34.4 %

GLMM neutral syllable trains: 1.06

Marginal R2: 5.5 %

Conditional R2: 17.4 %

GLMM aggressive syllable trains: 1.05

Marginal R2:19.6 %

Conditional R2: 72.0 %

GLMM affiliative syllable trains: 1.05

Marginal R2: 13.7 %

Conditional R2: 53.4 %

Additional results on pup behavior development

Besides the social influences on the vocal ontogeny we investigated if maternal behaviors also influenced the behavioral development of the pup. We hypothesized that the number of maternal behaviors would positively influence the number of behaviors produced by the pup.

We performed a GLMM to analyze the influence of the maternal behaviors, pup age and social group size on the pups' behaviors (fixed effects: number of maternal behaviors, centered pup age and social group size, random effects: pup ID and observation level random effect, Poisson family with log link function).

To test for potential sex differences in behaviors produced by pups and by mothers we additionally performed Mann-Whitney-U tests.

Influence of female behavior on pup behavioral development

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB; N=12´796 behaviors)

Dependent variable:

Pup behaviors

Fixed effects:

Mothers´ behaviors (1)

Centered pup age (2)

Social group size: (i.e. number of tutors and other pups present in colony) (3)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Fixed effects	Estimate	z-value	p-value	Trajectory
(1)	0.039592	13.823	< 0.001	Increasing***
(2)	0.011098	3.220	0.001	Increasing**
(3)	0.035335	1.337	0.18	Increasing

Check for overdispersion: 1.06

Marginal R2:62.4 %

Conditional R2: 81.8 %

No sex differences in behaviors

There was no significant difference between the total number of behaviors between male and female pups (Whitney-U-Test: W=2067, p=0.70). Likewise, there was no difference between the number of behaviors displayed by mothers depending on the sex of their pups (Whitney-U-Test: W=1945, p =0.83).

Some of the behaviors occurring during babbling are reminiscent of behaviors produced by adult males during courtship. For example, how adult males approach adult females (by crawling towards them) is very much reminiscent of the sequential interactive "crawl towards" and "crawl away" behavior observed between mother and pups during babbling bouts (Fig. S1, (c) & (d)). Other interactions often involved the pup hovering in front of the mother and the mother switching position within the day roost as a response. This behavior provoked that the pup followed the mother by also performing a short flight landing close to her or combining a short flight with another hover display (Fig. S1).

Pups´ hover displays

Moreover, we investigated the behavioral display "hover" separately since this is the typical courtship display of adult males (N=17 mother-pup dyads (2 dyads were excluded since two females were not observed hovering), N=139 babbling bouts, N=586 hover displays). We investigated if the mothers´ hover behavior and the social group size (i.e. sometimes adult males hovered in front of pups before foraging and sometimes other pups hovered close to each other which might influence the hover display rate) influenced the pups´ hover display behavior.

We calculated two Mann-Whitney-U-Tests to investigate potential differences between male and female pups, respectively the number of maternal hover flights depending on pup sex. Furthermore, to analyze the influence of maternal hover displays on the number of pups´ hover displays during babbling bouts we calculated a GLMM (number of the females' hover displays, pup age and social group size (i.e. number of tutors and number of other pups present in the colony, GLMM 17) as fixed factors, pup ID and a randomized observer factor as random factors, Poisson family with log-link function).

Hover displays

No significant difference was detected between the total number of hovering flights between male and female pups (Whitney-U-Test: W=1845.5, p=0.46) nor the number of hover flights of the females depending on pup sex (Whitney-U-Test: W=1702.5, p=0.10).

Dataset behaviors (2016: Costa Rica & 2017: Panama): 19 mother-pup dyads, N=183 BB; N=586 hover flights)

Dependent variable:

Hover flights (pup)

Fixed effects:

Number of hover flights (mother) (1)

Pup age (2)

Social group size (3)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Fixed effects	Estimate	z-value	p-value	Trajectory
(1)	0.259519	3.964	< 0.001	Increasing***
(2)	0.015175	1.793	0.07	Trend (Increasing)
(3)	0.145295	1.548	0.12	Increasing

Check for overdispersion: 0.92

Marginal R2:13.8 %

Conditional R2: 88.3 %

Interactive behavior sequences

Furthermore, we investigated behavioral interactions occurring exclusively between mother and pup. An interaction was defined as such if a behavior (e.g. pup hovers) succeeded within maximal one second after a previous behavior (e.g. mother flies next to pup). They were either short (i.e. initial behavior – response behavior) or were longer, including a suite of successive behaviors. We established four categories, ranging from simple sequence to a suite of multiple interactions (N=17 mother-pup dyads, N=139 babbling bouts, N=2199 interactions). We analyzed if the number of total interactions occurring during a babbling bout was influenced by the number of interactions started by the mother. Additionally, we analyzed if the duration of behavioral interactions occurring during babbling bouts were influenced by the number of maternal behaviors.

We performed a GLMM to analyze if the total number of interactions was influenced by the number of interactions initiated by the mother (number of interactions started by female, pup age and pup sex as covariates, pup ID and Unit (to deal with large eigenvalue) as random factors, Poisson family with log link function). We performed four separate GLMMs to investigate if the duration of the interactive sequences was influenced by the maternal behaviors performed during a babbling bout (number of maternal behaviors and centered pup age as covariates, pup ID and an additional random observer effect as random factors, Poisson family with log link function).

Total number of interactions

Dataset interactive behaviors (2016: Costa Rica & 2017: Panama): 17 mother-pup dyads, N=139 BB, N=2´199 interactions)

Dependent variable:

Number of total interactions

Fixed effects:

Interactions started by mother (1)

Pup age (2)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Fixed effects	Estimate	z-value	p-value	Trajectory
(1)	0.102142	6.567	< 0.001	Increasing***
(2)	0.029129	5.167	< 0.001	Increasing***

Check for overdispersion: 1.06

Marginal R2:42.5 %

Conditional R2: 72.3 %

Duration of interactive sequences

Dataset interactive sequences (2016: Costa Rica & 2017: Panama): 17 mother-pup dyads, N=102 BB; N=1800 interactive sequences cat.1-4)

Dependent variable:

Number of interactions belonging to category I-IV (for each category of interactive sequences a separate GLMM was calculated, i.e. four independent GLMMs).

Fixed effects:

Maternal behaviors (1)

Centered pup age (2)

Random effects:

pup ID

Random observation level factor

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: Poisson (log link function)

Dependent variable	Fixed effects	Estimate	z-value	p-value	Trajectory
Interactive sequence cat. I	(1)	0.006714	0.534	0.59	ns
	(2)	0.029238	4.968	< 0.001	Increasing***
Interactive sequence cat. II	(1)	0.03464	2.187	0.02	Increasing*
	(2)	0.02598	3.146	0.001	Increasing**
Interactive sequence cat. III	(1)	0.03229	2.229	0.02	Increasing*
	(2)	0.01702	0.904	0.36	ns
Interactive sequence cat. IV	(1)	0.07002	3.172	0.001	Increasing**
	(2)	0.02433	1.916	0.05	Trend (Increasing)

Check for overdispersion sequence 1: 1.03

Marginal R2: 62.4 %

Conditional R2: 81.8 %

Check for overdispersion sequence 2: 1.02

Model has a singular fit, convergence is 0.

Marginal R2: 17.3 %

Conditional R2: 27.6 %

Check for overdispersion sequence 3: 0.78

Marginal R2: 8.8 %

Conditional R2: 39.5 %

Check for overdispersion sequence 4: 1.04

Model has a singular fit, convergence is 0.

Marginal R2: 8.2 %

Conditional R2: 16.4 %

Short Discussion on pup behavior development

Our results showed that the maternal behaviors influenced the number of behaviors a pup produced during a babbling bout. Whether this influence is still significant after accounting for pup age remains to be investigated. Some behaviors observed during babbling are reminiscent of the adult male courtship displays such as hovering in front of an individual or short flights landing close to another individual, which require a good maneuverability and therefore, probably a fair amount of practice (Fig. S1, p. 157). It is likely that the practice of behaviors during pup ontogeny influences the endurance, strength and skillfulness during adulthood. Interestingly, we detected no sex-specific differences in the behavior of pups or in the behavior of mothers towards their pups, although the behaviors of both sexes differ remarkably as adults (Voigt et al., 2008).

Furthermore, our findings show that pups practice interactive behavioral sequences with their mothers during babbling bouts. Interactive sequences between mother and pup resemble interactive courtship displays to some degree (Strauss, von Helversen, & Knörnschild, 2010). Furthermore, according to our prediction, the interactions became more longer in duration with increasing pup age. This is most likely a result of an increase in endurance in pups.

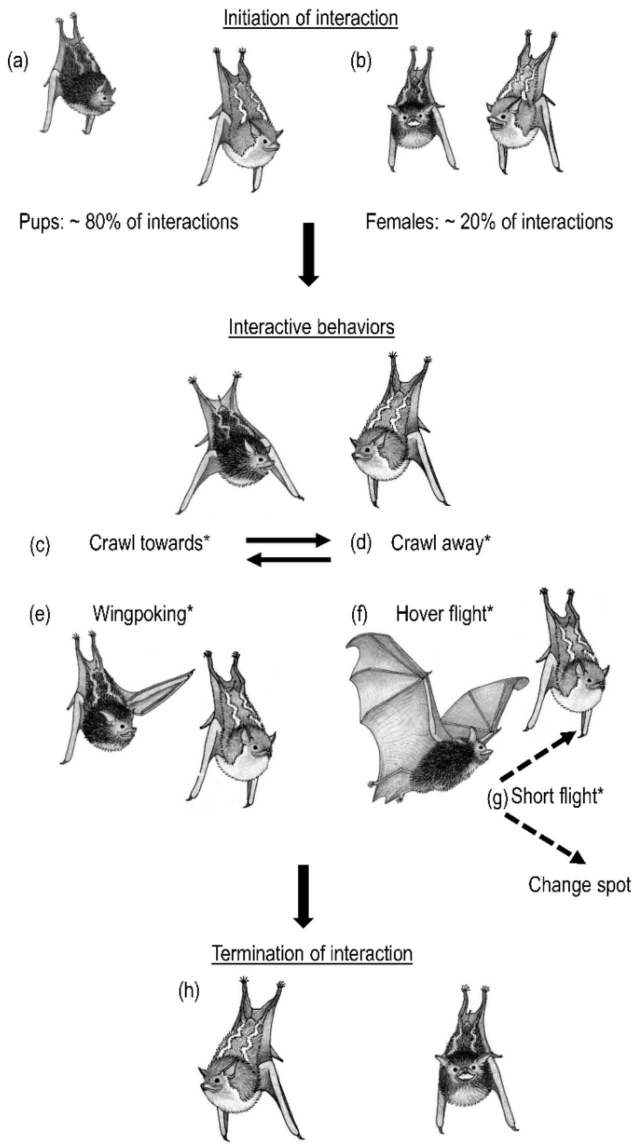


Fig. S1. Behavioral interactions between mother and pup during a babbling bout.

The figure illustrates the most common behaviors occurring during interactions between mother and pup. In the majority of cases (80%), the interaction was initiated by the pup, normally by approaching the mother while babbling (a). A mother initiates interactions by hovering in front of or landing next to her pup (b). Afterwards, mother and pup engage in a sequence of behaviors, the most common of which are crawl towards (c) and crawl away (d) which frequently occur in an alternating sequence (illustrated by two black arrows). This leads to the impression that mother and pup are “circling” around each other. In interactive sequences the pup often gently pokes its mother with the folded forearm (e), a behavior which is not frequently produced by females towards pups. Hover flights (f) in front of the interactive partner are the most conspicuous behavior, followed by the short flights either next to the interaction partner or within the day roost (g). With increasing pup age, short flights get progressively longer, e.g. pups might even briefly leave the day roost and land on an adjacent tree. The interaction between mother and pup terminates if one of the interaction partners ceases to produce a succeeding behaviors (within max. 1 sec). The pup may still continue to babble (h) while perching in the day roost without showing any behavioral display. Drawing credit: Dr. C. A. S. Mumm

Supporting methods:

Table S1. Ethogram of behaviors during babbling

Behavior	Description	Observed in:
crawl away*	One individual crawls a certain distance (> 2 cm) away from the other.	P / M
crawl towards*	One individual crawls towards another individual (approaches up to 2 cm).	P / M
hover*	Hover flight in front of an individual (for a few seconds).	P / M
wing-poke*	Poking a conspecific with wrist. Pups normally repeatedly poke their mothers. Mothers usually poke the pup only once to terminate the babbling bout. Includes wing pokes with and without physical contact.	P / M
short flight*	Individual flies to another spot in the day roost, or individual flies towards interaction partner).	P / M
rocking	Rocking with entire body from side to side; also used when mother wants juvenile to detach from teat.	P / M
wing stretch	Wing is stretched out completely.	P / M
head stretch	Head and neck are conspicuously bent and stretched to either left or right side.	P
push up	Push wrists against wall and push body up.	P / M
dribble	Pound with wrists against wall, very fast and repetitive.	P
wrist lift	Pound with wrists against wall once, not hitting other individual.	P
sit and babble	Babbling.	P

Table S1: Description of the behaviors observed during babbling bouts. The behaviors labelled with a * either occurred as solitary behavior or as behavioral sequence (i.e. interaction). The third column depicts if the behavior was observed in the mother (M), the (P) and or both.

References

Strauss, M., von Helversen, O., & Knörnschild, M. (2010). The ontogeny of courtship behaviours in bat pups (*Saccopteryx bilineata*). *Behaviour, 147*(5), 661-676.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat, *Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

Pup directed vocalizations of adult females and males in a vocal learning bat

Authors: Ahana Aurora Fernandez and Mirjam Knörnschild

ABSTRACT

In humans, infant-directed speech – or motherese –facilitates language learning and is socially beneficial by increasing attention and arousal in the child. It is characterized by high pitch, expanded intonation contours and slower speech tempo. Furthermore, the vocal timbre (i.e. 'color' of voice) of motherese differs from the timbre of adult-directed speech. In non-human animals, pup-directed vocalizations are produced during social interactions between adults and young and are often female-biased. The greater sac-winged bat, *Saccopteryx bilineata,* is a vocal production learner with a large vocal repertoire that is acquired during ontogeny. We investigated pup-directed vocalizations of adult males and females during the pups' vocal ontogeny. We compared acoustic features between female pup-directed and adult-directed vocalizations and demonstrated that they differed in timbre and peak frequency. Furthermore, we described pup-directed vocalizations of adult males. During the ontogenetic period when pups' isolation calls (used to solicit maternal care) are converging towards each other to form a group signature, adult males also produce isolation calls. Pups' isolation calls are acoustically more similar to those of males from the same social group than to other males. This suggests that adult isolation calls could serve as template for the group signature encoded in pup isolation calls. In conclusion, our novel findings indicate that parent-offspring communication in bats is more complex and multifaceted than previously thought, with female pup-directed vocalizations reminiscent of human motherese and male pup-directed vocalizations that may facilitate the transmission of a vocal signature across generations.

Keywords: motherese, vocal ontogeny, timbre, maternal directive call, pup-directed male vocalization

INTRODUCTION

Infant- directed vocalizations are produced in many birds and mammals, primates (e.g. Whitham, Gerald, & Maestripieri, 2007), bats (e.g. Esser & Schmidt, 1989), seals (Charrier, Mathevon, & Jouventin, 2001), cliff swallows (Beecher, Stoddard, & Loesche, 1985), and king penguins (Jouventin, Aubin, & Lengagne, 1999). The function of infant-directed vocalizations is to mediate social interactions between adults and young (e.g. parent-offspring reunions) and to influence the vocal development of offspring (e.g. call reinforcement) (Balcombe & McCracken, 1992; Charrier, Mathevon, & Jouventin, 2001; Takahashi et al., 2015; Whitham, Gerald, & Maestripieri, 2007). In humans, the use of infant-directed speech by which adults address the child is a well-known phenomenon (Fernald & Kuhl, 1987; Kuhl et al., 1997). This infant-directed speech – or motherese - is characterized by unique universal prosodic features such as higher pitch, increased frequency range and slow tempo and is significantly different from adult-directed speech (Broesch & Bryant, 2015; Grieser & Kuhl, 1988). The prosodic attributes of motherese support speech learning in infants (e.g. Kuhl et al., 1997) and promote the infants´ attention towards the acoustic input (Grieser & Kuhl, 1988). Besides the differences in general acoustic features a recent study reported that the timbre (i.e. the unique tone 'color' of a voice or an instrument) of motherese is significantly different from adult directed speech timbre (Piazza, Iordan, & Lew-Williams, 2017). Both, infant-directed vocalizations in animals and motherese in humans are thus specific acoustic signals which influence young-caregiver interactions. To our knowledge there is only one study in squirrel monkeys comparing the acoustic parameters between infant-directed calls and other adult vocalizations and discussing the results in relation to motherese in human infants (Biben, Symmes, & Bernhards, 1989). Moreover, differences in timbre between infant-directed and adult-directed vocalizations in non-human animals have never been addressed.

In this study, we want to investigate if we can detect a phenomenon reminiscent of motherese in infant-directed female vocalizations of the greater sac-winged bat, *Saccopteryx bilineata.* This highly social bat species lives in stable perennial groups (i.e. colonies) and possesses a large vocal repertoire (reviewed in Voigt et al., 2008). *S. bilineata* is a vocal production learner (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010, 2012) and exhibits a distinct vocal practice phase during ontogeny (Knörnschild, Behr, & von Helversen, 2006). During vocal ontogeny, mothers produce a so-called maternal directive call (MD, Fig. 1A) to communicate with their single pups (Knörnschild & von Helversen, 2008). Like all bat pups studied to date, *S. bilineata* pups produce isolation calls (ICs; Fig. 1B) to solicit maternal care (Knörnschild &

von Helversen, 2008). Pup ICs encode information about individual identity, age and social group affiliation (Fernandez & Knörnschild, 2017; Knörnschild, Nagy, Metz, Mayer & von Helversen, 2012; Knörnschild & von Helversen, 2008). During ontogeny, ICs of pups from the same social group become progressively more similar to one another, i.e. develop a group signature based on social modification (Knörnschild, Nagy, Metz, Mayer & von Helversen, 2012). Recent new observations suggest that adult males also produce pup-directed vocalizations that resemble normal ICs.

Since both adult females and males produce pup-directed vocalizations in *S. bilineata*, we wanted to investigate (1) whether pup-directed and adult-directed female vocalizations differ in their acoustic characteristics and (2) whether pup-directed vocalizations of adult males have the potential to influence the pups' vocal ontogeny. We hypothesized that the acoustic characteristics of female MDs, including timbre, would differ from those of adult-directed vocalizations produced by the same females. Furthermore, since MDs often occur during mother-pup reunions, we predicted that MDs contain an individual signature. Additionally, we hypothesized that ICs of pups are more similar to ICs of adult males from the same social group than to ICs of adult males from different social groups.

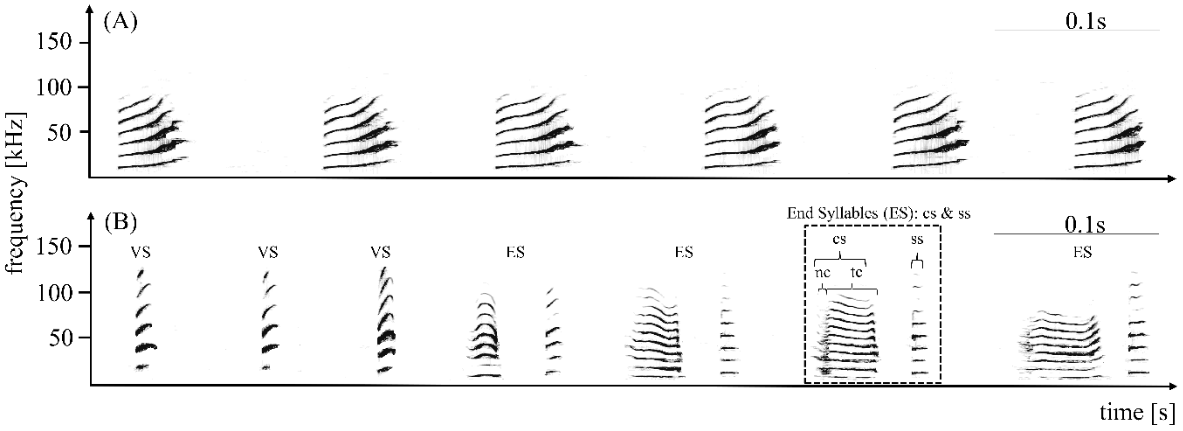


Fig.1. Pup-directed vocalizations from an adult female (A) and an adult male (B).

Panel A shows six successive directive calls produced by an adult female (the mother of the addressed pup), panel B shows one multisyllabic isolation call produced by an adult male. Maternal directive calls (A) are always produced in monosyllabic series, from three up to 15 syllables. An isolation call (B) is composed of simple tonal calls (variable simple syllable: VS) followed by the end syllables (ES). End syllables are composed of two syllable types, the composite (cs) and the stereotyped short syllable (ss). The cs part is further composed of a facultative noisy part (nc) succeeded by a tonal part (tc). Several simple frequency modulated syllables followed by several "end syllables" result in a typical isolation call. The spectrograms depict frequency (in kHz) as a function over time (in seconds) and were generated using a 1042 point fast Fourier transform and a Hamming window with 87.5 % overlap.

METHODS

Study sites and animals

We conducted sound recordings at three different locations in Central America during three consecutive field seasons (May-September in 2015-2017). We recorded the vocal and social behaviors of pups and pup-directed vocalizations of adult males and females *at libitum* throughout the pups' ontogeny (i.e. from birth until weaning at 10-12 weeks of age). In 2015, we conducted sound recordings at Barro Colorado Island, a field station of the Smithsonian Tropical Research Institute located in the Gatún lake of the Panamá Canal. We recorded vocalizations of six females belonging to four colonies. In 2016, we conducted sound recordings in the natural reserve Curú in Costa Rica and obtained vocalizations from seven females belonging to three colonies. Moreover, we recorded pup-directed vocalizations of 11 adult males in four colonies. In 2017, we conducted sound recordings in Gamboa, a field station of the Smithsonian Tropical Research Institute which is located at the Panamá Canal. We recorded pup-directed vocalizations of 11 adult males from three colonies. In total, we recorded pup-directed vocalizations from 13 females (maternal directive calls) and 22 males (ICs). For subsequent acoustic analyses we only included male IC recordings with good signal-to-noise ratio which reduced our sample size to 8 males. For 6 females, we had additional recordings of adult-directed vocalizations (simple calls and echolocation calls). We also recorded isolation calls of 14 pups.

Sound recordings

All recordings were performed throughout the day in the day-roosts which were located in tree cavities or on the outside of man-made structures. Focal recordings were feasible because the bats were individually banded with colored plastic rings on their forearms (see supplementary information). Furthermore, the colonies are part of a long-term project and bats are well habituated to human observers allowing close-range (2-4 meters) recordings and observations. Vocalizations were recorded using a high-quality ultrasonic sound recording equipment (500kHz sampling rate, 16-bit depth resolution, for details see supplementary material). The recording set-up consisted of a microphone (Avisoft UltraSoundGate 116Hm, with condenser microphone CM16, frequency range 1-200kHz ± 3dB) connected to a laptop (Lenovo S21e) running the software Avisoft RECORDER (v4.2.05 R. Specht, Avisoft Bioacoustics, Glienecke, Germany).

Pup-directed and adult-directed female vocalizations

Adult females produced both pup-directed calls (MDs) and adult-directed calls (short variable calls and echolocation calls). MD calls consist of frequency modulated tonal syllables (i.e. smallest acoustic unit surrounded by silence) which are produced in sequences of up to 15 calls (min: three calls; Fig. 1A). To investigate whether the acoustic characteristics of pup-directed versus adult-directed female vocalizations differed we analyzed not only MDs but also a set of adult-directed vocalizations from the same females, namely short variable calls (VS) and echolocation calls (EC); both were summarized as simple calls (SI). In total, we analyzed 26 MD call type sequences and 26 simple call type sequences from six females (range per female: 4-5 MD and simple call type sequences, for each call type the same number of call sequences was analyzed per female). Additionally, we analyzed 120 MD call type sequences composed of at least three syllables from 13 females (range: 7-12 MD call type sequences per female) belonging to seven colonies to investigate whether they encoded an individual signature.

Isolation call recordings

We analyzed 120 ICs of 14 pups (range: 6-10 calls per pup) and 39 ICs of 8 males from six colonies (range: 4-9 calls per male). The sound recordings of adult males were challenging to obtain because it was not predictable whether a male would produce an IC after a pup ceased its IC production. Moreover, males did not direct their isolation calls towards a specific pup. Hence, recording ICs from males required a fast change of microphone orientation (i.e. from pup to male) which resulted in fewer recordings with sufficient quality for subsequent acoustic analyses compared to ICs produced by pups.

Acoustic analyses

Each sound file was prepared in Cool Edit (Cool Edit 2000 Inc., Syntrillium Software Corporation P. O. Box 62255, Phoenix, AZ, USA) for subsequent acoustic measurements (see supplementary information). The acoustic analyses of ICs and MD calls were conducted using the software Avisoft SasLab Pro (v.5.2.09; R. Specht, Avisoft Bioacoustics, Glienicke, Germany).

For the pup-directed and adult-directed female vocalizations we extracted acoustic features that were based on linear-frequency cepstral coefficients (LFCCs) since those capture important acoustic characteristics of bat vocalizations (Knörnschild, Blüml, Steidl, Eckenweber, & Nagy, 2017). Each LFCC describes the spectral properties of an entire acoustic signal, comprising its most important features in a compact form. LFCC extraction is comparable to the MFCC extraction (mel frequency cepstral coefficient) used in human voice recognition (reviewed in Jain & Sharma, 2013) but it uses a linear scale instead of the mel scale to account for the bats' high frequency hearing. Extracted acoustic features represent not only common acoustic parameters such as peak frequency but also the timbre in a voice (Piazza, Iordan, & Lew-Williams, 2017).We used a customized MATLAB script in the toolbox "voicebox" (v. R2014a) for the feature extraction. Each vocalization sequence (i.e. MD sequence and SI sequence) was composed of three syllables containing the first three harmonics (F0-F2). Because we compared different call types with different durations (i.e. average simple call duration: 0.01s versus average MD call duration: 0.03s) we adapted the frame length of the feature extraction accordingly (i.e. MD calls: 24ms, SI calls: 8ms) to obtain comparable amounts of information. We extracted 20 LFCCs from each sequence and used them for subsequent statistical analyses. Furthermore, we measured the minimum, maximum and peak frequencies for each call type (MD, EC, VS).

To test for an individual signature in MDs we measured several temporal and spectral parameters for each syllable (n=120 MD sequences, see supplementary information). Principal component analyses were performed on the original acoustic parameters and derived acoustic parameters were used for subsequent statistical analyses (see supplementary information).

In the case of ICs we focused our analyses on the end syllables because former studies found that both the individual and the group signature are encoded in the end syllables (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012; Knörnschild & von Helversen, 2008). For each syllable type or part (Fig. 1B), we measured several temporal and spectral parameters (see supplementary information). We measured at least three end syllables per IC and subsequently averaged measurements per syllable type and part to minimize temporal dependence among syllable produced in direct succession. Principal component analyses (PCAs) were performed to reduce multicollinearity between original parameters and to obtain uncorrelated derived acoustic parameters (see supplementary information). Additionally, we extracted LFCCs of each IC. To obtain comparable acoustic features for each IC we extracted features from the first three harmonics (F0-F2) of the end syllables (without the noisy part since it was not always present). For each end syllable sequence we extracted 5 LFCCs using overlapping 6ms frames. A set of original acoustic parameters, derived parameters from the PCA and extracted LFCCs was used for subsequent multivariate analyses (see supplementary information).

Statistical analyses

We first conducted a multivariate GLM (with female ID, call type and their interaction as fixed factors) in which all acoustic features (LFCC1-20) and three original parameters (peak frequency, minimum and maximum frequency of the entire signal) were included. Subsequently, we selected the dependent variables which showed the same pattern for all females (no overlapping estimated marginal means for ID and call type, i.e. the differences between call types were all either de- or increasing) to calculate a second multivariate GLM with the same fixed factors as the first GLM. Six features (LFCC 2, 5, 6, 7, 9, and 12) and peak frequency were included as dependent variables in our second GLM. Minimum and maximum frequencies were not included because they were strongly correlated with peak frequency.

To test for the existence of an individual signature in maternal directive calls, we performed a DFA (n=120 MD call sequences from 13 females). We adjusted the DFA to the unequal number of analyzed call sequences per female by computing group sizes based on prior probabilities. We used a cross-validation procedure to estimate the correct classification success (n-1 cross-validation procedure), which classified each sequence based on discriminant functions established with all sequences except the one being classified. We selected one original acoustic parameter, namely duration, and five derived parameters, namely frequency curvature 1-3 and entropy curvature 1-2 (see supplementary information). All parameters were checked for multicollinearity and included simultaneously into the DFA.

To assess the acoustic similarity between ICs of pups and males we performed a discriminant function analysis (DFA) and subsequently calculated the Euclidean distances between individual centroids in the DFA signal space (see supplementary information). For each pup, we calculated the distance between itself and the male from the same colony and the average distance to all other males. Distances were compared with a paired Wilcoxon test. All statistical analyses were conducted in SPSS (v.20; IBM SPSS Statistics Chicago, IL, USA) and R (RStudio 2018, version 3.5.2).

RESULTS

Acoustic differences between pup-directed and adult-directed female vocalizations

Pup-directed and adult-directed female vocalizations differed significantly in their acoustic parameters (F (1,40) = 9.73, p < 0.001, η^2=0.66, Fig. 2A) whereas female ID had no significant effect (ID: F(5,40)=0.93, p = 0.57, η^2=0.15; call type*ID: F(5,40)=1.30, p=0.20, η^2=0.20). Results for each acoustic parameter a are shown in Table 1. Details on the GLMs (Table S1 & S2) and additional paired Wilcoxon tests can be found in the supplementary information.

Table 1. Difference in acoustic parameters between pup-directed and adult-directed vocalizations of females.

Dependent variable	F-value (1,40)	η^2	p-value	Trajectory: adult-directed to pup-directed
LFCC2	21.074	0.34	< 0.001	Increase
LFCC5	46.032	0.53	< 0.001	Increase
LFCC6	22.161	0.35	< 0.001	Increase
LFCC7	7.682	0.16	0.008	Increase
LFCC9	24.028	0.37	< 0.001	Increase
LFCC12	9.127	0.18	0.004	Increase
Peak freq.	23.295	0.36	< 0.001	Decrease

Table 1: Multivariate GLM results for 7 dependent variables and call type, the only independent variable which had a significant influence. Pup directed=maternal directive calls. Adult-directed=simple calls.

No individual signature in maternal directive calls

Although the overall classification success (25%) of the DFA was higher than expected by chance (7.7%), most MD sequences were not correctly classified to the respective female (Table S3). The overall classification success resulted from a few females that had an average score of 50%, whereas in most females the classification success was 0%. Therefore, MDs do not seem to encode sufficient interindividual variation to allow for individual discrimination.

Pup-directed vocalizations of adult males

In each monitored colony, both harem males and peripheral males produced complete ICs in response to pup ICs (Table 1, columns 5 and 6). Male IC production was usually restricted to a single IC, only in a few cases males produced several successive ICs. Male IC production was observed when pups were between 10 and 30 days old (observed during 5 weeks, at least once up to three times per week in the same colony). In most cases (78%), males produced ICs after a pup emitted ICs. Male IC production seemed not to be directed to a specific pup. In 11% of cases, males produced ICs after a pup uttered a short vocal practice bout (i.e. multisyllabic vocal sequence) which contained mainly IC end syllables. In the remaining 11% of cases, male IC production could not be related to any preceding pup vocalization, but was sometimes followed by pup ICs or vocal practice sequences. During IC production males and pups never engaged in any behavioral activity with one another.

Table 2. Male isolation call production.

Colonies	Harem males (HM)	Peripheral males (PM)	IC from HM	IC from PM	Females	Pups	Harems
1: TO	1	0	Y	na	3 [2]	3	1
2: TR	3	2	Y (3)	Y (2)	8 [7]	7 [6]	3
3: B	1	2	Y	Y (1)	5	4	2
4: CVV	3	2	Y (1)	Y (2)	9	9	3
5: INH	2	1	Y (2)	Y (1)	6	3	2
6: PH	3	2	Y (2)	Y (2)	11	8	3
7: LH	2	1	Y (1)	Y (1)	6	5	3

Table 2: Representation of the IC production and the social group composition of each monitored colony. Numbers in parentheses depict how many males of the respective colony residents were observed to produce ICs (column 5 and 6). Numbers in square parentheses depict changes in colony social group composition during the field seasons as females (and their respective pups) sometimes disappeared. Colonies 1-4 belong to the Costa Rican population, colonies 5-7 belong to the Panamanian population.

Acoustic similarity between males and pups

Pup ICs had a higher acoustic similarity to the ICs of males that belonged to their colony than to ICs of males from other colonies (paired Wilcoxon Test: V=105, p=0.0001, effect size: r=0.881, Fig. 2B). For all 14 pups, the Euclidean distances to the male from the same colony was smaller than to the mean value for the males from the other colonies (Table S1, column 5).

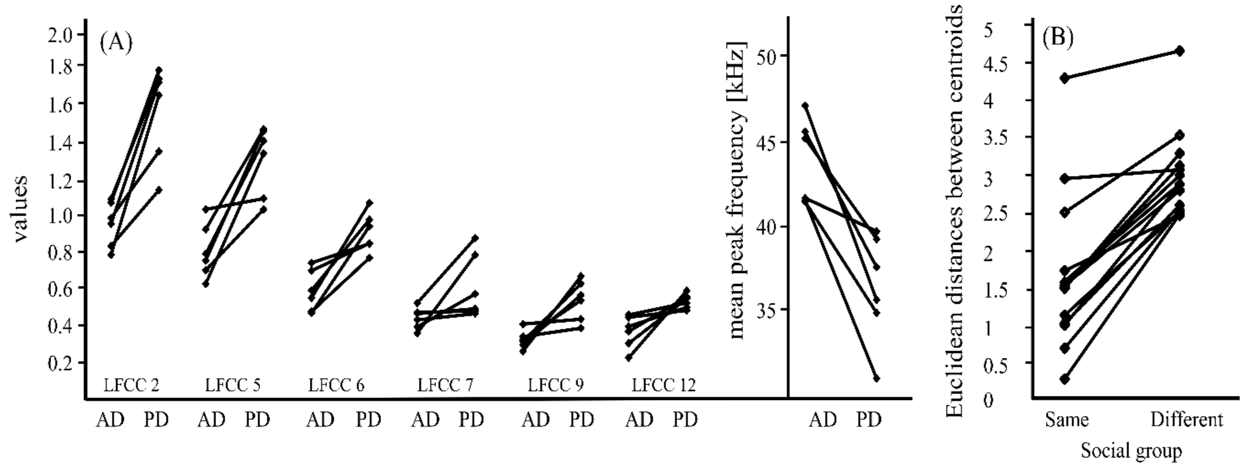


Fig. 2. Main results for pup-directed vocalizations of females and males.

(A) Pup-directed (PD) and adult-directed (AD) vocalizations of six females differ significantly for six LFCCs (linear frequency cepstral coefficients; LFCC 2, 5, 6, 7, 9, and 12); this suggests that pup-directed vocalizations have different acoustic properties (e.g. timbre) than adult-directed vocalizations of the same individual. Means for each call category (PD, AD) and female are shown. The differences in peak frequency of PD and AD vocalizations (mean: PD vocalizations: 36.5 kHz, AD vocalizations: 44 kHz) are shown next to the LFCC results. (B): The Euclidean distance, a proxy for acoustic similarity, between isolation calls of pups and males from their social group is smaller than the distance between isolation calls of pups and males from a different social group; this indicates that isolation calls from males and pups of the same social group share a group signature. The data includes calls from 8 males and 14 pups.

DISCUSSION

We detected pronounced acoustic differences between pup-directed and adult-directed female vocalizations which were consistent for all tested females. While the peak frequency of pup-directed calls was lower than that of adult-directed vocalizations (36.5 kHz vs. 44 kHz), values for all six LFCCs increased. Thus, our data indicates that both the pitch and the timbre of female vocalizations differed between adult-directed and pup directed calls. Pup-directed and adult directed calls are different vocalization types, so differences in peak frequency are not surprising. However, the large and consistent differences in LFCCs, which encode information on both pitch and timbre (De Poli & Prandoni, 1997; Piazza, Iordan, & Lew-Williams, 2017), suggest that the sound of the females' voice changed depending whether they were addressing their pups or adult conspecifics. This is similar to findings from human mothers which, irrespective of language, consistently shifted the timbre between adult-directed speech and motherese (Piazza, Iordan, & Lew-Williams, 2017). Our study describes for the first time a phenomenon that could be interpreted as reminiscent to motherese in bats. However, since our data set is very small, further investigations are needed before any final conclusions can be drawn.

In humans, motherese facilitates language learning (Kuhl et al., 1997) and its prosodic salience draws the infants attention towards the linguistic input (Grieser & Kuhl, 1988).
Despite the seemingly effortless language acquisition by infants, language learning is a complex task. Infants must learn the phonetic repertoire; they have to learn which speech subunits mark word boundaries (i.e. meaningful units) and which syllabic compositions occur in their native language. Motherese supports language learning by exaggerating lexical and grammatical structures (e.g. exaggeration of formant frequencies crucial for vowel discrimination) (Kuhl et al., 1997; Thiessen, Hill, & Saffran, 2005). Furthermore, motherese not only enhances the infants' attention towards the speech input but also promotes turn taking (Fernald, 1985; Fernald & Kuhl, 1987; Grieser & Kuhl, 1988). Therefore, it is suggested that motherese might also function as positive feedback for the vocalizing child, promoting further speech production (Fernald, 1985; Grieser & Kuhl, 1988).

The function of female MDs in our focal bat is not yet fully conclusive. The onset of MD call production coincides with increased pup independence, increased vocal practice behavior and increased behavioral activity (e.g. short flights within the day-roost). The production of MDs was observed in two contexts, during mother-pup reunions and during vocal practice bouts of the pup. Contrary to our expectation, we did not detect an individual signature in MDs, suggesting that they do not support mother-pup reunions as is the case in other bats (e.g. (Balcombe & McCracken, 1992; Brown, 1976; Esser & Schmidt, 1989). In *S. bilineata*, mothers are able to discriminate between own and alien pups based on ICs (Knörnschild & von Helversen, 2008) and females do not react aggressively towards alien pups, even when pups persistently and unsuccessfully solicit for maternal care from an alien female (personal observation A.A.F). Hence, pups may not need to discriminate between females because unidirectional recognition is sufficient. As aforementioned, MDs were also observed during vocal practice bouts of pups. We suggest that MDs may serve as a general positive feedback to pups during vocal practice but we need further investigations to validate this suggestion.

Furthermore, we describe a pup-directed adult male vocalization (adult IC) which seems strongly related to IC production in pups. Pups` ICs were acoustically more similar to ICs of males of the own social group than to ICs of males from other groups (Fig. 2B, Table S4). In most cases, pup IC bouts triggered the IC production of adult males. Considering these findings, we hypothesize that ICs of adult males may serve as guidance for the group signature in pup ICs. Thus, call convergence towards an already existing IC group signature could occur which would render vocal group signatures stable over time. The function of the group signature in ICs of S. *bilineata* remains to be investigated experimentally. Observations suggest that it may be of use when ICs of adult males are directed towards other adults. Two scenarios have been observed so far: 1) During agonistic interactions, submissive male produced ICs after which the dominant male ceased to be aggressive (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012). 2) Philopatric harem males produce ICs when courting newly immigrated females for the first time (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012). These observations suggest that adult-directed ICs are used for appeasement and to signal natal group affiliation; in both cases, a group signature would be beneficial. Thus, our new finding that adult males may influence the group signature of pups' ICs by producing ICs themselves is intriguing but we need further investigations with a considerably higher sample size (i.e. calls per individual and individual males per colony) to be certain. So far, other studies investigating the influence of adult vocal input on the formation of group signatures in juvenile vocalizations are restricted to songbirds (for review see Boughman & Moss, 2003) and two parrot species (Berg, Delgado, Cortopassi, Beissinger, & Bradbury, 2011; Farabaugh, Linzenbold, & Dooling, 1994).

To conclude, our study indicates that parent-offspring communication in bats is more complex than anticipated. Female pup-directed vocalizations seem to be reminiscent of human motherese, an interesting phenomenon that warrants further detailed studies. Moreover, male pup-directed vocalizations may facilitate the transmission of a vocal signature across generations, thus adding a new aspect to the study of social influences on vocal development.

AUTHOR CONTRIBUTIONS

AAF and MK designed the study. AAF collected the data and conducted the acoustic analyses. AAF preformed data analyses. AAF and MK wrote the manuscript.

FUNDING

This study as financed by a stipend from the Elsa-Neumann Foundation to AA. Fernandez and a Heisenberg Fellowship (DFG KN935 3-1) and a research grant (DFG KN935 4-1)
from the German Research Foundation to M. Knörnschild.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGEMENTS

We thank the Smithsonian Tropical Research Institute, the field station BCI, and the wildlife nature reserve Curú for providing excellent research conditions.

REFERENCES

Balcombe, J. P., & McCracken, G. F. (1992). Vocal recognition in Mexican free-tailed bats: do pups recognize mothers? *Animal Behaviour, 43*(1), 79-87.

Beecher, M. D., Stoddard, P. K., & Loesche, P. (1985). Recognition of parents' voices by young cliff swallows. *The Auk, 102*(3), 600-605.

Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R., & Bradbury, J. W. (2011). Vertical transmission of learned signatures in a wild parrot. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 585-591.

Biben, M., Symmes, D., & Bernhards, D. (1989). Contour variables in vocal communication between squirrel monkey mothers and infants. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 22*(6), 617-631.

Boughman, J. W., & Moss, C. F. (2003). Social sounds: vocal learning and development of mammal and bird calls. In *Acoustic communication* (pp. 138-224): Springer.

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development, 16*(1), 31-43.

Brown, P. (1976). Vocal communication in the pallid bat, *Antrozous pallidus*. *Zeitschrift für Tierpsychologie, 41*(1), 34-54.

Charrier, I., Mathevon, N., & Jouventin, P. (2001). Mother's voice recognition by seal pups. *Nature, 412*(6850), 873.

De Poli, G., & Prandoni, P. (1997). Sonological models for timbre characterization. *Journal of New Music Research, 26*(2), 170-197.

Esser, K. H., & Schmidt, U. (1989). Mother-Infant Communication in the Lesser Spear-Nosed Bat *Phyllostomus-Discolor* (Chiroptera, Phyllostomidae) - Evidence for Acoustic Learning. *Ethology, 82*(2), 156-168.

Esser, K. H., & Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat Phyllostomus discolor (Chiroptera, Phyllostomidae)—evidence for acoustic learning. *Ethology, 82*(2), 156-168.

Farabaugh, S. M., Linzenbold, A., & Dooling, R. J. (1994). Vocal plasticity in Budgerigars (*Melopsittacus undulatus*): evidence for social factors in the learning of contact calls. *Journal of Comparative Psychology, 108*(1), 81.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*(2), 181-195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279-293.

Fernandez, A. A., & Knörnschild, M. (2017). Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages. *Animal Behavior and Cognition, 4*(2), 169-186.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*(1), 14.

Jain, A., & Sharma, O. (2013). A Vector Quantization Approach for Voice Recognition Using Mel Frequency Cepstral Coefficient (MFCC): A Review 1.

Jouventin, P., Aubin, T., & Lengagne, T. (1999). Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Animal Behaviour, 57*(6), 1175-1183.

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*(9), 451-454. doi:10.1007/s00114-006-0127-9.

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. doi:10.1098/rsbl.2009.0685.

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*(4), 761-769. doi:10.1016/j.anbehav.2012.06.029.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., . . . Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684-686. doi:10.1126/science.277.5326.684.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology, 27*(20), 3162-3167. e3163.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science, 349*(6249), 734-738.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*(1), 53-71.

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, Scents, and Senses: Sexual Selection in the Greater Sac-Winged Bat, *Saccopteryx bilineata*. *Journal of Mammalogy, 89*(6), 1401-1410. doi:10.1644/08-mamm-s-060.1.

Whitham, J. C., Gerald, M. S., & Maestripieri, D. (2007). Intended receivers and functional significance of grunt and girney vocalizations in free-ranging female rhesus Macaques. *Ethology, 113*(9), 862-874.

Supplementary Information for

Pup directed vocalizations of adult females and males in a vocal learning bat

Authors: Ahana Aurora Fernandez and Mirjam Knörnschild

Table of Contents:

Results

Acoustic differences between pup-directed and adult-directed female vocalizations

GLM 1: The first multivariate GLM included all measured acoustic features, namely LFCCs 1-20, peak frequency, minimum and maximum frequency as dependent variables, as well as call type, ID and ID*call type as independent variables.

Table S1. Statistics of the multivariate GLM 1.

Independent variables	Wilks-Lambda	F-value (df)	η^2	Significance
ID	0.004	1.483 (120, 88.4)	0.66	0.026
call type	0.87	7.466 (24, 17)	0.91	> 0.001
ID*call type	0.005	1.469 (120, 88.4)	0.66	0.029

GLM2: The second GLM included the dependent variables which had no overlapping estimated marginal means for ID and call type (no interaction between ID and call type), i.e. only variables in which the acoustic changes between call types were consistent across all individuals.

Table S2. Statistics of the multivariate GLM 2.

Independent variable	Wilks-Lambda	F-value (df)	η^2	Significance
ID	0.424	0.939 (35, 145.5)	0.15	0.57
call type	0.333	9.732 (7, 34)	0.66	> 0.001
ID*call type	0.316	1.308 (35, 145.5)	0.20	0.13

Additional Paired Wilcoxon tests: Each dependent variable differed significantly between call types (same statistics for each variable: $V=21$, $p = 0.031$). Since our sample size was restricted to six individuals the p-value cannot drop below 0.031.

No individual signature in MDs

Although the overall classification success (25%) of the DFA was higher than expected by chance (7.7%), most MD sequences were not correctly classified to the respective female. The obtained classification success for each female was thus not significantly different from chance success (paired Wilcoxon test; Z = -1.732, p = 0.083).

Table S3. Individual classification success of MDs by a DFA.

ID	Colony	Classification success	No of females in social group
F2_2	SL1	42.9%	2 (1)
F4_4	SL2	50%	5 (4)
F7_7	E3	30%	9 (8)
F8_8	E3	0%	9 (8)
F9_9	F1	50%	8
F10_10	F1	0%	8
F11_11	B	0%	4
F12_12	B	20%	4
F14_14	TO	66.7%	3(2)
F15_15	TO	0%	3(2)
F17_17	TR	0%	2
F19_19	TR	0%	3
F20_20	TR	20%	3

Table S3: Individual classification success of the maternal directive call trains is shown in column 3 (in %). The last column delineates the number of females present in the colony, numbers in brackets indicate a change in number of present females since some females disappeared over the course of the study.

Methods

Study species

Adult bats were individually identified by colored plastic rings on their forearms (A.C. Hughes Ltd. UK, size XCL; one band per forearm). Pups were first identified via their banded mothers, since mothers only nurse their own pups and behave aggressively towards alien pups. At an age of approximately 60 days, when pups had almost adult mass and size, they were caught with mist nets (Ultrathin Mist Nets M-14; Ecotone, Gdynia, Poland) in the hvicinity of the day-roosts, sexed and banded with individually distinct colored bands. Immediately after the procedure pups were released at the site of capture. The banding method is well established and did not cause any noticeable negative effect on the bats' behavior or welfare. It was not possible to catch all the pups because we wanted to reduce the disturbance at the day-roosts to a minimum and thus did not perform re-capture attempts. Therefore, some pups remained unsexed and unbanded (but, nevertheless, individual identification was possible via their banded mothers).

Sound recordings

During the first two weeks after pup birth, sound recordings were restricted to the main vocal activity period of *S. bilineata*, one hour at dawn and dusk. Pups attached to the mothers' belly exclusively produced several isolation calls (i.e. isolation call bouts) when being awake. After the first two weeks when pups were more active and detached from their mothers, acoustic recording sessions were extended to entire mornings (between 6am – 1 pm) and afternoons (1pm - 6pm) where focal pups were recorded *ad libitum*. For each colony, we alternated between morning and afternoon sessions to avoid potential daytime biases in the vocal activity of pups.

Acoustic Analyses

Individual signature in maternal directive calls

Prior to acoustic measurements each sound file was bandpass filtered (5-120kHz) and normalized to 100% (Cool Edit 2000 Inc., Syntrillium Software Corporation P. O. Box 62255, Phoenix, AZ, USA). We used the software Avisoft SASLab Pro (v.5.2.09; R. Specht, Avisoft Bioacoustics, Glienicke, Germany) to measure acoustic parameters of the maternal directive calls (MDs) to test for an individual signature. Start and end of MDs within the train were determined automatically (-20dB relative to maximum amplitude, hold time 15ms). Automatic call detection was possible since the recordings had an excellent signal-to-noise ration and MDs are monosyllabic. Subsequent to automatic call detection, acoustic parameters were measured automatically. Spectrograms were created using a Hamming window with 1024-point fast Fourier transform and 93.75% overlap (resulting in a 488 Hz frequency resolution and a time resolution of 0.128 ms). We restricted the acoustic parameter measurements to the fundamental frequency because it contained most of the sound energy. To reduce the temporal dependence of successive MDs within a sequence we subsequently averaged the acoustic parameter measurements over the entire sequence. For each MD we measured several temporal (duration, interval between syllables, time to maximum amplitude), and spectral parameters (minimum, maximum and peak frequency, bandwidth, entropy, harmonics-to-noise ratio) over the entire syllable. To estimate the frequency and entropy curvature of MDs we measured six spectral parameters (minimum, maximum and peak frequency, bandwidth, entropy, harmonic-to-noise ratio) at eleven different locations distributed evenly over the fundamental frequency of each syllable. Curvature parameters were obtained by performing principal component analyses (PCAs) with varimax rotation on the above mentioned parameters (one PCA on all 44 frequency parameters and another PCA on all 22 entropy and harmonic-to-noise ratio parameters), thus reducing multicollinearity between original acoustic parameters considerably. For the frequency curvature, we extracted six principal components (with eigenvalues > 1) which explained 95.5% of the total variance and for the entropy curvature five principal components (with eigenvalues > 1) which explained 71.1% of the total variance. Both PCAs fulfilled KMO and Bartlett’s test criteria, thus assuring the appropriateness of our data for PCAs. We used the five most important derived acoustic parameters (three parameters describing the frequency curvature and two parameters describing the entropy curvature) per MD for subsequent statistical analyses.

Isolation calls

Prior to acoustic measurements, all sound files were bandpass filtered (5-120 kHz) and normalized to 100% (Cool Edit 2000 Inc., Syntrillium Software Corporation P. O. Box 62255, Phoenix, AZ, USA). For the acoustic parameter measurements we used the software Avisoft SasLab Pro. The start and end of the syllables were defined manually, based on the oscillogram, and subsequently, automatic parameter measurements provided by the software were applied. Spectrograms were created using a Hamming window with 1024-point fast Fourier transform and 87.5% overlap resulting in a frequency resolution of 488 Hz and a temporal resolution of 0.256 ms. Although isolation calls are multiharmonic, we used only the fundamental frequency for automated parameter measurements because it contained most of the sound energy. The multisyllabic isolation call is composed of simple tonal calls (variable simple syllable: VS) succeeded by the end syllables (ES). The endsyllable is composed of two syllable types, the composite (cs) and the short stereotyped syllable (ss). The composite syllable is constituted of a facultative noisy part (nc) followed by a tonal (tc) part. For each syllable type or syllable part we measured several temporal (duration, interval between syllables, time to maximum amplitude) and spectral (minimum, maximum and peak frequency, bandwidth, entropy, harmonics-to-noise ratio) parameters. Entropy is a measure of the width and uniformity of the power spectrum and assesses how tonal or noisy a signal is (on a scale of 0-1, pure tone=0, white noise=1). We decided to not include the measurements of the noisy part of composite syllables in statistical analyses because this part was not always present in all end syllables. To estimate the entropy and frequency curvatures of end syllables, we measured the spectral parameters at seven different locations evenly distributed over the fundamental frequency of each syllable type and part. We obtained the curvature parameters by performing a principal component analyses (PCAs) with varimax rotation on the aforementioned spectral parameters (one PCA on all 28 frequency parameters and another on all 14 entropy and harmonics to noise ratio parameters). For the frequency curvature we extracted four principal components (with eigenvalues >1) which explained 94.8% of the total variance for tc and five principal components (with eigenvalues >1) which explained 94.7% of the total variance for ss. For the entropy and harmonic-to-noise ratio (hnr) curvature we extracted four principal components (with eigenvalues > 1) which explained 73.5% of the total variance for tc and three principal components (with eigenvalues >1) which explained 69.5% of the total variance of ss. All PCAs fulfilled Kaiser-Meyer-Olkin (KMO) and Bartlett´s criteria, thus assuring the appropriateness of our dataset for PCAs. Hence, we obtained 16 derived acoustic parameters per isolation call for subsequent statistical analyses.

Assessment of acoustic similarity between ICs of adult males and pups

To obtain the centroids of each individual (i.e. the position of each individual in acoustic space) we performed a discriminant function analysis (DFA). We adjusted the DFA to the unequal number of analyzed calls per individual by computing group sizes based on prior probabilities. The number of variables introduced into a DFA should be lower than the smallest number of cases per group to avoid statistical overfitting. In our case, this was not possible since we sometimes only had 4 cases per group (i.e. 4 IC per individual). We selected 12 acoustic parameters (LFCC1, frequency curvatures 1-4 of the tc part, frequency curvatures 1-2 of the ss syllable, the entropy-hnr curvatures 1-2 of the ss syllable, the entropy-hnr curvature 1 of the tc part of the composite syllable, mean maximal frequency of the tc part of the composite syllable and the ss syllable) to obtain the centroids for a subsequent assessment of acoustic similarity. All parameters were checked for multicollinearity and introduced simultaneously into the DFA. We obtained a multidimensional space defined by four discriminant functions in which group centroids marked the position of each individual in the DFA signal space. Subsequently, we calculated Euclidean distances between centroids of pups and males from the same or different social groups.

Chapter IV was published in the following journal: Animal Behavior and Cognition, 2017, volume 4, issue 2, pages 169-186 and can be found under the following link:

https://doi.org/10.12966/abc.04.05.2017

Isolation Calls of the Bat *Saccopteryx bilineata* Encode Multiple Messages

Ahana Aurora Fernandez[1,*] and Mirjam Knörnschild[1,2]

[1] Free University Berlin, Berlin, Germany
[2] Smithsonian Tropical Research Institute, Balboa, Panama
* Corresponding author (Email: fernandez.aurora.ahana@gmail.com)

Citation – Fernandez, A. A., & Knörnschild, M. (2017). Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages. *Animal Behavior and Cognition, 4*(2), 169–186. doi: 10.12966/abc.04.05.2017

Abstract - The information that can possibly be encoded in a given vocalization is limited by the available acoustic space. Vocalizations composed of several elements have the potential to distribute information among distinct elements and thus encode various layers of information simultaneously. Correspondingly, the multiple messages hypothesis states that different elements of a vocalization convey different information, for example, about identity, signallers' quality or external events, which is directed to one or several receivers. The isolation call of the greater sac-winged bat, *Saccopteryx bilineata*, is a complex vocalization composed of different syllable types (variable, composite and stereotyped syllables). Pups produce isolation calls to elicit maternal care. In contrast to most bats, isolation calls of *S. bilineata* remain part of the adult vocal repertoire. In this study, we investigated the acoustic trajectories of isolation calls during ontogeny. Our analyses demonstrated that ontogenetic changes of acoustic parameters were most pronounced in stereotyped syllables. Throughout ontogeny, stereotyped syllables became shorter in duration and higher in frequency and bandwidth, thus effectively encoding information about age. Additionally, former studies revealed that composite syllables encode both an individual and a group signature. In our study, we also found an individual signature, thus corroborating the previous study. We conclude that isolation calls of *S. bilineata* encode multiple messages which are temporally segregated. The composite syllables conveyed information about individual identity and social group affiliation, whereas the stereotyped syllables encoded information about age. The temporal segregation allows the separation of information conveyed by distinct elements, thus enhancing signal reliability.

Keywords – Multiple messages hypothesis, Age signature, Individual signature, Temporal segregation, Vocal communication, Bats

Communicative signals are used to transfer information which elicits a behavioral response in the receiver (Bradbury & Vehrencamp, 2011; Simmons, 2003). Animals' signals are often highly complex, incorporating several sensory modalities (Bradbury & Vehrencamp, 2011; Hebets & Papaj, 2004). Complex signals may arise as a result of different selective pressures acting on signals. Selection pressures may affect the content of a signal (referred to as content-based selection) or the reception and correct decoding of the message by the receiver (referred to as efficacy-based selection; Guilford & Dawkins, 1991; Hebets & Papaj, 2004). Signal complexity can be achieved by producing signals which are composed of multiple messages (Candolin, 2003; Hebets & Papaj, 2004) or repetitive signal parts providing the same information (redundant signals), which is used to increase attention, reception and accurate receiver response (Hebets & Papaj, 2004). Signals can be transmitted and received via different sensory modalities, allowing communication through visual, olfactory, tactile and/or acoustic signals (Bradbury & Vehrencamp, 2011). Acoustic signals can encode different types of information such as signallers' quality in terms of size or social rank, individual, group or species identity, sex or age, and, sometimes, different messages are directed to different receivers (e.g., Blumstein & Armitage, 1997; Charlton, Zhihe, & Snyder, 2009; Erb, Hodges, & Hammerschmidt, 2013; Ey, Hammerschmidt, Seyfarth, & Fischer, 2007; Fischer, Kitchen, Seyfarth, & Cheney, 2004; Pfefferle & Fischer, 2006; Riesch, Ford, & Thomsen, 2006; Townsend, Hollén, & Manser, 2010).

According to the multiple message hypothesis (*sensu* Hebets & Papaj, 2004), signals encoding multiple layers of information can be composed of a single element, which, in turn, is composed of distinct acoustic units (Fischer et al., 2004; Furrer & Manser, 2009; Jansen, Cant, & Manser, 2012) or of different elements emitted in a sequence (Jahelková, Horáček, & Bartonička, 2008; Koren & Geffen, 2009; Nelson & Poesel, 2007; Ryan, 1983; Templeton, Greene, & Davis, 2005). Therefore, information can be encoded in distinct acoustic units of a single element, in different elements, or in both.

When information is encoded in different elements or acoustic units of a single element, temporal segregation of information is achieved (Jahelková et al., 2008; Nelson & Poesel, 2007; Templeton et al., 2005). Temporal segregation is beneficial because it deals with problems arising from multiple message encoding. Elements have a defined acoustic "space" which can only encode a limited size of information (in bit; e.g., Beecher, 1989). Consequently, when several messages are conveyed within the same element, reliability of one or several messages diminishes. However, when multiple messages are encoded in different elements (or acoustic units of a single element) of a vocalization, each acoustic "space" can be fully exploited to encode a single message, thus enhancing signal reliability (Jahelková et al., 2008; Koren & Geffen, 2009; Nelson & Poesel, 2007). Furthermore, temporal segregation of information may allow a more precise decoding process of the receiver, therefore facilitating the behavioral response (Nelson & Poesel, 2007; Templeton et al., 2005).

Temporal segregation of information can be achieved by segmentation of a vocalization into different acoustic elements. In the black-capped chickadee, temporal segregation of a multisyllabic call is required to encode one of several messages (Templeton et al., 2005). The entire call conveys information about a predation event (e.g., stationary predator), whereas the number of "dee" elements specifies the size of the predator. A playback experiment revealed that chickadees respond accordingly to the conveyed information about predator size (Templeton et al., 2005). Temporal segregation can also be obtained by segmentation of a single element into distinct acoustic units, as for example in the contact call of banded mongooses, *Mungos mungo* (Jansen et al., 2012). This call is composed of two distinct acoustic units, a noisy and a harmonic part (Jansen et al., 2012). The noisy part is always present and encodes individual and group identity (Müller & Manser, 2008), whereas the different behavioral contexts are encoded in the duration of the harmonic part (Jansen et al., 2012).

Bats are a highly promising taxon for studying the temporal segregation of information in social vocalizations because many elements in their vocal repertoire consist of distinct acoustic units (i.e., a noisy and a tonal part) and vocalizations containing several different elements have been described for many species (e.g., Behr & von Helversen, 2004; Chaverri, Gillam, & Kunz, 2012; Chaverri, Gillam, & Vonhof, 2010; Davidson & Wilkinson, 2002; Leippert, 1994; Leippert, Goymann, Hofer, Marimuthu, & Balasingh, 2000; Pfalzer & Kusch, 2003). However, investigations on whether distinct elements in vocalizations convey different messages have only been conducted in few bat species (e.g., Bohn, Schmidt-French, Schwartz, Smotherman, & Pollak, 2009; Jahelková et al., 2008).

Saccopteryx bilineata is a Neotropical insectivorous bat species living in year-round stable colonies which can consist of up to 60 individuals (Bradbury & Vehrencamp, 1976; Tannenbaum, 1975). Colonies contain smaller social groups called harems, which are composed of a harem male guarding a territory of up to 2m^2 vertical surface area where females can choose to roost during the day (Bradbury & Vehrencamp, 1976). Additionally, nonharem males which are queuing for harem access in their natal colony are also often present (Nagy, Knörnschild, Voigt, & Mayer, 2012; Voigt & Streich, 2003). Males are not able to control female movements within the colony, which leads to intense courtship efforts throughout the whole year (reviewed in Nagy & Knörnschild, 2016; Voigt et al., 2008). Females disperse after weaning (at 12 – 14 weeks of age) to avoid inbreeding with older male relatives (Nagy, Heckel, Voigt, & Mayer, 2007). Pup births are highly synchronized, with females giving birth to one pup per year in May (Bradbury & Vehrencamp, 1976; Tannenbaum, 1975). New-born pups are already fully furred and have their eyes open (Tannenbaum, 1975). Throughout ontogeny, *S. bilineata* pups of both sexes engage in a conspicuous vocal behavior termed babbling (Knörnschild, Behr, & von Helversen, 2006). Babbling pups emit long vocal sequences containing different syllable types (the term "syllable" is used synonymously to "element") of the adult vocal repertoire as well as interspersed isolation calls and echolocation pulses (Knörnschild et al., 2006). Babbling behavior probably allows pups to acquire the highly diverse adult vocal repertoire (Behr & von Helversen, 2004) by repeatedly practicing the production of different syllable types and whole vocalization types (Knörnschild, 2014; Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). Pups utter isolation calls to elicit maternal care (Knörnschild & von Helversen, 2008). In contrast to isolation calls of other bat species, isolation calls of *S. bilineata* are comparatively long in duration (> 1 s), and composed of three different syllable types, namely variable syllables (vs), and composite (cs) and stereotyped (ss) end syllables (Knörnschild & von Helversen, 2008). Isolation calls encode an individual and a group signature located in the composite end syllables (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2012; Knörnschild & von Helversen, 2008). Female *S. bilineata* are capable of differentiating own versus alien pups based on isolation calls alone (Knörnschild & von Helversen, 2008), which is in line with other bat species (e.g., Balcombe, 1990; Bohn, Wilkinson, & Moss, 2007; Fanis & Jones, 1995; Knörnschild, Feifel, & Kalko, 2013). However, unlike isolation calls in other bats (e.g., Engler, Rose, & Knörnschild, 2017; Esser & Schmidt, 1989; Knörnschild, von Helversen, & Mayer, 2007; Scherrer & Wilkinson, 1993; Sterbing, 2002), isolation calls of *S. bilineata* remain part of the adult vocal repertoire. They are not only emitted during mother-pup interactions and in pups' babbling trains but also during agonistic male-male interactions and male courtship (Knörnschild et al., 2012).

The aim of this study was to investigate the ontogenetic change of pup isolation calls in *S. bilineata*. We hypothesized that the majority of acoustic parameters would not change considerably during ontogeny since isolation calls remain in the adult vocal repertoire. Nevertheless, we expected to find fine-scale acoustic differences which can be used as a reliable indicator for assessing pup age. Therefore, we hypothesized that isolation calls encode different layers of information — i.e., individual identity, group affiliation and age — in different syllable types (or in different acoustic units of syllables), according to the multiple message hypothesis.

Method

Study Site and Animals

We conducted sound recordings at Barro Colorado Island, a field station of the Smithsonian Tropical Research Institute. Barro Colorado Island is located in the artificial Gatun lake (9° 9' 0''N, 79° 51' 0''W) in Panamá. We recorded vocalizations of 14 bat pups from six colonies throughout their ontogeny (from mid-May to end of August 2015, see Table 1). Colonies were roosting on the outside walls of buildings belonging to the field station and contained up to four social groups (i.e., harems) each. Each social group consisted of the harem male, several females with their respective pups and, in two colonies, nonharem males in the periphery (see Table 1 for details on group and colony composition). Since these colonies are part of a long-term project, our focal bats were well habituated to human presence and not noticeably disturbed by behavioral observations or sound recordings. Adult bats were individually identified by colored plastic bands on their forearms (A.C. Hughes Ltd. UK; size XCL; one band per forearm). All young pups were individually identified via their banded mothers, since mothers only nurse their own pups and behave aggressively towards alien pups. At approximately 60 days of age, when they had almost adult mass and size, pups were caught with mist nets (Ultrathin Mist Nets M-14; Ecotone, Gdynia, Poland) in the vicinity of their day-roosts, sexed and equipped with individually distinct colored bands. Immediately after the procedure, pups were released at the site of capture. The banding method is well established and did not cause any noticeable negative effect on the bats' behavior or welfare. It was not possible to catch all pups because we did not want to disturb the bats by repeated capture attempts; therefore, five of 14 pups remained unsexed and unbanded (but, nevertheless, individual identification was possible via their banded mothers).

Table 1

Colony and Social Group Composition of Pups

Pup ID	Colony	Colony composition				Social group composition		Pup sex
		Harem males	*Non-harem males*	*Females*	*Pups*	*Females*	*Pups*	
1	C1	1	-	4	4 (3)	4	4 (3)	Unknown
2	E1	1	2	2	2	2	2	Unknown
3	E1	1	2	2	2	2	2	Male
4	E3	3	1	9 (8)	9 (8)	3	3	Female
5	E3	3	1	9 (8)	9 (8)	2	2	Male
6	F1	4	-	8	7	3	3	Unknown
7	F1	4	-	8	7	3	3	Female
8	SL1	1	-	2 (1)	2 (1)	2 (1)	2 (1)	Male
9	SL2	2	-	5 (4)	5 (3)	4 (3)	5 (3)	Unknown
10	SL2	2	-	5 (4)	5 (3)	4 (3)	5 (3)	Unknown
11	E3	3	1	9 (8)	9 (8)	3	3	Male
12	E3	3	1	9 (8)	9 (8)	3	3	Male
13	E3	3	1	9 (8)	9 (8)	2	2	Female
14	SL2	1	-	5 (4)	5 (3)	1	1	Female

Note. Representation of the colony and social group composition for each pup (*N* = 14). Numbers in parentheses depict changes in colony or social group composition during the field season as females and their respective pups sometime switched to a different social group or disappeared.

Sound Recordings

Sound recordings and accompanying behavioral observations of pups in the day-roost were performed during their whole ontogeny. During the first two weeks after pups were born, we restricted recordings and observations to the main activity period of bats, one hour after dawn and before dusk, respectively. With increasing age pups started to be more active throughout the entire day. Consequently, we then conducted sound recordings and behavioral observations throughout the entire day (*ad libitum* sampling *sensu* Altmann, 1974). We recorded each colony twice a week, once from dawn to midday and once from midday to dusk. We used high-quality ultrasonic recording equipment (500 kHz sampling rate, 16-bit depth resolution) to record vocalizations of individual pups. The set-up consisted of a microphone (Avisoft UltraSoundGate 116Hm, with condenser microphone CM16, frequency range 1-200 kHz ± 3 dB, connected to a laptop computer Lenovo S21e) running the software Avisoft-Recorder (v4.2.05 R. Specht, Avisoft Bioacoustics, Glienicke, Germany). Due to the excellent visibility in the day-roosts and the short recording distance (2-4 m on average), we could reliably associate isolation calls to individual pups based on mouth movements and the accompanying behavior (i.e., approaching the mother to solicit maternal care).

Acoustic Analysis

In this study, the terms "element" and "syllable" are used synonymously and are defined as an acoustic entity surrounded by silence. Syllables were classified into different categories ("syllable types") according to their acoustic properties. Furthermore, syllables were either compositional (i.e., composed of two or more acoustic units) or simple (i.e., composed of only one acoustic unit). We defined acoustic units as separately identifiable parts with distinct acoustic features (e.g., a noisy part followed by a purely tonal part). Isolation calls of *S. bilineata* pups are multisyllabic and consist of three different syllable types, namely variable (vs), composite (cs) and stereotyped (ss) syllables (Knörnschild & von Helversen, 2008, Figure 1). Composite syllables have a noisy part ('nc') and a tonal part ('tc'), but the noisy part may be reduced or absent in some cases. Composite syllables ('nc+tc') and stereotyped syllables ('ss') are combined into so-called 'end syllables' because isolation calls always end with them. In total, we measured 356 isolation calls from 14 different pups belonging to six different social groups. To minimize temporal dependence between calls, we never measured more than two isolation calls per pup and day.

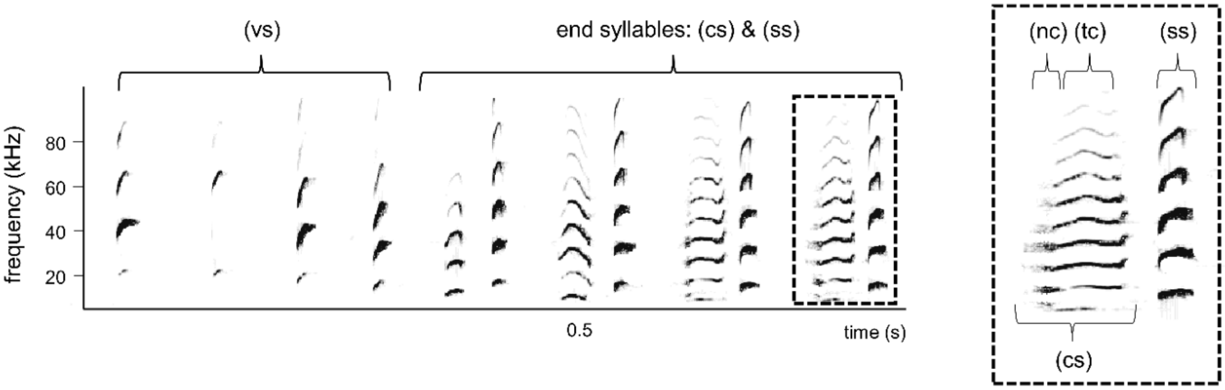


Figure 1. Entire isolation call of one pup composed of three different syllable types. **(vs)** variable simple syllables. **(tc)** tonal part of the composite syllable. **(nc)** noisy part of the composite syllable. **(cs)** composite syllable (noisy part connected to tonal part). **(ss)** stereotyped syllables. The spectrogram depicts frequency (kHz) as a function of time (s) and was generated using a 1042 point fast Fourier transform and a Hamming window with 50% overlap.

Prior to acoustic measurements, isolation calls were bandpass filtered (5-90 kHz) and normalized to 100% (Cool Edit 2000 Inc., Syntrillium Software Corporation P.O. Box 62255, Phoenix, AZ, USA). We used the software Avisoft SASLab Pro (v.5.2.09; R. Specht, Avisoft Bioacoustics, Glienicke, Germany) for our acoustic analyses. Start and end of syllables and syllable parts were determined manually based on the oscillogram and, subsequently, automatic parameter measurements provided by the software were applied. Spectrograms for measuring single syllable types were created using a Hamming window with 1024-point fast Fourier transform and 87.5% overlap (frequency resolution: 488 Hz; time resolution: 0.256 ms). Spectrograms for measuring entire call parameters were created using a Hamming window with 1024-point fast Fourier transform and 50% overlap (frequency resolution: 488 Hz; time resolution: 1.024 ms).

Some acoustic measurements were taken over entire isolation calls, while others focused on the end syllables of isolation calls because earlier work found that two vocal signatures (individual identity and social group affiliation) are encoded there (Knörnschild et al., 2012; Knörnschild & von Helversen, 2008). Isolation calls are multiharmonic and we incorporated all harmonics when measuring entire isolation calls. When measuring end syllables, however, we used only the fundamental frequency (first harmonic) because it contained most of the sound energy. For entire isolation calls, we measured duration, mean peak, minimum and maximum frequency, mean bandwidth and mean entropy. Entropy is a measure of the width and uniformity of the power spectrum and assesses how tonal or noisy a signal is (on a scale of 0-1; pure tone = 0, white noise = 1). For end syllables, different syllable types or syllable parts were measured separately (nc, tc, ss). We did not use measurements of the noisy part of composite syllables (nc) in statistical analyses because nc was not present in all end syllables (but we report all acoustic measurements, see Appendix A, Tables A1-A4). For each isolation call, we measured three end syllables and subsequently averaged measurements per syllable type and syllable part. We measured the duration, time to maximum amplitude and interval of each syllable type and syllable part. To estimate the frequency and entropy curvature of end syllables, we measured five spectral parameters (peak frequency, minimum and maximum frequency, bandwidth, entropy) at five different locations distributed equally over the fundamental frequency of each syllable type and syllable part. Curvature parameters were obtained by performing principal component analyses (PCAs) with varimax rotation on the above mentioned parameters (one PCA on all 20 frequency parameters and another PCA on all five entropy parameters), thus reducing multicollinearity between original acoustic parameters considerably. For the frequency curvature, we extracted three principal components (with eigenvalues >1) which explained 93.4% of the total variance for tc and two principal components (with eigenvalues >1) which explained 94.6% of the total variance for ss. For the entropy curvature, we extracted one principal component for tc and ss each (with an eigenvalue >1) which explained 56.2% (tc) and 63.4% (ss) of the total variance, respectively. Both PCAs fulfilled Kaiser-Meyer-Olkin (KMO) and Bartlett's test criteria, thus assuring the appropriateness of our data for PCAs. Thus, we obtained seven derived acoustic parameters (five parameters describing the frequency curvature and two parameters describing the entropy curvature for tc and ss) per isolation call for subsequent statistical analyses.

Statistical Analysis

All statistical tests were conducted in SPSS (v.20; IBM SPSS Statistics Chicago, IL, USA) and R (v.3.02; R Development Core Team 2008). To estimate the ontogenetic trajectory of isolation call parameters, we performed separate General Linear Mixed Models (GLMMs) for each call parameter (with age in 10-day steps as covariate, sex as fixed factor and pup ID as random factor; Gamma distribution with log link function). In addition, we performed separate linear regressions for each pup and call parameter to obtain the respective individual slopes of the linear regressions. Slopes were used as a basic proxy for visualizing ontogenetic development (i.e., a positive slope value indicated that a certain parameter increased during ontogeny). GLMMs and linear regressions were conducted for 19 acoustic parameters, namely tc duration, tc time to maximum amplitude, tc mean peak frequency, tc mean bandwidth, tc frequency curvature 1-3, tc entropy curvature 1, ss duration, ss time to maximum amplitude, ss mean peak frequency, ss mean bandwidth, ss frequency curvature 1-2, ss entropy curvature 1, entire call duration, entire mean peak frequency, entire mean bandwidth, and entire mean entropy.

To test for the existence of an individual signature in isolation calls, we performed a discriminant function analysis (DFA) on the whole data set. To test whether the strength of the individual signature remained stable or changed during ontogeny, we performed two more DFAs on a subset of the data (only including pups for which we had a sufficient amount of recorded isolation calls to split the data into two ontogenetic phases). We adjusted the DFAs to the unequal number of analyzed calls per pup by computing group sizes based on prior probabilities (calls per pup varied within each DFA but not between DFAs). We used a cross-validation procedure to estimate the correct classification success ('n-1' cross-validation procedure), which classified each call based on discriminant functions established with all calls except the call being classified. Subsequently, we performed a Binomial test to check whether the obtained classification success was better than a random classification.

In the first DFA (testing the existence of an individual signature), we used 14 pups for which we analyzed at least 17 isolation calls each (17-38 isolation calls per pup, 356 calls in total). We selected 16 acoustic parameters, namely tc duration, tc interval, tc frequency curvature 1-3, tc entropy curvature 1, ss duration, ss interval, ss frequency curvature 1-2, ss entropy curvature 1, entire call duration, entire mean peak frequency, entire mean bandwidth, entire mean entropy, and number of variable syllables of entire call. All parameters were checked for multicollinearity and included simultaneously into the DFA.

In two subsequent DFAs (testing the strength of the individual signature during ontogeny), we used ten pups for which we analyzed at least nine isolation calls for each of two ontogenetic phases (9-15 isolation calls per ontogeny phase and pup, 220 calls in total). The number of calls per pup was equal for both ontogenetic phases to ensure a direct comparison of the classification success obtained by both DFAs. We split our data set into two ontogenetic phases based on pup age, resulting in one data set with pups younger than 34 days and a second data set with pups older than 34 days. This age corresponded to the onset of independent foraging which inaugurates a phase of greater independence of pups, even though they are still nursed by their mothers (Knörnschild et al., 2012). We selected eight acoustic parameters, namely tc duration, tc frequency curvature 1-2, tc entropy curvature 1, ss duration, ss frequency curvature 1-2, and ss entropy curvature 1. Again, all parameters were checked for multicollinearity and included simultaneously into the DFA. The classification success of individual pups was compared between both ontogenetic phases using a Wilcoxon test. All statistical tests were conducted in SPSS and R.

Results

Ontogenetic Trajectories of Isolation Calls

We conducted separate General Linear Mixed Models (GLMMs) to assess the ontogenetic trajectory of isolation call parameters. Taken together, ontogenetic changes were more pronounced in the stereotyped syllables than in the composite syllables' tonal part (for details see Table 2). As pups matured, the mean peak frequency of the tonal component of composite syllables showed a strong trend to decrease, $t = -1.93$; $p = 0.053$. In contrast to this, the mean peak frequency and bandwidth of the stereotyped syllables increased significantly, mean peak frequency: $t = 4.94$, $p < 0.001$; bandwidth: $t = 6.51$, $p < 0.001$, (Figure 2). This also resulted in a significant increase of the entire isolation calls' peak frequency, $t = 3.47$, $p < 0.001$ because the stereotyped syllables had a higher relative amplitude than the composite syllables. The duration of entire isolation calls increased with age as well, $t = 8.72$, $p < 0.001$. Most other significant ontogenetic changes concerned the stereotyped syllables. For example, duration and the time to maximum amplitude decreased significantly during ontogeny, whereas entropy significantly increased, duration: $t = -3.37$, $p < 0.001$; time to maximum amplitude: $t = -3.07$, $p = 0.002$; entropy: $t = 2.70$, $p = 0.006$). Pup sex had a negligible effect in all of the conducted GLMMs (see Appendix B, Table B1).

Individual Signature

Most isolation calls could be correctly classified to the respective pup. A DFA with 356 isolation calls of 14 pups classified 48.9 % of all calls to the correct individual, which was significantly higher than expected by chance (7.1%; binomial test: $p = 0.0009$, $N = 14$). Differences between isolation calls of different pups were subtle (Figure 3), but a previous study showed that these differences are sufficient for correct maternal offspring recognition (Knörnschild & von Helversen, 2008). Misclassifications were significantly more likely to occur between pups from the same social group (intra-group misclassification) than between pups from different social groups (inter-group misclassifications; Wilcoxon test: Z = -2.667; $p = 0.008$, $N = 12$; see Table 3).

In both ontogenetic stages, pups could be statistically distinguished based on acoustical parameters of isolation calls (pups younger than 34 days: 46.4% classification success; pups older than 34 days: 56.4% classification success; $N = 10$ pups in each case). Again, the obtained classification success was significantly better than expected by chance (10%), binomial test: $p = 0.001$ for ontogeny phase 1, $p = 0.01$ for ontogeny phase 2. We found no significant difference between the classification success for individual pups from both ontogenetic phases, Wilcoxon test: Z = -1.174; $p = 0.24$, *ns* ($N = 10$; see Table 4), suggesting that the strength of the individual signature remains stable during ontogeny.

Table 2

Ontogenetic Development of Acoustic Parameters

Syllable type	**Dependent variable**	**Estimate**	**t**	***p***	**Trajectory**
Tonal part of CS	Duration (ms)	-0.010135	-1.160	0.247	n.s.
Tonal part of CS	Time to max. amplitude (ms)	0.014780	1.002	0.316	n.s.
Tonal part of CS	Mean peak freq. (kHz)	-0.008092	-1.930	0.053	n.s.
Tonal part of CS	Mean bandwidth (kHz)	-0.002916	-0.268	0.788	n.s.
Tonal part of CS	PC1 entropy	-0.013856	-2.142	0.032	decrease*
Tonal part of CS	PC1 call curv.	-0.021493	-2.636	0.008	decrease**
Tonal part of CS	PC2 call curv.	-0.001279	-0.191	0.848	n.s.
Tonal part of CS	PC3 call curv.	0.008929	1.300	0.194	n.s.
Stereotyped	Duration (ms)	-0.018117	-3.370	< 0.001	decrease***
Stereotyped	Time to max. amplitude (ms)	-0.040350	-3.072	0.002	decrease**
Stereotyped	Mean peak freq. (kHz)	0.020765	4.940	< 0.001	increase***
Stereotyped	Mean bandwidth (kHz)	0.072170	6.510	< 0.001	increase***
Stereotyped	PC1 entropy	0.021806	2.707	0.006	increase**
Stereotyped	PC1 call curv.	0.038622	4.701	< 0.001	increase***
Stereotyped	PC2 call curv.	0.034560	5.800	< 0.001	increase***
Entire call	Duration (s)	0.085696	8.729	< 0.001	increase***
Entire call	Mean peak freq. (kHz)	0.039714	3.470	< 0.001	increase***
Entire call	Mean bandwidth (kHz)	-0.007587	-1.610	0.108	n.s.
Entire call	Mean entropy	-0.009124	-2.470	0.013	decrease*
Entire call	Harmonics to noise ratio	0.003677	0.530	0.598	n.s.

Note. GLMMs with two fixed factors, age (in 10-day blocks) and sex (female, male, unknown), and one random factor, pup ID 1-14. Results for sex are not listed in the table (but can be found in the appendix) because they were non-significant in most of the cases. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

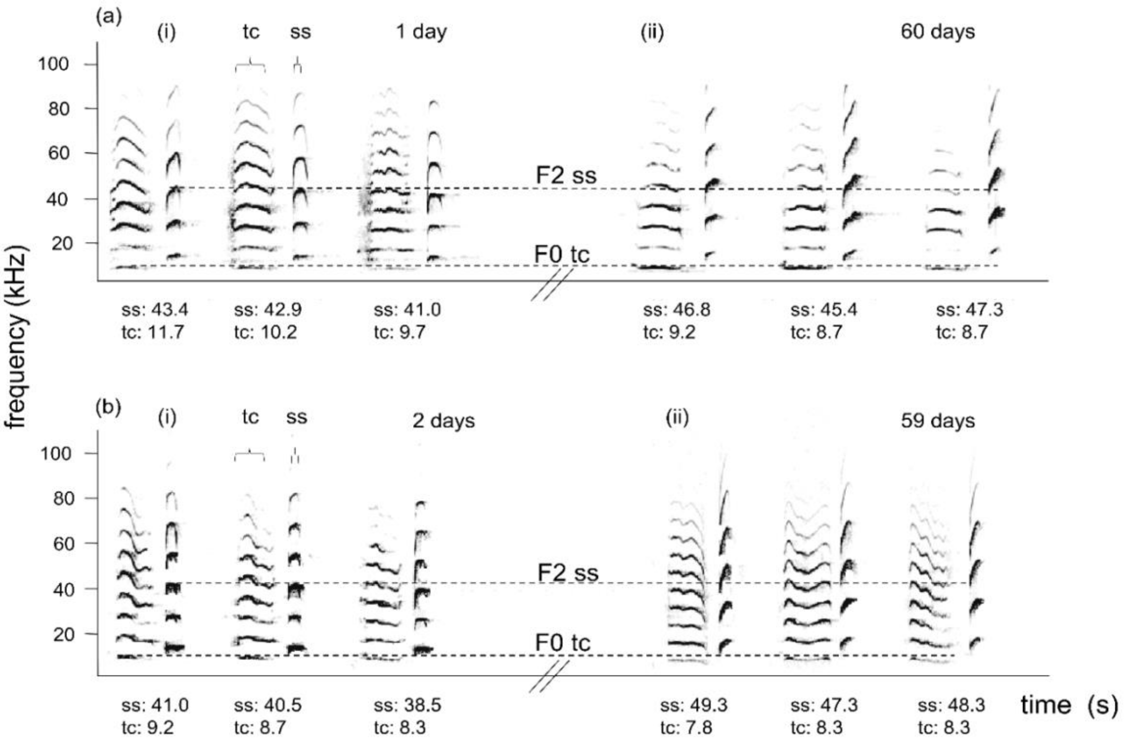


Figure 2. Isolation call end syllables of two pups at different ages. Three consecutive end syllables of one pup (unknown sex) **(a)** at one day (i) and 60 days of age (ii) and of a male pup **(b)** at two days (i) and 59 days of age (ii). Dashed lines serve as visual orientation for the ontogenetic change of acoustic parameters of tc and ss with age. F0 (fundamental frequency): depicts the decrease of mean peak frequency of the tonal part of the composite syllable. The rather subtle decrease in frequency of the tonal unit of the composite syllable with age is further illustrated by mean values of the peak frequency (kHz) of each tonal part (tc). F2 (second harmonic): depicts the increase of the mean peak frequency of the stereotyped syllable. Mean values of the peak frequency (kHz) of each stereotyped syllable (ss) are given as well. Spectrograms depict frequency (kHz) as a function of time (s) and were generated using a 1042 point fast Fourier transform and a Hamming window with 50% overlap.

Table 3

Intra- and Inter-group Misclassifications of Isolation Calls to Individual Pups (N = 14) Obtained by a DFA and the Classification Success Calculated Separately for Each Social Group (N = 4 groups)

Pup ID	Group ID	Intra-group misclassification	Inter-group misclassification	Classification success within social groups (%)
2	2	3.20	2.96	78.2% (50% random classification success)
3	2	16.70	4.86	
4	3	9.08	3.03	58.8% (20% random classification success)
5	3	4.78	3.43	
11	3	2.63	2.63	
12	3	10.83	3.40	
13	3	5.95	3.70	
6	4	16.70	5.55	70.8% (50% random classification success)
7	4	3.30	3.33	
9	6	16.70	0.75	73.2% (33.3% random classification success)
10	6	1.65	1.52	
14	6	14.70	6.42	

Note. Misclassifications were averaged per pup. Two of 14 pups are not listed here because we did not have any recordings for fellow pups from their social group.

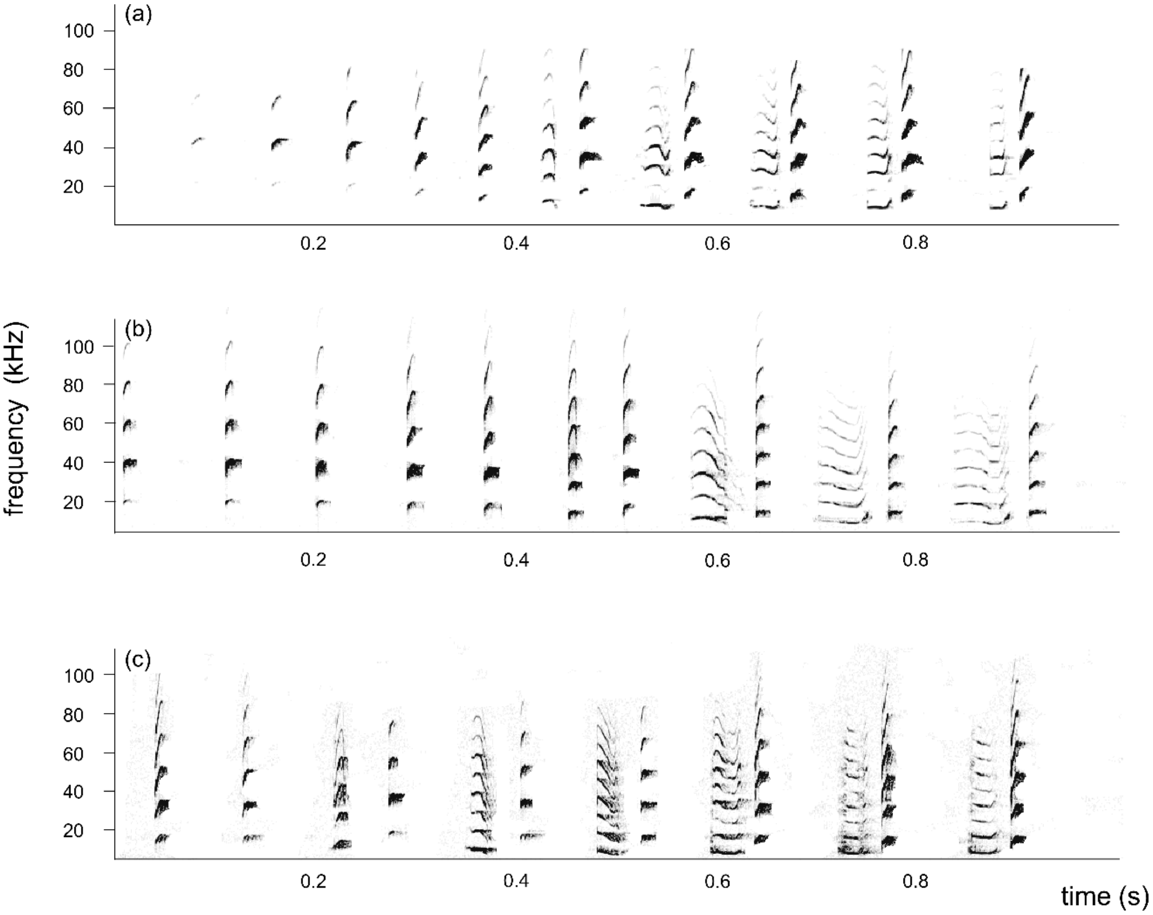


Figure 3. Spectrograms of isolation calls encoding an individual signature from three pups at similar ages. **(a)** female bat pup, 46 days of age. **(b)** unknown sex, 49 days of age. **(c)** male bat pup, 44 days of age. Pups belonged to three different colonies. Spectrograms depict frequency (kHz) as a function of time (s) and were generated using a 1042 point fast Fourier transform and a Hamming window with 50% overlap.

Table 4

Strength of Individual Signature During Ontogeny (N = 10 pups)

Pup ID	Classification success < 34 days of age (%)	Classification success > 34 days of age (%)	Relative change of classification success
2	80.0	66.7	decrease
3	30.0	10.0	decrease
4	11.1	44.4	increase
5	70.0	80.0	increase
7	46.7	80.0	increase
8	45.5	90.9	increase
9	75.0	66.7	decrease
11	20.0	30.0	increase
12	11.1	33.3	increase
13	44.4	33.3	decrease

Note. Four of 14 pups had to be excluded from the analysis because we did not have sufficient recordings throughout ontogeny for a comparison of signature strength.

Discussion

This study shows that multisyllabic isolation calls of *S. bilineata* pups encode multiple messages in different syllable types (multiple messages hypothesis *sensu* Gerhardt, 1992; Hebets & Papaj, 2004). While it was already established that the tonal part of the composite syllables in isolation calls encodes information about individual identity and group affiliation (Knörnschild et al., 2012; Knörnschild & von Helversen, 2008), this study found conclusive evidence that the stereotyped syllables convey information about age. In contrast to the stereotyped syllables, the acoustic parameters of composite syllables remained generally stable during ontogeny.

The ontogenetic change in stereotyped syllables is most probably a by-product of vocal tract maturation and not influenced by the vocal learning processes shaping the group signature encoded in the tonal part of the composite syllables (Knörnschild et al., 2012). The ontogenetic increase in body mass and the accompanying growth of the vocal folds does not contradict the frequency increase in the stereotyped syllables. Fundamental frequencies are not only influenced by the length and mass of the vocal folds but also by their tension (Bradbury & Vehrencamp, 2011; Fitch & Hauser, 2003). Tension of vocal folds is achieved through laryngeal muscle control and muscular control becomes more sophisticated during ontogeny (Bradbury & Vehrencamp, 2011). Until now, the vocal tract of *S. bilineata* has not been investigated. Nevertheless, the most parsimonious assumption is that the acoustic changes in stereotyped syllables result from maturation processes, suggesting that they are an honest signal about the signaller (Bradbury & Vehrencamp, 2011; Fitch & Hauser, 2003).

As a receiver, information about signaller age could be useful in different social interactions such as during agonistic encounters, for mate choice, or during predator encounters where signaller reliability is crucial for a receiver's response (Blumstein & Daniel, 2004; Blumstein & Munos, 2005; Charlton et al., 2009; Erb et al., 2013; Ey et al., 2007; Fischer et al., 2004; Mathevon, Koralek, Weldele, Glickman, & Theunissen, 2010; Reby & McComb, 2003). In *S. bilineata*, pup age might be useful for females to adequately allocate maternal care. However, it is unlikely that females would rely on a single signal for assessing how much maternal care should be invested. Most probably, other signals such as pup behavior (solicitation for food), foraging efficacy, and morphological cues, such as size, influence the allocation of maternal care as well. To test the mothers` ability to discriminate between pup ages based on stereotyped syllables alone, one could perform a playback experiment in which mothers had to distinguish between correct versions of their own pups` isolation calls and incorrect versions in which the stereotyped syllables have been artificially shifted upwards or downwards in frequency. If mothers evaluate age information in pup isolation calls, they should react most strongly to the correct versions.

The multiple messages conveyed in *S. bilineata's* isolation calls (individual ID, group affiliation, age) are encoded in different syllable types and thus temporally segregated. This segregation has the advantage that information can be encoded without a reduction in signal reliability or receivers' decoding accuracy (Hebets & Papaj, 2004; Jahelková et al., 2008; Koren & Geffen, 2009; Nelson & Poesel, 2007; Ryan, 1983). A former study (Knörnschild & von Helversen, 2008) showed that the individual signature, encoded in the tonal part of composite syllables, facilitated maternal offspring recognition based on isolation calls alone. The strengths of the individual signature in both the latter and our present study are comparable and surprisingly low for a signal under strong natural selection pressure (allonursing does not occur in *S. bilineata*; Knörnschild & von Helversen, 2008). The comparatively low strength of the individual signature may be caused by the fact that a second signature is encoded in the tonal part of composite syllables as well, suggesting that the individual signature is, to a certain degree, masked by the group signature (Knörnschild et al., 2012). This signature, conveying information about social group affiliation, increases in strength during ontogeny, while the individual signature strength seems to remain stable as pups mature (Knörnschild et al., 2012, and this study). Since information about age is temporally segregated from information about individual identity and group affiliation (age information is encoded in a different syllable type, the stereotyped syllables), it might be easier for conspecifics to decode the information content of pup isolation calls.

Our findings correspond to other incidents of temporal segregation of information in various taxa. In white-crowned sparrows, for example, males produce a multisyllabic song encoding multiple messages, namely identity and local dialect, which receivers perceive accordingly (Nelson & Poesel, 2007). In the rock hyrax (*Procavia capensis*), males sing complex multisyllabic songs which are presumably used for mate attraction and rival deterrence (Koren & Geffen, 2009). Hyrax song encodes information about body weight, size, body condition, hormonal levels and social status in different song elements (Koren & Geffen, 2009). The above mentioned studies and others (e.g., Bohn et al., 2009; Jahelková et al., 2008; Müller & Manser, 2008; Templeton et al., 2005) demonstrate that the temporal segregation of information occurs in vocalizations under both natural selection and sexual selection pressures.

To conclude, our study shows that multisyllabic isolation calls of *S. bilineata* pups convey different layers of information encoded in two different syllable types. This finding highlights the possibility that more animal vocalizations than previously anticipated encode multiple messages. Future work will benefit especially from focussing on multisyllabic vocalizations which are emitted in different social contexts and/or directed towards different receivers.

Acknowledgements

We thank the Smithsonian Tropical Research Institute and the field station BCI for providing excellent research conditions. Ethics: The process of acquiring data and protocols for capturing and handling bats complied with the current laws of Panama and the Smithsonian Tropical Research Institute (IACUC 2013-1015-2016). This study was financed by a stipend from the Elsa-Neumann Foundation to A. Fernandez and both a Heisenberg Fellowship (DFG KN935 3-1) and a research grant (DFG KN935 4-1) from the German Research Foundation to M. Knörnschild.

References

Altmann, J. (1974). Observational study of behaviour: Sampling methods. *Behaviour, 49*, 227–266.

Balcombe, J. P. (1990). Vocal recognition of pups by mother Mexican free-tailed bats, *Tadarida-Brasiliensis-Mexicana*. *Animal Behaviour, 39*, 960–966. doi:10.1016/S0003-3472(05)80961-3

Beecher, M. D. (1989). Signaling systems for individual recognition - an information-theory approach. *Animal Behaviour, 38*, 248–261. doi:10.1016/S0003-3472(89)80087-9

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

Blumstein, D. T., & Armitage, K. B. (1997). Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Animal Behaviour, 53*, 143–171. doi:10.1006/anbe.1996.0285

Blumstein, D. T., & Daniel, J. C. (2004). Yellow-bellied marmots discriminate between the alarm calls of individuals and are more responsive to calls from juveniles. *Animal Behaviour, 68*, 1257–1265. doi: 10.1016/j.anbehav.2003.12.024

Blumstein, D. T., & Munos, O. (2005). Individual, age and sex-specific information is contained in yellow-bellied marmot alarm calls. *Animal Behaviour, 69*, 353–361. doi:10.1016/j.anbehav.2004.10.001

Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and stereotypy of free-tailed bat songs. *Plos One, 4,* e6746*.* doi.org/10.1371/journal.pone.0006746

Bohn, K. M., Wilkinson, G. S., & Moss, C. F. (2007). Discrimination of infant isolation calls by female greater spear-nosed bats, *Phyllostomus hastatus*. *Animal Behaviour, 73*, 423–432. doi:10.1016/j.anbehav.2006.09.003

Bradbury, J. W., & Vehrencamp, S. L. (1976). Social-organization and foraging in Emballonurid bats. 2. Model for determination of group-size. *Behavioral Ecology and Sociobiology, 1*, 383–404. doi:10.1007/Bf00299400

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication* (2nd ed.). Sunderland, MA: Sinauer Associates Inc.

Candolin, U. (2003). The use of multiple cues in mate choice. *Biological Reviews, 78*, 575-595. doi:10.1017/s14647 93103006158

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009). The information content of giant panda, *Ailuropoda melanoleuca*, bleats: Acoustic cues to sex, age and size. *Animal Behaviour, 78*, 893─898. doi:10.1016/j.anbehav.2009.06.029

Chaverri, G., Gillam, E. H., & Kunz, T. H. (2012). A call-and-response system facilitates group cohesion among disc-winged bats. *Behavioral Ecology, 24*, 481─487. doi:10.1093/beheco/ars188

Chaverri, G., Gillam, E. H., & Vonhof, M. J. (2010). Social calls used by a leaf-roosting bat to signal location. *Biology Letters, 6*, 441─444. doi:10.1098/rsbl.2009.0964

Davidson, S. M., & Wilkinson, G. S. (2002). Geographic and individual variation in vocalizations by male *Saccopteryx bilineata* (Chiroptera: Emballonuridae). *Journal of Mammalogy, 83*, 526─535. doi:10.1644/1545-1542(2002)083<0526:Gaiviv>2.0.Co;2

Engler, S., Rose, A., & Knörnschild, M. (2017). Isolation call ontogeny in bat pups (*Glossophaga soricina*). *Behaviour, 154,* 267─286. doi:10.1163/1568539x-00003421

Erb, W. M., Hodges, J. K., & Hammerschmidt, K. (2013). Individual, contextual, and age-related acoustic variation in Simakobu (*Simias concolor*) loud calls. *Plos One, 8*, e83131. doi:10.1371/journal.pone.0083131

Esser, K. H., & Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae) - Evidence for acoustic learning. *Ethology, 82*, 156─168.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., & Fischer, J. (2007). Age- and sex-related variations in clear calls of *Papio ursinus*. *International Journal of Primatology, 28*, 947─960. doi:10.1007/s10764-007-9139-3

Fanis, E., & Jones, G. (1995). Post-natal growth, mother-infant interactions and development of vocalizations in the vespertilionid bat *Plecotus auritus*. *Journal of Zoology, 235*, 85─97. doi:10.1111/j.1469-7998.1995.tb05130.x

Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: Acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology, 56*, 140─148. doi:10.1007/s00265-003-0739-4

Fitch, W. T., & Hauser, M. D. (2003). Unpacking "honesty": Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (pp. 65 ─ 137). New York, NY: Springer.

Furrer, R. D., & Manser, M. B. (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *American Naturalist, 173*, 400─410. doi:10.1086/596541

Gerhardt, H. C. (1992). Multiple messages in acoustic signals. In P. Marler (Ed.), *Seminars in neuroscience,* (Vol 4, No 6, pp. 391 ─ 400). Columbia, MO: Academic Press Ltd.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour, 42*, 1─14. doi:10.1016/s0003-3472(05)80600-1

Hebets, E. A., & Papaj, D. R. (2004). Complex signal function: Developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology, 57*, 197─214. doi:10.1007/s00265-004-0865-7

Jahelková, H., Horáček, I., & Bartonička, T. (2008). The advertisement song of *Pipistrellus nathusii* (Chiroptera, Vespertilionidae): A complex message containing acoustic signatures of individuals. *Acta Chiropterologica, 10*, 103─126. doi:10.3161/150811008x331144

Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biology, 10*, 97. doi:10.1186/1741-7007-10-97

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

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*, 451─454. doi:10.1007/s00114-006-0127-9

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

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters, 6*, 156─159. doi:10.1098/rsbl.2009.0685

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. doi:10.1016/j.anbehav.2012.06.029

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*, 1001─1009. doi:10.1016/j.anbehav.2008.05.018

Knörnschild, M., von Helversen, O., & Mayer, F. (2007). Twin siblings sound alike: Isolation call variation in the noctule bat, *Nyctalus noctula*. *Animal Behaviour, 74*, 1055─1063. doi:10.1016/j.anbehav.2006.12.024

Koren, L., & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): A multi-information distributing channel. *Behavioral Ecology and Sociobiology, 63*, 581-590. doi:10.1007/s00265-008-0693-2

Leippert, D. (1994). Social-behavior on the wing in the false vampire, *Megaderma-Lyra*. *Ethology, 98*, 111─127.

Leippert, D., Goymann, W., Hofer, H., Marimuthu, G., & Balasingh, J. (2000). Roost-mate communication in adult Indian false vampire bats (*Megaderma lyra*): An indication of individuality in temporal and spectral pattern. *Animal Cognition, 3*, 99─106.

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

Müller, C. A., & Manser, M. B. (2008). Mutual recognition of pups and providers in the cooperatively breeding banded mongoose. *Animal Behaviour, 75*, 1683─1692. doi:10.1016/j.anbehav.2007.10.021

Nagy, M., Heckel, G., Voigt, C. C., & Mayer, F. (2007). Female-biased dispersal and patrilocal kin groups in a mammal with resource-defence polygyny. *Proceedings of the Royal Society of London B: Biological Sciences, 274*, 3019─3025. doi:10.1098/rspb.2007.1008

Nagy, M. & Knörnschild, M. (2016). Sex-biased dispersal and social systems of Neotropical Emballonurids. In J. Ortega (Ed.), *Sociality in bats* (pp. 47 ─ 63). Zurich, Switzerland: Springer International Publishing.

Nagy, M., Knörnschild, M., Voigt, C. C., & Mayer, F. (2012). Male greater sac-winged bats gain direct fitness benefits when roosting in multimale colonies. *Behavioral Ecology, 23*, 597─606. doi:10.1093/beheco/ars003

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*, 1073─1084. doi:10.1016/j.anbehav.2007.01.018

Pfalzer, G., & Kusch, J. (2003). Structure and variability of bat social calls: Implications for specificity and individual recognition. *Journal of Zoology, 261*, 21─33. doi:10.1017/s0952836903003935

Pfefferle, D., & Fischer, J. (2006). Sounds and size: Identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*. *Animal Behaviour, 72*, 43─51. doi:10.1016/j.anbehav.2005.08.021

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour, 65*, 519─530. doi:10.1006/anbe.2003.2078

Riesch, R., Ford, J. K. B., & Thomsen, F. (2006). Stability and group specificity of stereotyped whistles in resident killer whales, *Orcinus orca*, off British Columbia. *Animal Behaviour, 71*, 79─91. doi:10.1016/j.anbehav.2005.03.026

Ryan, M. J. (1983). Frequency modulated calls and species recognition in a neotropical frog. *Journal of Comparative Physiology, 150*, 217─221.

Scherrer, J. A., & Wilkinson, G. S. (1993). Evening bat isolation calls provide evidence for heritable signatures. *Animal Behaviour, 46*, 847─860. doi:10.1006/anbe.1993.1270

Simmons, A. M. (2003). Perspectives and progress in animal acoustic communication. In A. M. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (pp. 1-14). New York, NY: Springer

Sterbing, S. J. (2002). Postnatal development of vocalizations and hearing in the phyllostomid bat, *Carollia perspicillata*. *Journal of Mammalogy, 83*, 516─525. doi:10.1644/1545-1542(2002)083

Tannenbaum, B. R. (1975). *Reproductive strategies in the white-lined bat.* (Doctoral dissertation). Cornell University, Cornell, New York. Retrieved from SIRIS database http://www.siris.si.edu/.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science, 308*, 1934─1937. doi:10.1126/science.1108841

Townsend, S. W., Hollén, L. I., & Manser, M. B. (2010). Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. *Animal Behaviour, 80*, 133─138. doi:10.1016/j.anbehav.2010.04.010

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., & Nagy, M. (2008). Songs, scents, and senses: Sexual selection in the greater sac-winged bat, *Saccopteryx bilineata*. *Journal of Mammalogy, 89*, 1401─1410. doi:10.1644/08-mamm-s-060.1

Voigt, C. C., & Streich, W. J. (2003). Queuing for harem access in colonies of the greater sac-winged bat. *Animal Behaviour*, *65*, 149–156. doi:10.1006/anbe.2002.2031

Appendix A

Table A1

Acoustic Parameters: Noisy Part of Composite Syllables

Pup ID	Duration (ms)	Time to max. amplitude (ms)	Mean peak freq. (kHz)	Mean bandwidth (kHz)	PC1 entropy	PC1 call curv.	PC2 call curv.	PC3 call curv.
1 (16)	2.99	2.33	27.6	5.6	-0.413	-0.103	-0.320	-0.033
2 (21)	2.54	2.01	33.7	4.9	-0.376	0.606	-0.344	-0.231
3 (14)	4.66	4.20	29.6	5.7	-0.031	0.187	-0.235	-0.131
4 (7)	3.43	2.81	26.1	4.0	-1.164	-0.218	-1.001	-0.215
5 (8)	3.45	2.57	22.9	6.5	-0.019	-0.679	0.144	-0.088
6 (14)	4.10	3.42	22.8	6.1	0.156	-0.595	0.187	-0.224
7 (24)	7.06	6.36	27.2	5.6	-0.192	-0.117	-0.410	0.109
8 (31)	9.01	7.37	29.5	7.0	0.610	0.085	0.672	-0.457
9 (5)	3.45	3.17	26.5	4.6	-1.018	-0.164	-0.690	-0.137
10 (15)	3.43	2.70	26.4	6.9	0.755	-0.267	0.477	-0.153
11 (27)	4.61	3.59	32.9	6.3	-0.045	0.361	0.237	0.733
12 (15)	14.38	13.11	28.5	7.4	0.660	-0.081	0.396	0.735
13 (7)	4.78	4.08	30.9	5.7	-0.120	0.293	-0.154	-0.268
14 (10)	5.45	3.73	24.9	4.2	-0.731	-0.396	-0.709	-0.178

Note. Raw data of the different parameters are presented as mean ($\bar{x}$) of all measurements per pup. Note that the noisy part was not present in every call and thus the mean presented here is an average of only those calls which had a noisy part. Call number per pup is given in parentheses after the pup ID.

Table A2

Acoustic Parameters: Tonal Part of Composite Syllables

Pup ID	Duration (ms)	Time to max. amplitude (ms)	Mean peak freq. (kHz)	Mean bandwidth (kHz)	PC1 entropy	PC1 call curv.	PC2 call curv.	PC3 call curv.
1 (19)	23.84	8.92	8.9	1.8	-0.305	-0.134	0.100	-0.097
2 (32)	22.59	9.1	8.8	1.8	0.331	-0.206	-0.211	0.061
3 (24)	23.38	8.66	8.8	1.8	-0.029	0.049	-0.347	-0.254
4 (22)	26.09	14.42	8.9	1.4	-0.828	0.044	0.001	-0.229
5 (26)	32.57	16.45	9.2	2.2	0.207	0.090	-0.015	0.341
6 (18)	24.82	9.91	9.2	1.9	-0.355	0.058	0.054	0.355
7 (30)	27.46	12.77	9.2	1.4	-0.645	-0.088	0.470	-0.237
8 (33)	26.33	18.13	10.6	3.3	1.111	-0.622	1.695	0.291
9 (24)	18.94	5.55	9.2	2.4	-0.065	0.411	-0.669	0.113
10 (30)	16.86	6.78	9.1	2.5	0.739	0.367	-0.825	0.079
11 (38)	17.55	8.05	8.6	1.3	-0.206	-0.178	-0.274	-0.315
12 (23)	20.45	7.47	10.3	2.2	0.264	0.844	0.082	0.107
13 (21)	25.83	14.28	8.4	1.6	-0.490	-0.312	-0.256	-0.147
14 (17)	18.93	9.43	8.9	1.7	-0.633	-0.014	-0.158	0.042

Note. Raw data of the different parameters are presented as mean ($\bar{x}$) of all measurements per pup. Note that the total call numbers differ between individual pups. Call number per pup is given in parentheses after the pup ID.

Table A3

Acoustic Parameters: Stereotyped Syllables

Pup ID	Duration (ms)	Time to max. amplitude (ms)	Mean peak freq. (kHz)	Mean bandwidth (kHz)	PC1 entropy	PC1 call curv.	PC2 call curv.
1 (19)	6.96	2.30	16.5	2.9	0.724	0.724	-0.195
2 (32)	7.22	2.73	15.2	2.2	0.771	0.771	-0.281
3 (24)	6.42	2.09	14.9	2.7	0.445	0.445	0.018
4 (22)	5.36	1.86	13.9	1.7	-0.351	-0.351	-0.253
5 (26)	6.61	2.29	16.0	1.7	-0.048	-0.048	-0.184
6 (18)	7.21	2.43	14.6	2.6	0.392	0.392	0.341
7 (30)	7.58	2.37	15.2	3.0	0.162	0.162	0.510
8 (33)	7.18	2.85	14.6	2.1	0.013	0.013	-0.186
9 (24)	6.98	2.19	12.3	2.8	-0.630	-0.630	0.518
10 (30)	6.87	2.41	14.5	3.6	0.133	0.133	0.704
11 (38)	5.48	2.21	12.9	1.6	-0.392	-0.392	-0.464
12 (23)	6.11	2.15	14.1	1.7	-0.434	-0.434	-0.473
13 (21)	5.66	2.52	14.6	2.3	-0.179	-0.179	0.171
14 (17)	6.73	2.36	13.34	1.9	-0.749	-0.749	-0.076

Note. Raw data of the different parameters are presented as mean ($\bar{x}$) of all measurements per pup. Note that the total call numbers differ between individual pups. Call number per pup is given in parentheses after the pup ID.

Table A4

Acoustic Parameters: Entire Call

Pup ID	Duration (s)	Mean peak freq. (kHz)	Mean bandwidth (kHz)	Mean entropy	Harmonics to noise ratio
1 (19)	0.91	31.8	52.9	0.362	21.61
2 (32)	1.08	28.1	52.7	0.353	20.54
3 (24)	1.54	26.6	56.1	0.385	18.25
4 (22)	1.50	32.3	50.7	0.342	21.95
5 (26)	1.50	31.2	49.9	0.313	22.69
6 (18)	1.67	28.0	52.6	0.345	23.56
7 (30)	1.20	31.6	51.5	0.349	20.54
8 (33)	1.01	30.1	53.6	0.369	21.18
9 (24)	1.41	26.2	43.7	0.317	23.16
10 (30)	1.74	30.5	44.4	0.330	22.62
11 (38)	1.41	30.2	48.8	0.334	20.26
12 (23)	1.29	25.9	49.7	0.351	18.98
13 (21)	1.14	28.2	47.5	0.347	19.40
14 (17)	1.15	31.2	46.9	0.325	20.84

Note. Raw data of the different parameters are presented as mean ($\bar{x}$) of all measurements per pup. Note that the total call numbers differ between individual pups. Call number per pup is given in parentheses after the pup ID.

Appendix B

Table B1

Ontogenetic Development of Acoustic Parameters in Female and Male Pups

Syllable type	**Dependent variable**	**Sex**	**Estimate**	**t**	***p***	**Trajectory**
Tonal part of CS	Duration (ms)	Female	0.127925	1.46	0.146	n.s.
Tonal part of CS	Duration (ms)	Male	0.107603	1.32	0.187	n.s.
Tonal part of CS	Distance to max. amplitude (ms)	Female	0.47125	3.166	0.001	Increase **
Tonal part of CS	Distance to max. amplitude (ms)	Male	0.31888	2.308	0.021	Increase *
Tonal part of CS	Mean peak freq (kHz)	Female	-0.018560	-0.51	0.607	n.s.
Tonal part of CS	Mean peak freq (kHz)	Male	0.051538	1.55	0.121	n.s.
Tonal part of CS	Mean bandwidth (kHz)	Female	-0.257784	-2.287	0.022	Decrease *
Tonal part of CS	Mean bandwidth (kHz)	Male	0.047766	0.456	0.648	n.s.
Tonal part of CS	PC1 Entropy	Female	-0.237559	-3.802	0.0001	Decrease ***
Tonal part of CS	PC1 Entropy	Male	0.091533	1.581	0.113	n.s.
Tonal part of CS	PC1 call curv.	Female	-0.065979	-0.916	0.359	n.s.
Tonal part of CS	PC1 call curv.	Male	-0.020092	-0.302	0.762	n.s.
Tonal part of CS	PC2 call curv.	Female	0.101359	1.363	0.173	n.s.
Tonal part of CS	PC2 call curv.	Male	0.150005	2.167	0.030	Increase *
Tonal part of CS	PC3 call curv.	Female	-0.056057	-1.064	0.287	n.s.
Tonal part of CS	PC3 call curv.	Male	-0.003933	-0.081	0.935	n.s.
Stereotyped	Duration (ms)	Female	-0.125893	-2.58	0.010	Decrease *
Stereotyped	Duration (ms)	Male	-0.096351	-2.13	0.033	Decrease *
Stereotyped	Distance to max. amplitude (ms)	Female	-0.076780	-0.96	0.336	n.s.
Stereotyped	Distance to max. amplitude (ms)	Male	-0.031500	-0.44	0.661	n.s.
Stereotyped	Mean peak freq (kHz)	Female	-0.022286	-0.51	0.613	n.s.
Stereotyped	Mean peak freq (kHz)	Male	-0.011453	-0.28	0.780	n.s.
Stereotyped	Mean bandwidth (kHz)	Female	-0.212450	-1.87	0.061	n.s.
Stereotyped	Mean bandwidth (kHz)	Male	-0.388830	-3.69	0.0002	Decrease ***
Stereotyped	PC1 entropy	Female	-0.176338	-2.24	0.024	Decrease *
Stereotyped	PC1 entropy	Male	-0.111754	-1.53	0.125	n.s.
Stereotyped	PC1 call curv.	Female	-0.011949	-0.14	0.892	n.s.
Stereotyped	PC1 call curv.	Male	-0.003267	-0.04	0.968	n.s.
Stereotyped	PC2 call curvature	Female	-0.013777	-0.23	0.819	n.s.
Stereotyped	PC2 call curv.	Male	-0.141397	-2.53	0.011	Decrease *
Entire call	Duration (s)	Female	-0.032539	-0.34	0.733	n.s.
Entire call	Duration (s)	Male	-0.011165	-0.13	0.900	n.s.
Entire call	Mean peak freq (kHz)	Female	0.087851	1.42	0.156	n.s.
Entire call	Mean peak freq (kHz)	Male	0.003581	0.06	0.948	n.s.
Entire call	Mean bandwidth (kHz)	Female	0.000422	0.01	0.991	n.s.
Entire call	Mean bandwidth (kHz)	Male	0.054037	1.59	0.112	n.s.
Entire call	Mean entropy	Female	-0.004142	-0.13	0.893	n.s.
Entire call	Mean entropy	Male	0.028272	1.00	0.319	n.s.
Entire call	Harmonics to noise ratio	Female	-0.072516	-1.52	0.129	n.s.
Entire call	Harmonics to noise ratio	Male	-0.095553	-2.20	0.028	Increase *

Note. GLMMs with two fixed factors, age (in 10-day blocks) and sex (female, male, unknown), and one random factor, pup ID 1-14. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Chapter V

Chapter V was published in Functional Ecology in 2019, volume 34, issue 2, pages 322-331 and can be found under the following link:

https://doi.org/10.1111/1365-2435.13407

General discussion

The key finding of my thesis is that infant canonical babbling and pup babbling are defined by the same characteristic features. Furthermore, I demonstrate that two main aspects of pup babbling, namely the amount of vocal practice and the babbling bout composition are influenced by social feedback from conspecifics. My thesis includes the first description of a phenomenon reminiscent of motherese in pup-directed female vocalizations. Additionally, I describe an adult male pup-directed vocalization: adult males produce isolation calls which may serve as template for the learned group signature encoded in pup isolation calls. In addition to the group signature, the pup isolation call encodes individual identity. In my thesis, I show that this call further encodes information about pup age and that the different information contents are temporally segregated according to the "multiple messages" hypothesis. Finally, my thesis demonstrates that increased social complexity (i.e. group size) is correlated with an increase in individual identity information encoded in social vocalizations as predicted by the "social complexity hypothesis for communication" hypothesis.

Babbling: an extraordinary vocal practice behavior during ontogeny

Babbling is a very scarce trait in the animal kingdom. A behavior reminiscent of canonical babbling in non-human animals is described in songbirds (for review see Doupe & Kuhl, 1999), in the pygmy marmoset (Elowson, Snowdon, & Lazaro-Perea, 1998) and the bat *Saccopteryx bilineata* (Chapter I, Knörnschild, Behr, & von Helversen, 2006). Up to now, a comparison of the characteristic babbling features across babbling species was missing. My in-depth study of pup babbling behavior provides the first qualitative comparison of babbling features across mammalian vocal production learners. Yet, for what reason is it important to study this pronounced vocal practice behavior and compare it across species?

It is of prime importance that babbling behavior is an indicator of learning taking place, both in human infants and *S. bilineata* pups. I believe that vocal ontogenetic processes offer the possibility to investigate which proximate mechanisms and ultimate causes underlie complex cognitive tasks such as vocal production learning. Therefore, a cross-species comparison of vocal ontogenetic processes is a promising avenue to investigate the evolution of complex communication systems such as the human language faculty. In the following, I will discuss a) what the functions of babbling are, b) why babbling can reflect how social feedback influences

learning processes, and c) an outlook how babbling can be used to investigate proximate mechanisms such as neurogenetics in vocal production learners.
During human speech acquisition, infants imitate the speech sound subunits to acquire the phonetic repertoire of their maternal language (Oller, 1980; Vihman, 2014). Similarly, bat pups acquire a part of their adult vocal repertoire through vocal production while babbling (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010, Chapter I), and juvenile male songbirds produce plastic song while learning the tutors' song (Marler, 1970). The commonalities in the three aforementioned examples are the ability of vocal production learning and the translation of the inner template into motor activity by practicing and refining the own vocal output. It is known – for songbirds and human infants – that practicing is indispensable for the development of song, respectively speech (Fagan, 2015; Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974; Konishi, 1985).

Infant canonical babbling allows the child to get refined control over the laryngeal muscles by repeatedly practicing the production of canonical syllables (i.e. syllables with adult-like resonance properties and mature temporal vowel-consonant transitions) (Oller, 1980; Oller, Griebel, & Warlaumont, 2016; Vihman, 2014). The production of protophones (i.e. non-speech sounds like raspberry vocalizations, squeals and growls), which can be understood as vocal play exploration, probably supports this goal (Oller, 1980; Oller, Griebel, & Warlaumont, 2016). Babbling in *S. bilineata* pups likely serves the same purpose, to refine the control over the vocal apparatus while rehearsing the adult-like syllable types (Chapter I). In songbirds, the production of the plastic song is likewise necessary for successful tutor song learning (Doupe & Kuhl, 1999; Marler, 1970). On the contrary, the pygmy marmoset is a non-vocal learner but, nevertheless, young infants go through a vocal practice phase reminiscent of infant canonical babbling (Elowson, Snowdon, & Lazaro-Perea, 1998). Unlike in the other babbling species, in marmoset infants, the primary function of babbling is to solicit care (Elowson, Snowdon, & Lazaro-Perea, 1998). Young babbling infants are more likely to be approached and interacted with compared to non-babbling ones (Elowson, Snowdon, & Lazaro-Perea, 1998).

This vocal practice behavior also influences social interactions in the other babbling species. In humans, studies investigating interactions between primary caregivers and babbling infants demonstrate that babbling tends to elicit immediate adult responses (e.g. Goldstein, King, & West, 2003) and shapes the parental feedback, e.g. elicits simplified linguistic responses (e.g. Albert, Schwade, & Goldstein, 2018; Elmlinger, Schwade, & Goldstein, 2019). A well-known form of parental feedback is infant-directed speech. Infant-directed speech, also called “motherese” is characterized by simple grammatical structures, increased pitch, exaggerated frequency contours and slow speech tempo (Fernald & Kuhl, 1987; Grieser & Kuhl, 1988). Motherese has linguistic and social benefits (Kuhl et al., 1997). Motherese supports linguistic learning by highlighting word boundaries, by intensifying frequency contours of phonetic units and by increased pitch, which probably attracts the infants´ attention towards the acoustic input (Grieser & Kuhl, 1988; Kuhl et al., 1997). The social benefit is that motherese promotes the infants´ attention and arousal as well as vocal turn-taking behavior by the infant (Grieser & Kuhl, 1988).

I also investigated specific pup-directed adult vocalizations occurring during babbling. To my knowledge, my study is first to describe a phenomenon that could be interpreted as reminiscent to motherese (Chapter III). I compared pup-directed versus adult-directed vocalizations of the same individuals and demonstrated that timbral differences exist between the two contexts (Chapter III). The study was inspired by a recent study in humans demonstrating that infant-directed is defined by a universal shift in timbre, which makes it distinguishable from adult-directed speech (Piazza, Iordan, & Lew-Williams, 2017). If and how the maternal directive call influences pup babbling remains to be investigated.

Social feedback includes not only vocal interactions but also behavioral responses (Goldstein & Schwade, 2010). Likewise, I found that maternal behaviors (i.e. non-vocal feedback) in *S. bilineata* influence main aspects of pup babbling (Chapter II). Furthermore, I demonstrated that the auditory input (i.e. the number of singing males) influences the syllabic composition of vocal practice. Noteworthy, maternal behaviors and the number of singing males influenced in particular the production of babbling trains composed of territorial song syllables, which are acquired through vocal production learning.

In humans, social interactions – both vocal and non-vocal – during infant speech acquisition are probably influential because they can increase the attention of the child towards spoken acoustic inputs and increase arousal. Both factors are crucial during learning and probably play a major role in memory development, thus influencing the quality of speech acquisition (Kuhl, 2003). This hypothesis is called the “social brain hypothesis” (Kuhl, 2003, 2007). In general, social feedback in humans shapes the infants vocal practice behavior; for example, contingent maternal responses were correlated with an increase in developmentally advanced consonant-vowel vocalizations (Gros-Louis, West, Goldstein, & King, 2006; Gros-Louis, West, & King, 2014), contingent maternal responses lead to an increase in the number and quality (i.e. more canonical syllables) of babbling (Goldstein, King, & West, 2003), and general volubility is enhanced by social interactions (Franklin et al., 2014). If the social feedback in *S. bilineata* likewise influences song learning success remains to be investigated. However, learning processes are shaped by the social environment and it can be assumed that social feedback and a pronounced vocal practice behavior positively influence and reinforce each other, both in the human infant and in *S. bilineata* pups.

As mentioned before, an interesting field for cross-species comparisons is the proximate mechanisms of vocal production learning, more specifically, neurobiology and genetics.

To study proximate mechanisms such as neurogenetics of vocal production learning two things are crucial: a) a suitable study species and b) a possibility to infer when learning takes place. Regarding point a), songbirds have so far been the preferred taxon to perform comparative studies (Doupe & Kuhl, 1999). Although neurogenetic studies in songbirds significantly advanced our knowledge about learning mechanisms, similar studies in mammalian vocal production learner would be crucial. Performing neurogenetic studies in other vocal production learners (e.g. cetaceans or elephants) is not feasible, mainly because of their CITES status and associated ethical concerns. Bats are a promising model taxon to study neurogenetics, given that several species are vocal learners (for review see Knörnschild, 2014). Point b) illustrates why babbling is such a valuable behavior to study genes related to vocal production learning.

In humans, FOXP2 is the most famous gene related to speech and language (Fisher & Scharff, 2009; Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). FOXP2 is a transcription factor that encodes for a regulatory protein, a class of proteins which is responsible for switching off and on other genes. FOXP2 probably involved in the development of speech, hence is an interesting candidate to investigate the molecular pathways and genetic networks underlying speech acquisition in infants. This gene was discovered in the KE family, half of whom suffered from severe speech disorders and impaired language comprehension (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998; Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). This discovery has prompted a large number of studies on FOXP2, also in non-human vocal learners and also non-vocal learners (for an overview see White, Fisher, Geschwind, Scharff, & Holy, 2006). The zebra finch advanced to a model species in FoxP2 related research. Similar to the findings in humans, where point mutations of FOXP2 led to impaired speech, the knockdown of FoxP2 in juvenile male zebra finches resulted in incomplete and inaccurate song imitation and production (Haesler et al., 2007).

FoxP2 has already been suggested to play a role in the development of echolocation (Li, Wang, Rossiter, Jones, & Zhang, 2007). Although several bat species are capable of vocal production learning, only *S. bilineata* is currently known to acquire a complete song through vocal production learning during ontogeny (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010) and even more important this learning process is indicated by babbling behavior. As mentioned before, pup babbling is the behavioral readout of learning taking place. This offers the unique opportunity to study the expression of FoxP2 during vocal learning processes in a mammal.

In humans, neuronal control underlying speech is characterized by direct (monosynaptic) connections between the cortex and the vocal motor neurons (Fitch, 2018; Jarvis, 2004). Studies about neuronal circuitry in vocal learning versus non-vocal learning species suggest that these direct connections are the reason for our ability to control tongue and lip movement during speech (Fitch, De Boer, Mathur, & Ghazanfar, 2016). This “direct connections hypothesis” was formulated based on comparative studies of vocal production learning in birds and other mammals (Fitch & Jarvis, 2013; Janik & Slater, 2000). In two out of three vocal learning avian taxa – in parrots and songbirds – it was demonstrated that direct connections between syringeal motor neurons and a specialized forebrain region exist (Jarvis, 2004). So far, neurogenetic studies in vocal learning songbirds demonstrated that non-human animal models can help to shed light on shared underlying traits (Brainard & Doupe, 2013; Doupe & Kuhl, 1999; Haesler et al., 2007; White, Fisher, Geschwind, Scharff, & Holy, 2006). The benefit of babbling behavior could also be useful to identify other genes, and neuronal and molecular networks involved in learning. Neurogenetic investigations are usually time-constrained. For example, the expression of immediate early genes (IEGs) in response to a particular behavior is constrained by a short time window. Some IEGs have been implicated with learning and memory processes (Bozon, Davis, & Laroche, 2003). Pup babbling in *S. bilineata* exactly pinpoints the time window of learning, hence allows investigating molecular mechanisms related to vocal production learning.

The interplay between social and vocal complexity

Social vocalizations are found in numerous species – but why are there social communication systems which are composed of a few vocalizations whereas others comprise large and diverse vocal repertoires? There are numerous benefits of sociality (e.g. shelter allocation, territorial defense, access to foraging grounds) but sociality also evokes additional cognitive difficulties to deal with (Freeberg, Dunbar, & Ord, 2012). Being a member of a group can introduce challenges such as recognizing and distinguishing between numerous individuals, assessment of social hierarchies and potential mates and remembering cooperative and deceptive partners (Freeberg, Dunbar, & Ord, 2012). These demands increase with the size of the social group, i.e. the number of interaction partners who could potentially be confounded. Social interactions are mediated by communication and a logical consequence of increasingly complex social structures is an increasingly complex communication system. The question whether social complexity requires and therefore promotes increased vocal complexity has led to the formulation of the "social complexity hypothesis for communication" (Bouchet, Blois-Heulin, & Lemasson, 2013; Freeberg, Dunbar, & Ord, 2012; Krams, Krama, Freeberg, Kullberg, & Lucas, 2012).

Vocal complexity can be obtained by different means. Possibilities are to increase the vocal repertoire size (i.e. the number of distinct units) or the compositional structure (i.e. combination of smaller units into larger ones) and to increase the amount of information encoded in an acoustic unit (i.e. vocal signatures). Vocal signatures encode a variety of information types, e.g. individual identity (e.g. Charrier, Mathevon, & Jouventin, 2003), population- (e.g. Riesch, Ford, & Thomsen, 2006) and social group affiliation (e.g. Jansen, Cant, & Manser, 2012), sex (e.g. Ey, Hammerschmidt, Seyfarth, & Fischer, 2007) and age (Erb, Hodges, & Hammerschmidt, 2013, results Chapter IV). Uncertainty decreases with increasing amount of information encoded by signalers but also decoded correctly by receivers (Bradbury & Vehrencamp, 2011). This gets more challenging if different information content is present in a vocalization. Vocalizations composed of multiple syllables (i.e. the smallest acoustic element surrounded by silence) and/or compositional syllables (i.e. the segmentation of a single element into distinct acoustic units) offer more available space to encode multiple messages and support signal reliability and error-free decoding by temporal segregation. Most studies investigating the "multiple-messages" hypothesis involve adult vocalizations (e.g. Koren & Geffen, 2009; Templeton, Greene, & Davis, 2005) or the investigation of a single call gradually changing during development (e.g. Ey, Hammerschmidt, Seyfarth, & Fischer, 2007). In my study (Chapter IV) I show that also pup vocalizations have the potential to encode multiple messages. Pups that are part of social groups probably not only benefit from signaling identity cues (e.g. individual and/or group signatures) but also information about age and sex. Identity information is crucial for mother-pup recognition (Knörnschild & von Helversen, 2008) whereas age could be supportive of nursing investment decisions. My result shows that multiple messages are not only found in adult vocalizations but that pup vocalizations encode more information than anticipated.

Vocal signatures can be measured in terms of entropy (how much uncertainty exists in communication) in bits of information (Shannon & Weaver, 1998). Beecher describes a method to analyze information in a communicative system measuring the stereotyped index H_s based on the information theory of Shannon & Weaver (Beecher, 1989; Shannon & Weaver, 1998). H indicates how much binary information is encoded in an acoustic signal, practically comparing the variance within and between individuals (i.e. noise and signal strength). I used this information theory approach as a proxy for vocal complexity (Chapter V). It is a useful method since this measure of H_s is comparatively simple to obtain from general acoustic parameter measurements. Especially the investigation of vocal signatures offers a perfect opportunity for making use of this approach since this type of information is usually always present in social vocalizations. I demonstrated a positive correlation between the amount of identity information encoded in social vocalizations (i.e. a proxy for vocal complexity) and the relevant social group size (i.e. a proxy for social complexity) (Chapter V). Evidence for this hypothesis has also been found in primates (McComb & Semple, 2005), rodents (Pollard & Blumstein, 2012), cetaceans (May-Collado, Agnarsson, & Wartzok, 2007), carnivores (Manser et al., 2014) and birds (Krams, Krama, Freeberg, Kullberg, & Lucas 2012).

In conclusion, bats are promising candidates for comparative acoustic analyses, especially concerning social vocalizations. Their diversity of social and vocal systems makes them particularly interesting for cross-species investigations. Moreover, the results of the first chapters of my thesis demonstrate that *Saccopteryx bilineata* is an excellent mammalian model species for comparative biolinguistic studies. The commonalities between infant speech acquisition and vocal ontogenetic processes in *S. bilineata pups* are a promising basis to discover common mechanisms and key factors which are necessary to acquire a complex communication system. Ultimately, this research could contribute to a better understanding of what may be unique to humans and bring us one step closer to the mystery of the evolution of human language.

References

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018). The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science, 21*(5), e12641. doi:10.1111/desc.12641.

Beecher, M. D. (1989). Signalling systems for individual recognition: an information theory approach. *Animal Behaviour, 38*(2), 248-261.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Frontiers in Psychology, 4*, 390. doi:10.3389/fpsyg.2013.00390.

Bozon, B., Davis, S., & Laroche, S. (2003). A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval. *Neuron, 40*(4), 695-701.

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication* (Vol. Second edition): Sinauer Associates Inc., Sunderland USA.

Brainard, M. S., & Doupe, A. J. (2013). Translating birdsong: songbirds as a model for basic and applied medical research. *Annual Review of Neuroscience, 36*, 489-517.

Charrier, I., Mathevon, N., & Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Animal Behaviour, 65*(3), 543-550. doi:10.1006/anbe.2003.2073.

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

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: parents simplify the structure of their speech in response to babbling. *Journal of Child Language, 46*(5), 998-1011.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998). 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends in Cognitive Sciences, 2*(1), 31-37.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998). Infant'babbling'in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour, 135*(5), 643-664.

Erb, W. M., Hodges, J. K., & Hammerschmidt, K. (2013). Individual, contextual, and age-related acoustic variation in Simakobu (Simias concolor) loud calls. *PLoS One, 8*(12), e83131. doi:10.1371/journal.pone.0083131.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., & Fischer, J. (2007). Age- and Sex-Related Variations in Clear Calls of Papio ursinus. *International Journal of Primatology, 28*(4), 947-960. doi:10.1007/s10764-007-9139-3.

Fagan, M. K. (2015). Why repetition? Repetitive babbling, auditory feedback, and cochlear implantation. *Journal of Experimental Child Psychology, 137*, 125-136. doi:10.1016/j.jecp.2015.04.005.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279-293.

Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics, 25*(4), 166-177.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature genetics, 18*(2), 168-170.

Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annual Review of Linguistics, 4*, 255-279.

Fitch, W. T., De Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science advances, 2*(12), e1600723.

Fitch, W. T., & Jarvis, E. D. (2013). Birdsong and Other Animal Models for Human Speech, Song, and Vocal Learning. *Language, Music, and the Brain: A Mysterious Relationship*, 499-540.

Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., . . . Oller, D. K. (2014). Effects of parental interaction on infant vocalization rate, variability and vocal type. *Language Learning and Development, 10*(3), 279-296.

Freeberg, T. M., Dunbar, R. I., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1597), 1785-1801. doi:10.1098/rstb.2011.0213.

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in Genie: a case of language acquisition beyond the “critical period”. *Brain and language, 1*(1), 81-107.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. *The Oxford handbook of developmental and comparative neuroscience*, 708-729.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*(1), 14.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*(6), 509-516.

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal Responsiveness and the Development of Directed Vocalizing in Social Interactions. *Infancy, 19*(4), 385-408. doi:10.1111/infa.12054.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS biology, 5*(12).

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

Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biology, 10*, 1741-7007.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016*(1), 749-777.

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

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften, 93*(9), 451-454. doi:10.1007/s00114-006-0127-9.

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. doi:10.1098/rsbl.2009.0685.

Knörnschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Animal Behaviour, 76*(3), 1001-1009. doi:10.1016/j.anbehav.2008.05.018.

Konishi, M. (1985). Birdsong: from behavior to neuron. *Annual Review of Neuroscience, 8*(1), 125-170.

Koren, L., & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. *Behavioral Ecology and Sociobiology, 63*(4), 581-590. doi:10.1007/s00265-008-0693-2.

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of Royal Society B: Biological Sciences, 367*(1597), 1879-1891. doi:10.1098/rstb.2011.0222.

Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(17), 9645-9646. doi:10.1073/pnas.1733998100.

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain? *Developmental science, 10*(1), 110-120.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., . . . Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684-686. doi:10.1126/science.277.5326.684.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413*(6855), 519-523.

Li, G., Wang, J., Rossiter, S. J., Jones, G., & Zhang, S. (2007). Accelerated FoxP2 evolution in echolocating bats. *PLoS One, 2*(9).

Manser, M. B., Jansen, D. A., Graw, B., Hollén, L. I., Bousquet, C. A., Furrer, R. D., & le Roux, A. (2014). Vocal complexity in meerkats and other mongoose species. In *Advances in the Study of Behavior* (Vol. 46, pp. 281-310): Elsevier.

Marler, P. (1970). Birdsong and Speech Development - Could There Be Parallels. *American Scientist, 58*(6), 669-673.

May-Collado, L. J., Agnarsson, I., & Wartzok, D. (2007). Phylogenetic review of tonal sound production in whales in relation to sociality. *BMC Evolutionary Biology, 7*(1), 136.

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters, 1*(4), 381-385. doi:10.1098/rsbl.2005.0366.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In J. F. K. a. C. A. F. G. H. Yeni-Komshian (Ed.), *Child phonology* (Vol. 1, pp. 93-112). New York: Academic Press, Inc.

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal Development as a Guide to Modeling the Evolution of Language. *Topics in Cognitive Science, 8*(2), 382-392. doi:10.1111/tops.12198.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology, 27*(20), 3162-3167. e3163.

Pollard, K. A., & Blumstein, D. T. (2012). Evolving communicative complexity: insights from rodents and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1597), 1869-1878. doi:10.1098/rstb.2011.0221.

Riesch, R., Ford, J. K. B., & Thomsen, F. (2006). Stability and group specificity of stereotyped whistles in resident killer whales, Orcinus orca, off British Columbia. *Animal Behaviour, 71*(1), 79-91. doi:10.1016/j.anbehav.2005.03.026.

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication*: University of Illinois press.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science, 308*(5730), 1934-1937. doi:10.1126/science.1108841.

Vihman, M. M. (2014). *Phonological development: The first two years*: Wiley-Blackwell Boston, MA.

White, S. A., Fisher, S. E., Geschwind, D. H., Scharff, C., & Holy, T. E. (2006). Singing mice, songbirds, and more: models for FOXP2 function and dysfunction in human speech and language. In: Society for Neuroscience.

Erklärung über die eigenständige Verfassung der vorgelegten Dissertation

Hiermit erkläre ich, dass die vorliegende Dissertation gemäß §7 Abs. 4 der Promotionsordnung des Fachbereichs Biologie, Chemie, Pharmazie der Freien Universität Berlin vom 25. April 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

Ahana Aurora Fernandez

The CV is not included in the online version for reasons of data protection.