Maternal behavior influences vocal practice and learning processes in the greater sac-winged bat

Authors: Ahana A. Fernandez[1,2,*], Nora Serve[2], Sarah-Cecil Fabian[2] & Mirjam Knörnschild[1,2,3]

Author affiliation:

[1] Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstrasse 43, 10115 Berlin, Germany.

[2] Department of Biology, Chemistry, and Pharmacy, Institute of Biology, Free University Berlin, Takustrasse 6, 14195 Berlin, Germany.

[3] Smithsonian Tropical Research Institute, Luis Clement Avenue, Bldg. 401 Tupper, Balboa Ancon, Republic of Panama.

[*] Corresponding author: fernandez.aurora.ahana@gmail.com

Abstract

Learning, particularly vocal learning, is often a social process. In human infants, it is well-established that social interactions influence speech acquisition and are hypothesized to modulate attentiveness and sensory processes, thereby affecting the speech-learning process. However, our understanding of how social interactions shape vocal ontogenetic processes in non-human mammals, particularly those which vocally learn, remains limited. In the bat *Saccopteryx bilineata*, pups acquire the adult vocal repertoire through a distinctive babbling behavior that shows interesting similarities to human infant babbling. While babbling encompasses many different syllable types, it is particularly noteworthy that pups learn song syllables by imitating adult singing males. The pups' social environment involves frequent interactions with their mothers, whereas adult males mainly serve as the primary source of

acoustic input. We monitored the vocal ontogeny of wild pups, investigating whether their social environment influenced three aspects of babbling: the amount of vocal practice, the pups´ final syllable repertoire size and the production of the syllable types acquired through vocal learning. The results demonstrate that maternal behavioral displays significantly influence the amount of vocal practice, the presence and versatility of song syllable types in babbling and the percentage of mature song syllables. Our findings show that maternal feedback plays a significant role in the vocal ontogeny and learning processes of *S. bilineata*, thus enhancing our understanding of the relationship between social feedback and vocal development in mammalian vocal learners.

Keywords: social feedback, babbling, vocal learning, vocal ontogeny, social environment, *Saccopteryx bilineata*

Introduction

Vocal development and learning are often fundamentally social processes (Carouso-Peck et al., 2021; Doupe & Kuhl, 1999; García, 2019; Gultekin & Hage, 2017; Janik & Slater, 1997; Takahashi et al., 2015). Vocal production learning (VPL) is a continuous trait encompassing various levels, including vocal convergence (i.e. gradual change of vocalizations based on social input) and vocal imitation of species-specific or heterospecific vocalizations and artificial sounds (Janik & Knörnschild, 2021). The social environment influences all levels of VPL to varying degrees, either passively through acoustic input or actively through vocal and behavioral feedback. For example, the social environment of goat kids passively influences the acoustic convergence of their innate contact call (Briefer & McElligott, 2012). Songbird species require acoustic input for successful song learning but vary in the degree to which they require social interactions (García, 2019). Social feedback has also been shown to influence the maturation patterns of the adult vocal repertoire in marmosets (Gultekin & Hage, 2017). To gain a deeper understanding of vocal learning and development, it is important to explore the social factors that influence it, the extent of their impact, and the consequences when these factors are absent. Questions about how the social environment influences vocal development and learning have begun to receive more research attention (Bistere et al., 2024; Carouso-Peck & Goldstein, 2019; Chen et al., 2016; Gultekin & Hage, 2017; Takahashi et al., 2016; Takahashi et al., 2017; West & King, 1988). However, there is need for more studies that explore interactions beyond the tutor-tutee relationship. Only a few studies have addressed this to date (i.e. zebra finch: Bistere et al., 2024; Carouso-Peck & Goldstein, 2019; i.e. brown-headed cowbird West & King, 1988), demonstrating the significance of this research direction. Further studies are needed – particularly in the wild (García, 2019) - to broaden the range of study species, with a special focus on mammalian vocal learners.

In this study, we explored the role of social feedback in one of the few non-human mammalian vocal learners (Janik & Knörnschild, 2021), the greater sac-winged bat *Saccopteryx bilineata*. This neotropical bat species is known for its extensive vocal repertoire, which includes male song (Behr & von Helversen, 2004). During ontogeny, pups acquire a part of the adult vocal repertoire by imitating the territorial song of adult male tutors (Knörnschild et al., 2010). This learning process is expressed through a vocal practice known as "pup babbling" (Fernandez et al., 2021; Knörnschild et al., 2006). Pup babbling shows parallels to human infant babbling (Fernandez et al., 2021), which makes *S. bilineata* a particularly exciting mammalian vocal-learning species to study. Babbling is a conspicuous and unique behavior: Over the course of 7 weeks, pups spend almost 30% of their active daytime hours babbling. Babbling bouts are 7 minutes on average but can last up to 43 minutes (Fernandez et al., 2021). Babbling bouts are composed of multisyllabic repetitive vocal sequences (i.e. syllable trains) containing adult-like syllable types (i.e. syllable types of the adult vocal repertoire) and undifferentiated proto-syllables (i.e. syllables only produced by babbling pups) (Fernandez et al., 2021). Both male and female pups babble (Fernandez et al., 2021), which is different from the male-biased plastic song production found in songbirds (Doupe & Kuhl, 1999; Marler, 1970) (but see (Odom et al., 2014) for female bird song). Although babbling bouts can result in maternal care in the form of nursing (Strauss et al., 2010a), babbling is clearly not a begging behavior; babbling differs from begging in structural composition (i.e. including adult and juvenile syllables) and its lack of context specificity (Ter Haar et al., 2021). To solicit maternal care pups specifically produce isolation calls which they deliver in a fast repetitive manner. Pup babbling is not tied to a specific function – pups either babble alone in the day-roost or interact with their mothers, mainly through various additional behaviors which can occur singly or in interactive sequences (Fig. 1A-C). Remarkably, females never produce these behavioral displays outside of the babbling context (Bradbury & Emmons, 1974; Tannenbaum, 1975; Voigt et al., 2008). Adult males do not

behaviorally interact with pups, but their vocalizations provide acoustic input in the form of conspecific adult song (Behr & von Helversen, 2004).

We studied the vocal ontogeny of 19 pups from 6 wild colonies during two consecutive field seasons (May-September 2016-2017) in Panama and Costa Rica (see methods). In the present study, we investigated how two primary social factors—the number of singing males (i.e. tutors) and the maternal behaviors displayed during a babbling bout—influenced the following three aspects of pup vocal development and learning: (1) the amount of vocal practice, (2) final syllable repertoire size and, (3) various characteristics of the song syllables learned through vocal imitation while babbling.

To assess maternal influence, we counted the total number of maternal behavioral displays occurring during babbling, focusing only on interactive behaviors and excluding comfort behaviors (Table S1). Territorial song is produced by adult territorial males daily at dusk and dawn for acoustic territorial defence (Behr & von Helversen, 2004). Thus, the number of singing males (i.e. tutors) present in a colony was used as a proxy for the amount of social acoustic input a given pup received during its ontogeny. Pup age was included in analyses to account for physical maturation.

Fig. 1.

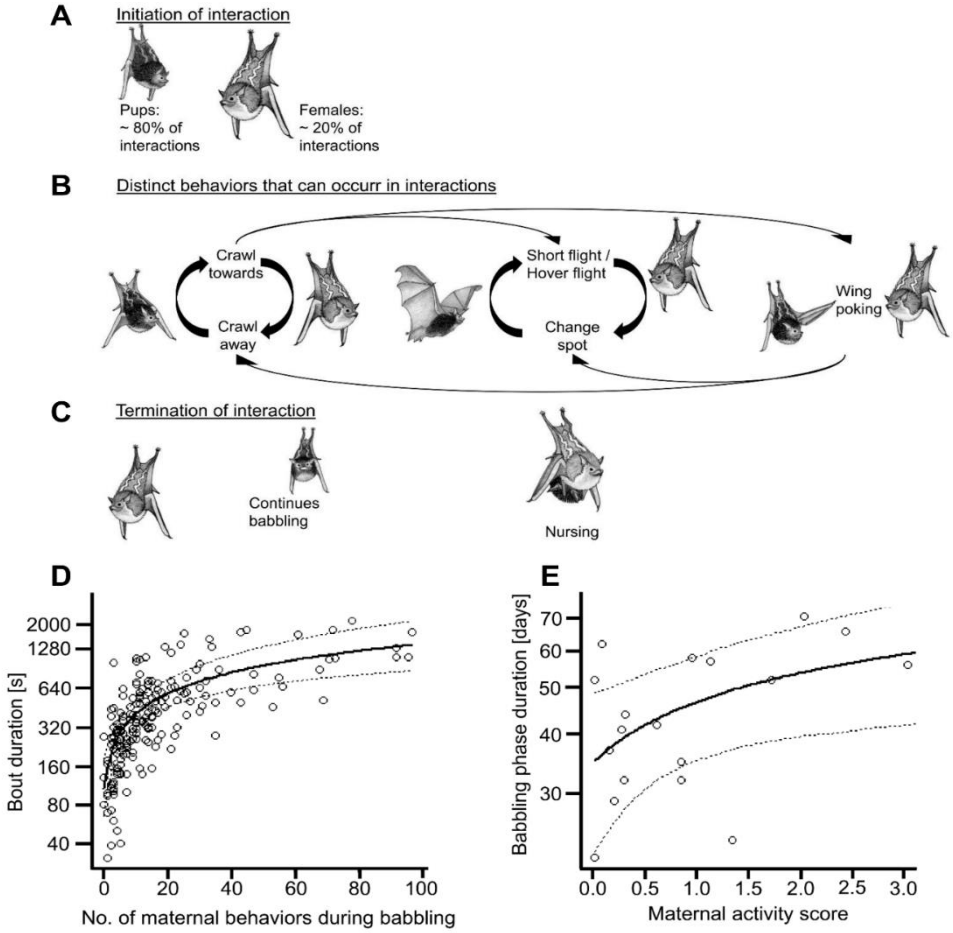

Fig. 1. Behavioral displays and interactions during babbling and maternal influence on the amount of vocal practice.

Fig. 1 (A) to (C) shows the typical social behaviors and interactions of mother-pup pairs during a babbling bout. Note: all these behaviors could occur singly or in an interactive way (i.e. a behavioral display was followed by another in less than one second). (A) In on average 80% of cases, the interaction was initiated by the pup, normally by crawling towards the mother while babbling. A mother initiated interactions by hovering in front of or landing next to her pup. (B) After the start of an interaction, mother and pup engaged in a number of different behaviors, which were performed equally by both mother and pup (Table S1), with the exception of "wing poking", which was mostly observed in pups (Table S1). The different behaviors could variably succeed after one another (large arrows); however, "crawl towards" and "crawl away", as well as "short flights" respectively "hover flights" and a temporarily "change in the roosting spot" were often observed in repetitive sequences (bold black arrows). Hover flights in front of the interactive partner were the most conspicuous behavior, followed by the short flights either next to the interaction partner or within the day roost. With increasing pup age, short flights became progressively longer, e.g. pups even briefly left the day roost and landed on an adjacent tree. (C)

Interactions between mother and pup could be terminated in two ways: Either the pup continued to babble alone or the mother allowed nursing. Maternal behavior did not differ with regard to pup sex (SI). Drawing credit: C. A. S. Mumm. (D) shows the positive influence of the maternal behaviors displayed during babbling bouts on babbling bout duration [s] (N=19 pups, N=186 babbling bouts, GLMM, family Gamma with log link). (E) shows that higher maternal activity scores led to a longer babbling phase duration [days] (N=19 pups, LMER). The figure legends represent the raw data whereas the fitted lines (with lower (2.5^{th} percentile) and upper (97.5^{th} percentile) confidence intervals) are based on the calculated model. Neither bout duration nor babbling phase duration differed between pup sex (SI).

Results

Amount of vocal practice

The amount of vocal practice was defined as a) the length of daily practice, i.e. babbling bout duration in seconds, and b) the overall babbling phase duration, i.e. time in days from the first day of babbling until the last (see methods). The total number of maternal behaviors performed during babbling bouts, as well as pup age, had a significant effect on the duration of daily babbling bouts, whereas the number of tutors had no effect (Table 1, Fig 1D). This finding does not result from the fact that longer bouts simply offer the possibility for more maternal displays (see Table S2). Interestingly, some mothers were more active than others in the sense that they showed more and different behavioral displays when interacting with pups. Therefore, we calculated a maternal activity score for each female to capture its overall behavioural activity (see methods). We found that the amount of maternal activity significantly influenced the duration of the babbling phase whereas, once again, the number of male tutors did not (Table 1, Fig. 1E). These findings show that maternal behaviors significantly shape the amount of time pups spend vocally practicing—age-related increases in endurance and physical maturity explain only part of the variance (Table 1).

Table 1

Response variable	Parameter	Estimate (SE)	Significance
A *Babbling bout duration [s]*	*Fixed effects*		
	Intercept	6.038 (0.074)	p < 0.001
	Mat. behav. during babbling (standardized)	0.537 (0.044)	**p < 0.001**
	Pup age (standardized)	0.100 (0.044)	**p = 0.022**
	No. of tutors (standardized)	-0.041 (0.077)	p = 0.6
	Random effect (standard deviation)		
	Pup ID	0.19	
B *Babbling phase duration [days]*	*Fixed effects*		
	Intercept	44.322 (5.222)	p = 0.001
	Maternal activity score (standardized)	7.317 (2.824)	**p = 0.02**
	No. of tutors (standardized)	3.560 (5.167)	p = 0.530
	Random effect (standard deviation)		
	Colony ID	11.20	

The influence of the social environment on the amount of vocal practice.

Response variables (first column): **A:** *Babbling bout duration [s]:* the babbling bout duration measured in seconds (N=19 pups, N=186 babbling bouts, GLMM, family Gamma with log link, random factor pup ID: repeated measurements over time of the same focal pups). **B:** *Babbling phase duration:* The number of days pups spent babbling (N=19 pups, LMER, random factor colony ID: measurements of pups from the same colony). The predictor variables for both models were z-transformed, a standardisation procedure that facilitates convergence of the model. The second column depicts the fixed and random effects for both models. The third column depicts the estimate and the standard error as well as the standard deviation for the random effect, and the last column the p-value (bold if influence is significant). Abbreviations: "Mat. behav." = maternal behaviors.

Box 1: Terminology

Term	Description
Syllable type	Sound unit with characteristic spectro-temporal features
Undifferentiated proto-syllables	Syllables only produced by pups, occurring in babbling and vanish after weaning.
Adult-like syllable	Syllables which are present in the later adult vocal repertoire (after weaning).
Buzz syllable types of song	Territorial song syllables with discernible pulsed part; 5 distinct syllable types: B1, B2, B2 trill, B3, B4 (only syllable with smeared buzz part) (Fig. S1).
Mature syllable	Mature syllables show no gap (e.g. between the pulsed and tonal part, within the buzzed part) (Fig. S2).
Precursor syllable	Distinct gaps within the pulsed part and/or between pulsed and tonal part (Fig. S2).
Early babbling phase	First half of entire babbling phase
Late babbling phase	Second half of entire babbling phase
Song syllables	Syllables of which the territorial song is composed of (Fig. 2.A)

Final repertoire size

The final repertoire size is defined as the number of adult-like syllable types (i.e. syllables that were identifiable as syllables of the later adult vocal repertoire) present in pup babbling at the time point of weaning (analyzed for a prior study (Fernandez et al., 2021), which reports the relevant methods).

The maternal activity score was not significantly correlated with the pups´ final syllable repertoire size (Spearman rank correlation: r_s=0.05, p=0.88, N=10 mother-pup dyads). Likewise, the number of singing males in a day-roost did not correlate with the pups´ final syllable repertoire size (Spearman rank correlation: r_s=0.52, p=0.13, N=10 pups). However, we found a trend for a positive correlation between the mean babbling bout duration of pups and their final syllable repertoire size (r_s=0.58, p=0.08, N=10 pups).

Learned song syllables

Babbling *S. bilineata* pups use vocal imitation to acquire the syllables of their territorial song (Knörnschild et al., 2010). Babbling is therefore a behavioral indicator of learning. During babbling, the pup's song precursors strongly resemble the songs of their acoustic tutors, regardless of their genetic relatedness to the tutor. The acoustic changes which the song syllables undergo cannot solely be attributed to maturation processes or convergence towards a species mean, indicating that pups are influenced by the songs they hear on a daily basis (Knörnschild et al., 2010). To understand how the social environment impacts learning processes, we examined the characteristic "buzz" syllable types of the territorial song (Fig. 2A). The territorial song is composed of five different song syllable types (Fig. S1). However, pups differ in which and how many of these five different song syllable types are present in their babbling. We therefore calculated for each babbling bout the song syllable type versatility (see methods): high values indicate that the pups' vocal practice contained many of these different types, low values indicate that the pups' vocal practice contained fewer different syllable types. Furthermore, the most common song syllable type (Table S3) was categorized into mature and precursor syllables (see methods) to assess whether the social environment had an influence on the presence of mature syllables in babbling. Subsequently, we investigated whether the social environment influenced the a) total number of song syllables present in a babbling bout, b) the song syllable type versatility in a babbling bout and c) the percentage of mature versus precursor syllables of the most common song syllable type in babbling bouts. Both maternal behaviors and pup age positively influenced the total number of buzz syllables in babbling bouts, while the number of male tutors had no effect (Table 2, Fig. 2B). Similarly, the versatility was influenced by maternal behavior and pup age, with no influence from the number of male tutors (Table 2, Fig. 2C). The maternal behavior also influenced the percentage of mature song syllables present in babbling, with no effect of pup age and the number of male tutors (Table 2, Fig. 2D). Additionally, we quantified the

acoustic change of the most common learned syllable type (B2, Table S3) during ontogeny and explored its relationship with maternal activity. We used classical acoustic parameters (e.g., duration, peak frequency) and linear frequency cepstral coefficients (LFCCs) to assess both traditional acoustic features and timbre (voice color) of B2 syllables (see methods). Discriminant function analyses were performed to obtain centroids (mean canonical scores for each individual) for the early and late babbling phase (see methods). Subsequently, we calculated Euclidean distances between the early and late babbling phase for each individual, where larger values indicated greater acoustic changes during the babbling phase. Our findings revealed a positive correlation between acoustic changes and maternal activity (Spearman rank correlation: r_s=0.73, p = 0.0037, Fig. 2E).

Fig. 2.

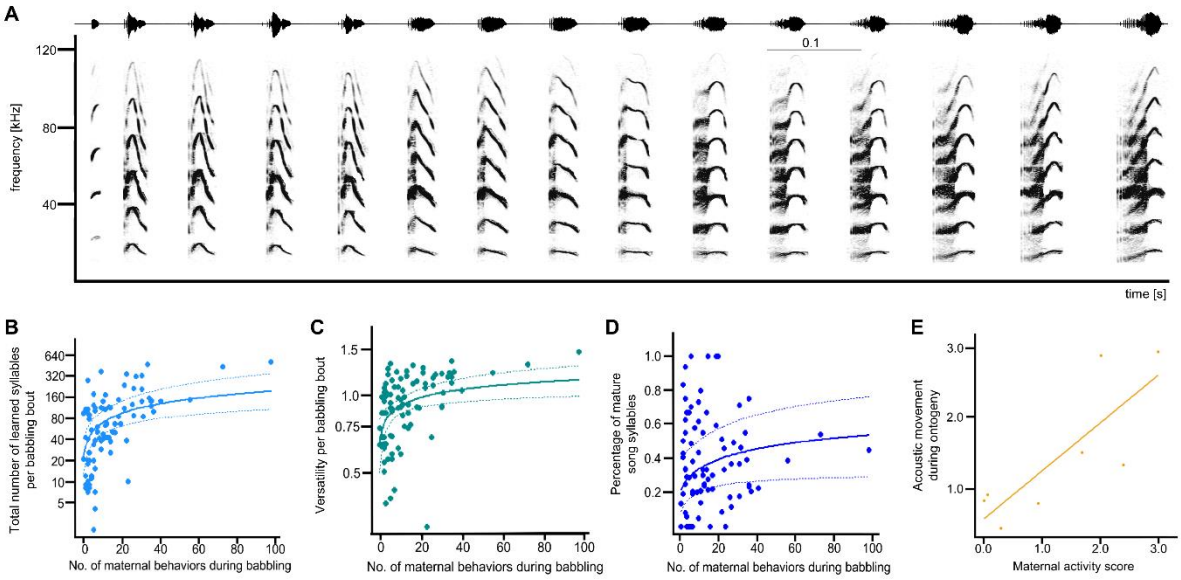

Fig. 2. Maternal influence on learned syllables in babbling bouts.

Fig. 2 (A) shows a typical multisyllabic territorial song produced by an adult male. The most characteristic syllable type is called "B2" (i.e. the last six syllables of this song), composed of a buzzed part connected to a tonal part. This is also the most common song syllable type in pup babbling (Table S3). The spectrogram was created in Avisoft SasLab Pro (Hamming window, 1024 FFT, 50% overlap resulting in 317 Hz frequency and 2.048ms time resolution). (B) shows the positive significant influence of maternal behavioral displays on the total number of learned syllables within babbling bouts (N=10 pups, N=90 babbling bouts, GLMM, family negative binomial with log link). Territorial songs can be flexibly composed of different buzz syllable types (Fig.

S1). (C) shows that maternal behavior positively influences the versatility (i.e. number of different learned syllable types) of babbling bouts (N=10 pups, N=90 babbling bouts, LMER). (D) The most commons song syllable type B2 (Table S3) was categorized into precursor and mature syllables. Maternal behavioral displays positively influenced the percentage of mature song syllables (N=10 pup, N=90 babbling bouts, GLMM family binomial with logit link). The figure legends represent the raw data whereas the fitted lines (with lower (2.5^{th} percentile) and upper (97.5^{th} percentile) confidence intervals) are based on the calculated model. (E) shows the acoustic movement of pups during their vocal ontogeny based on acoustic parameters of the syllable type B2, the most common song syllable type (N=8 pups, N=220 syllable trains). The larger the value on the y-axis, the larger the acoustic movement during the pup´s vocal ontogeny. The acoustic movement was positively correlated with the maternal activity score.

Table 2

Response variable	Parameter	Estimate (SE)	Significance
A *Total number of song syllables*	*Fixed effects*		
	Intercept	4.188 (0.194)	p < 0.001
	Maternal behavior during babbling (standardized)	0.460 (0.098)	**p < 0.001**
	Pup age (standardized)	0.446 (0.096)	**p < 0.002**
	No. of tutors (standardized)	0.036 (0.193)	p = 0.9
	Random effects (standard deviation)		
	Pup ID	0.55	
B *Song syllable versatility*	*Fixed effects*		
	Intercept	0.889 (0.030)	p < 0.001
	Maternal behavior during babbling (standardized)	0.105 (0.031)	**p < 0.001**
	Pup age (standardized)	0.125 (0.030)	**p < 0.001**
	No. of tutors (standardized)	-0.022 (0.030)	p = 0.479
	Random effects (standard deviation)		
	Pup ID	0.04	
C *Percentage of mature song syllables*	*Fixed effects*		
	Intercept	-0-698 (0.355)	p = 0.049
	Maternal behavior during babbling (standardized)	0.376 (0.188)	**p = 0.045**
	Pup age (standardized)	0.185 (0.177)	p = 0.296
	No. of tutors (standardized)	0.107 (0.341)	p = 0.754
	Random effects (standard deviation)		
	Pup ID	0.98	
	Observation-level	1.24	

The influence of the social environment on learned syllables.

Response variables: **A**: *Total number of song syllables*: the learned territorial song syllables which were present in a babbling bout, (N=10 pups, N=90 babbling bouts, GLMM, family negative binomial with log link, random factor ID: repeated measurements over time of the same focal pups). **B**: *Song syllable versatility*: how many of the five different song syllable types (Fig. S1) were present in a babbling bout, (N=10 pups, N=90 babbling bouts, LMER, random factor ID: repeated measurements over time of the same focal pups). Neither the total

number of song syllables nor song syllable versatility differed between pup sex (SI). **C**: *Percentage of mature song syllables*: We investigated the influence of different predictor variables on the percentage of mature song syllables of the most common syllable type, B2 (Table S3, N=10 pup, N=90 babbling bouts, GLMM family binomial with logit link, random factor ID: repeated measurements over time of the same focal pups, random factor observation-level: to avoid overdispersion). The predictor variables for both models were z-transformed, a standardisation procedure that facilitates convergence of the model. Fixed and random effects are depicted in the second column. The third column depicts the estimate with standard error, and the last column the p-value (significant influence indicated in bold). Abbreviations: "Mat. behav." = maternal behaviors.

Discussion

Our findings demonstrate that both maturational effects (i.e. pup age) and social feedback (i.e. maternal feedback) together influenced various aspects of pup babbling behavior whereas the presence of male tutors had no influence on any of these aspects. In *S. bilineata,* babbling very likely serves multiple functions which are not mutually exclusive. One function of babbling is the acquisition of the learned syllables of the territorial song. The territorial song is an important acoustic signal for adult bats (Behr et al., 2009; Knörnschild et al., 2017). In *S. bilineata*, unlike most other mammals, females disperse after weaning, while most males stay in their natal colony (Nagy et al., 2007). Adult males queue for territory access. Once they obtain a territory, they mark it through territorial singing, a daily activity influenced by various social factors (Eckenweber & Knörnschild, 2013). The territorial song encodes a group signature (Eckenweber & Knörnschild, 2013) that may function as a password to identify social group members and potentially assists in the cooperative defense of the colony against unrelated intruders (Nagy et al., 2012). This group signature might be acquired during the babbling phase, when pups listen to singing adult males. An intriguing aspect of this species is that, unlike most song-learning bird species, female pups show no differences from males in babbling behavior and vocal development (Fernandez et al., 2021; Knörnschild et al., 2010). This study corroborated this finding: female pups received the same maternal

feedback, and their song syllable imitation did not differ in any way from male pups (as observed as well in Knörnschild et al., 2010). This phenomenon is rare among vocal learners and raises the question of why female pups match male vocal development despite not using the learned vocalizations later in life. One potential explanation might lie in the function of the territorial song for adult females: it serves as an acoustic signal to help females locate new suitable colonies after dispersal. The territorial song exhibits different dialects, with females showing a preference for local over foreign dialects (Knörnschild et al., 2017). The own early practice and production of song might enhance the ability to evaluate male song and support mating decisions. The female preference for the dialect learned during ontogeny likely drives the selection of precise imitative learning in males. This might explain the exceptional and long vocal learning and practice phase during ontogeny. It could also explain why maternal feedback during this period positively influences the production (i.e. amount and versatility) of learned syllables. The process of vocal imitation involves learning through self-perception: An individual hears its own vocal output and matches it to the acoustic input of the tutor (Brainard & Doupe, 2000). Increased practice of learned syllable types could strengthen the connection between acoustic perception and vocal production in a manner akin to the mechanisms observed in human infants and songbirds (Doupe & Kuhl, 1999; Kuhl, 2007). Furthermore, such practice is likely to improve fine-scale control of the vocal apparatus in a way that will eventually allow for the production of a mature syllable or song. Our data show that *S. bilineata* mothers also influence the presence of mature song syllables of the most commonly produced syllable type within babbling bouts. In human infants, contingent social feedback, both vocal and non-vocal (e.g. touch, smiling, moving closer to the infant), can positively influence babbling (Goldstein et al., 2003; Goldstein & Schwade, 2008). For example, social feedback can increase the rate of production of mature speech sounds in babbling (Goldstein et al., 2003). However, it should be noted that this finding does not indicate that contingent social feedback increases the size of the infant's babbling

repertoire through imitation, but rather that it induces a shift towards the production of mature speech sounds that are already present in the infant's repertoire. The debate about how much imitation is present in babbling during social interactions, and how this imitation manifests itself, is still ongoing. One interesting possibility is that infants imitate prosodic contours in dyadic interactions with their caregivers (Gratier & Devouche, 2011). In most songbird studies, similarity between tutor and tutee defines song learning success (García, 2019). While this study did not assess the similarity between pups and their adult tutors (due to the limited availability of song data from all adult tutors), we were able to investigate the acoustic trajectories of individual pups during their babbling phase. Our data demonstrates that higher maternal activity was associated with larger overall changes in acoustic parameters in the most commonly learned syllable type (B2). This finding serves as an initial indication that non-vocal interactions with the mother may influence a pup´s individual learning trajectories. Future studies will focus on the relationship between acoustic change, maternal feedback, and learning success, specifically investigating contingencies between particular pup vocalizations and maternal displays in natural settings. Playback experiments are an additional approach to test the impact of contingency on vocal learning. For example, one study in zebra finches demonstrated that contingent non-vocal maternal feedback affects imitation success (Carouso-Peck & Goldstein, 2019), while another recent study found that female calls can promote song learning, but the role of contingency remains to be determined (Bistere et al., 2024).

Beyond the territorial song, it remains unclear whether and which other syllables from the adult vocal repertoire are acquired through vocal learning. However, producing a diverse vocal repertoire that includes syllables with distinct spectro-temporal features may necessitate practice, regardless of whether they are learned socially or not. Vocal practice also involves the transitions between the different syllable types, and possibly even the acquisition of specific sequential orders for later multisyllabic adult vocalizations.

While we did not find a direct correlation between maternal feedback and the final syllable repertoire size of pups, babbling time may play a role in expanding the vocal repertoire, as indicated by the positive correlation trend between the final syllable repertoire size and the mean bout duration. Maternal behaviors significantly increased the amount of vocal practice (Fig. 1D-E, Table 1), which might indirectly support the maturation and acquisition of a large vocal repertoire. Such a relationship has already been documented in common marmosets, where social feedback influences the maturation (i.e. the development of the adult calls) of the vocal repertoire (Gultekin & Hage, 2017).

It may appear surprising that the number of singing males did not influence any aspects of pup vocal practice and learning. In songbirds, acoustic input from and social interaction with the tutor plays a pivotal role in song learning (Chen et al., 2016; García, 2019). In *S. bilineata*, direct interactions between pups and adult males were absent. Nevertheless, it is possible that the amount of adult male singing could influence pup practice. The present study cannot directly address this relationship, since the measure of song activity applied—the number of singing males—is an indirect indicator. In future studies, we will include the total number and duration of songs to which pups are exposed to and examine whether this might influence various aspects of their vocal practice and learning. Furthermore, the findings of this study are purely observational and future captive playback studies will provide further insight into how social effects affect vocal practice and learning.

In this study, we focused on how social feedback shaped babbling. However, babbling—itself a salient vocal signal—likely also influences social feedback (Ter Haar et al., 2021). This might also apply to *S. bilineata* where pups initiated ~ 80% of social interactions, suggesting that maternal feedback is likely influenced by the pup´s vocal practice. In humans, parents are sensitive to their infants' vocalizations, can intuitively assess speech maturation, and are more likely to respond to sounds that resemble speech (Abney et al., 2017; Albert et al., 2018; Oller et al., 2001). This can initiate a positive feedback loop: when infants receive contingent adult

feedback for their speech-like utterances, they tend to produce more speech-like vocalizations (Goldstein & Schwade, 2008; Warlaumont et al., 2014). Babbling in humans is also hypothesized to serve as a vocal fitness signal, potentially evolved to elicit more parental care thus, gaining importance throughout human evolution (Oller et al., 2016). In this context, it would be highly interesting to study the role of hormones and neuromodulators (Theofanopoulou et al., 2017). In the case of *S. bilineata*, it would be intriguing to investigate whether babbling elicits oxytocin release in mothers and whether, in turn, this hormone modulates maternal social feedback.

In conclusion, our study shows that maternal social feedback shapes different aspects of the vocal practice and learning of pups. As bats are highly social mammals that share key aspects of human brain architecture and vocal development, our findings will substantially advance our knowledge about how social interactions shape vocal learning.

Materials and methods

Study species

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 are weaned at 10-12 weeks of age (Tannenbaum, 1975). Day-roosts of *S. bilineata* are located in tree cavities or outer walls of man-made structures (Yancey et al., 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; Fernandez et al., 2021; Knörnschild et al., 2006). Individuals in the day-roost maintain an inter-individual distance of five to eight centimeters (Voigt et al., 2008). Only pups are allowed to ignore this inter-individual distance without eliciting aggression.

This species possesses a large vocal repertoire (i.e. 25 distinct syllable types) which is fully delineated (Behr & von Helversen, 2004; Fernandez et al., 2021; Knörnschild & von Helversen, 2008). These syllable types are composed into mono- or multisyllabic vocalizations. For most vocalizations, the behavioral context and function is known as well (e.g. the territorial song is an aggressive signal used during territorial defense).

Study sites

Acoustic recordings and observations of social behaviors were recorded *ad libitum* sensu Altmann (Altmann, 1974) throughout the pups´ ontogeny (i.e. from birth until weaning) in their day-roost. In 2016, we conducted acoustic recordings and additionally monitored the accompanying behaviors from ten 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 in Panama.

Research permits

All data collection in Panama (2017) was approved by the local governmental authorities and the Smithsonian Tropical Research Institute Animal Care and Use Committee (ACUC 2017-0516-2020). All data collection in Costa Rica (2016) was approved by the Costa Rican Ministry of the Environment (SINAC-SE-CUS-PI-R-088-2016).

Acoustic and behavioral data collection

In 2016 and 2017, acoustic recordings and behavioral observations of mother-pup dyads during babbling were performed simultaneously. During all field seasons, we counted the number of singing males present in a colony, representing the indirect social environment of our focal pups. Since adult bats are individually banded with colored plastic rings on their forearms (A.C. Hughes Ltd. UK, size XCL, one band per forearm) and mothers only nurse their own, single pup, individual identity of our focal bats could be determined. The inter-individual distance outside of the mother-pup pairs and discrimination of the individually

banded mothers between own and alien pups allowed for focal acoustic recordings and behavioral monitoring. After habituation, it was possible to record and observe focal bats within a distance of two to four meters without noticeable disturbance.

During both field seasons, focal pups were recorded at least twice per week, during alternating morning or afternoon recording sessions (approx. five-hour total recording time) throughout their vocal ontogeny (i.e. from birth until weaning, for more details see methods in (Fernandez et al., 2021)). 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 $\pm$ 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. Behaviors of mother-pup dyads during babbling were monitored using a combination of video recordings and simultaneously taking notes, to ensure that all behaviors were captured. The video camera was mounted next to the microphone on the tripod.

Acoustic data analysis

Acoustic analyses of babbling were performed on two different levels; the syllable level (i.e. the basic unit) and the babbling bout level (i.e. composed of syllable trains and interspersed with silent intervals; see methods (Fernandez et al., 2021)).

The amount of vocal practice was measured in two ways: the duration (in seconds) of babbling bouts and the duration of the babbling phase. A babbling bout is composed of at least three syllable trains which belong to at least two different syllable train categories and has a minimum duration of 30 seconds (for details see (Fernandez et al., 2021)). A syllable train is defined as a sequence of at least five syllables (for details see (Fernandez et al., 2021)). A babbling bout was either terminated if a pup remained silent for longer than three

minutes and/or when a pup was allowed to nurse. The babbling phase duration (in days) spans the period during which the pups were observed to engage in babbling. The day of birth was used to calculate pup age in general, age at babbling onset and the duration of the babbling phase (for details see (Fernandez et al., 2021)). The daily bout duration (in seconds) and babbling phase duration (in days) was obtained during both field seasons (N=186 babbling bouts in total, Costa Rica 2016: N=123 bouts from ten pups, Panama 2017: N=63 bouts from nine pups). The final repertoire size of 10 pups (Costa Rica 2016) was analyzed for a previous study (Fernandez et al., 2021) (method details can be found there). The number of adult-like syllable types (i.e. the syllables that were identifiable as syllables of the later adult vocal repertoire) present in pup babbling at the time point of weaning represented the pups´ final syllable repertoire size. The territorial song is composed of different syllable types whose presence varies with behavioral context (Eckenweber & Knörnschild, 2013; Knörnschild et al., 2017; Knörnschild et al., 2016). The typical territorial song (Figure 2A) which is sung at dawn and dusk starts with simple, short tonal syllables and ends with composite syllables (with pulsed and tonal components) which are called buzz syllables. The tonal and buzz syllables are categorized into different syllable types (Fernandez et al., 2021). For this study, we focused on the buzz syllable types because those are the syllables imitated from tutors (Knörnschild et al., 2010) and they encode most information about the signaler (i.e. individual-, group and population signature) (Eckenweber & Knörnschild, 2013; Knörnschild et al., 2017; Knörnschild et al., 2016). Moreover, the simple tonal syllable types are also present in other vocalization types, some of which are present directly after pup birth. To analyze the influence of the mothers´ behavior on the buzz syllable types we selected babbling bouts of 10 pups (N=90 babbling bouts, Costa Rica 2016) where we had simultaneous behavioral observations and acoustic recordings (with good signal to noise ratio). If possible, we selected one babbling bout per week and pup during the babbling phase and manually labelled all present buzz syllables (Fig. S3). We defined five different buzz

syllable types based on their spectro-temporal properties, namely B1, B2, B2 trill, B3, and B4 (Fig. S1) following our categorization in (Fernandez et al., 2021). Our visual classification was statistically confirmed in an earlier study (Fernandez et al., 2021), and syllables in this study were classified by the same observer. In total, we manually labelled 9201 buzz syllables (N=90 babbling bouts, Costa Rica 2016). We then performed the following measures: a) count of all present buzz syllables (but without B1 because it was not present in each pup) and b) the buzz syllable versatility (R: *vegan* package, function diversity) including all five syllable types present in a babbling bout. Furthermore, we identified the number of mature and precursor syllables of the most common syllable type B2 (Table S3, N=4847 B2 syllables, N=90 babbling bouts, Costa Rica 2016). Syllables were classified as precursor if there were distinct gaps within the pulsed part and/ or within the pulsed and tonal part (Fig. S2). This is clearly not the case in mature syllables of adult males (Fig. 2A, Fig. S2). We focused on B2 syllables because this was the most abundant syllable type in babbling bouts for all pups and because it is the syllable type characteristic of the typical adult male song (Fig. 2A). To investigate the influence of the social environment on the change of acoustic properties during ontogeny, we measured 530 B2 syllables from 8 pups (for 2 pups of the Costa Rica dataset there was not enough data from the entire ontogeny; therefore, they were excluded for this analysis) and 205 B2 syllables from adult males (for explanation see statistical analysis below). Again, we focused on B2 syllables because this was the most abundant syllable type in babbling bouts for all pups and because it is the syllable type characteristic of the typical adult male song (Fig. 2A). To be able to investigate the acoustic trajectory during ontogeny, we measured B2 syllables from different babbling bouts throughout the entire vocal ontogeny of each pup. For each pup, we normally selected three B2 syllables from one to three syllable trains per babbling bout (N=220 trains in total). It was not always possible to measure three B2 syllables from three trains in a babbling bout either because there were simply not so many B2 syllables (especially at the beginning of the

babbling phase) or because the signal-to-noise ratio quality was not sufficient for acoustic measurements. To quantify the acoustic change during ontogeny, we separated the entire babbling period into an early and late babbling phase (for more details see statistical analysis below). For each pup, we analysed the same number of syllable trains per babbling phase (7-20 trains per pup for each babbling phase). Subsequently, we measured various acoustic parameters of the B2 syllables. For our acoustic parameter measurement analysis, we used the software Avisoft SASLab Pro (v.5.3.01). Before measurements, we converted the sampling frequency of each acoustic recording from 500kHz to 250kHz to obtain a higher frequency resolution. The buzz syllables are composed of a pulsed (buzz-like) part preceding a tonal part. We extracted the same acoustic parameters for the pulsed and tonal part separately (Fig. S4A). Acoustic parameters were subsequently averaged (separately for pulsed and tonal parts) for each train to minimize temporal dependence among syllables produced within close succession. Spectrograms for measuring syllable parts were created using a Hamming window with 1024–point Fourier transformation and 93.75% overlap (frequency resolution: 244Hz, time resolution: 0.256 ms). Syllable parts were multi-harmonic; measurements were taken from the harmonic containing most energy (i.e. clearest spectral shape) and the extracted acoustic parameters were subsequently calculated to the fundamental frequency (first harmonic) to make measurements comparable across the different syllable parts. For each syllable part, we measured one temporal (duration) and six spectrum-based parameters (peak-, minimum-, maximum frequency, bandwidth, entropy and harmonics-to-noise ratio). These spectrum-based parameters were measured at the center of each syllable part and at 9 points evenly distributed over each syllable (start, end, and seven intermediate locations). In addition, the six spectrum-based parameters were averaged over each syllable part (mean values) and the maximum spectral values of the entire syllable part were calculated. The spectrum-based parameters were used to estimate the frequency curvatures and entropy and harmonics-to-noise curvatures (henceforth referred to as entropy curvatures) of the syllable

parts. Furthermore, we extracted acoustic features that were based on linear-frequency cepstral coefficients (LFCCs) since these capture important acoustic characteristics of bat vocalizations (Fernandez & Knörnschild, 2020; Knörnschild et al., 2017). Each LFCC describes the spectral properties of an entire acoustic signal, comprising its most important features in a compact form. The extracted features summarize both common acoustic parameters (e.g. peak frequency) and timbre (voice color). We used a customized MATLAB script in the toolbox "voicebox" (v. R2014a) for the feature extraction. Each vocalization sequence was composed of B2 syllables of the same train containing the first four harmonics (F0-F3, frame length 0.01 s). We extracted 20 LFCCs from each vocal sequence and used them for subsequent statistical analyses.

Behavioral data analysis

To describe and quantify the behaviors of mothers and pups during babbling bouts, videos were analyzed using VLC (VLC media player), complemented by field notes. We established an ethogram (Table S1) based on our observations of mother-pup behaviors and the behavioral observations of a former study (Strauss et al., 2010b). Inter-observer-reliability was 100% among the observers. Based on this ethogram, we counted the maternal behavioral displays (N=186 babbling bouts, N=3221 interactive behaviors, N=19 mothers) occurring during babbling, focusing only on the interactive behaviors, excluding comfort behaviors (see Table S1).

To assess the mothers` influence over the course of the entire pup babbling phase, we calculated a maternal activity score (based on 186 babbling bouts). We calculated behavioral quotient for each babbling bout (i.e. number of maternal interactive behaviors divided by the bout duration) and subsequently averaged these quotients for each female. We calculated also the mean hover rate per bout. We counted hover behaviors (i.e. hovering in mid-air in front of a roosting conspecific) separately because it is a very distinctive behavior that is never performed outside of a social context; for instance, hovering is produced by adult males

during courtship in a multimodal display (Behr & von Helversen, 2004). The mean behavioral quotient and the mean hover rate were summed up, and we obtained a maternal activity score ranging from 0.03-3.03. Lower numbers accounted 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= 0.27, p=0.25).

Statistical analysis

All statistical analyses were performed in R (RStudio 2021, version 1.4.1106) and SPSS (IBM SPSS Statistics 2022, version 29.0.0.0.(241)). Statistical differences were considered significant for p <0.05, *p<0.05, **p<0.01, ***p<0.001. Since we did not obtain enough high-quality recordings and/or behavioral observations for all mother-pup dyads, we had to restrict some analyses to a data subset. Therefore, sample sizes differed for our analyses. All generalized linear mixed models (GLMMs) and LMERs were fitted using the package *lme4* and *lmerTest* in R. In all multivariate models, the predictor variables were z-transformed (standardized: i.e. meaning the substraction of the mean and division by the standard deviation), a standardization procedure that facilitates convergence of the model (Gelman & Hill, 2006). For standardization we used the function *scale()*. In all models, the variable maternal behavior was log-transformed before it was standardized. This transformation was applied to reduce the influence of individual high values, which can have a disproportionate effect on the model. The (negative) binomial models were checked for overdispersion using the package *blemco*. Overdispersion has only to be checked for in distributions which do not have variance parameters (e.g. poisson or binomial models).

The amount of vocal practice

To assess the influence of maternal behaviors on bout duration we counted all interactive maternal behaviors (Table S1) during a babbling bout. We analyzed a total of 3221 maternal behaviors produced by mothers during 186 babbling bouts (Costa Rica 2016: N=123 babbling

bouts, ten mother-pup dyads, Panama 2017: N=63 babbling bouts, nine mother-pup dyads). To assess the maternal influence on bout duration we calculated a GLMM (Gamma family, with log link), with bout duration as response variable, maternal behaviors (log transformed and subsequently standardized), pup age (standardized) and the number of tutors (standardized) as fixed factors and pup ID as random factor (based on the datasets of 2016 and 2017, N=19 mother-pup dyads, N=186 babbling bouts). To assess the influence of the social environment on babbling phase duration (in days), we calculated a LMER. An LMER was chosen because it fits the distribution of the raw data better than a GLMM. A GLMM with Gamma-family would not adequately represent such a symmetrical distribution. The response variable was babbling phase duration (in days). The fixed effects were the maternal activity score (standardized) and the number of tutors (standardized) (N=19 pups, N=186 babbling bouts), and colony ID was included as random factor. To investigate potential sex differences in bout duration, babbling phase duration and maternal interactive behaviors we calculated Mann-Whitney-U tests (N=131 babbling bouts, N=14 mothers; using the r-package "coin"). Not all pups were sexed; therefore, we calculated sex differences only for 14 out of the 19 pups.

The pups´ final repertoire size

We calculated three separate Spearman rank correlation tests (*cor.test* function, method "spearman") to investigate the relationship between the female activity score, respectively the number of tutors, the mean bout duration and the pups` final repertoire size ((N=10 mother-pup dyads from 2016, repertoire size see Fernandez et al., 2021)).

The properties of the learned buzz syllables

To assess the maternal influence on the number of buzz syllables in a babbling bout we calculated a GLMM (negative binomial model) with number of buzz syllables as response variable (without syllable type B1 because it was not produced by each pup), maternal behaviors (log transformed and subsequently standardized), pup age (standardized) and the

number of tutors (standardized) as fixed factors and pup ID as random factor (based on the dataset of 2016, N=10 mother-pup dyads, N=90 babbling bouts, N=8837 buzz syllables). A negative binomial model was chosen over a poisson model because it better fitted the data structure. To assess the influence of the social environment on the buzz syllable versatility within a babbling bout, we calculated the Shannon-wiener Index (*vegan* package) for each babbling bout (N=10 mother-pup dyads, N=90 babbling bouts, N=9201 buzz syllables). This index was the response variable, whereas the maternal behaviors (log transformed and subsequently standardized), pup age (standardized) and the number of tutors (standardized) were the fixed factors with pup ID as random factor (lmer model). An LMER was chosen because a Gamma-family cannot deal with the value 0. Furthermore, we calculated the effect of the social environment on the presence of mature song syllable types (focusing on the most common syllable type B2, Table S3) in babbling bouts. We used the function *cbind(Mature syllables, precursor syllables)* to analyze the effect of the predictor variables on the presence of mature syllables. The maternal behavior (log transformed and subsequently standardized), the pup age (standardized) and the number of tutors (standardized) were the fixed factors, including pup ID and an observation level (because this model showed overdispersion) as random factors (GLMM, binomial, N=90 babbling bouts, N=4847 buzz syllables).

The pups´ acoustic movement during ontogeny

For each syllable part of the B2 syllable (i.e. buzz part and tonal part) we separately obtained derived parameters (i.e. frequency and entropy curvatures) by performing four separate principal component analyses (PCA) with varimax rotation on the parameters mentioned above (see acoustic data analysis) which reduced multicollinearity between original acoustic parameters. Furthermore, we decided to calculate separate PCAs for the colonies that were acoustically separated (i.e. the bats could not hear each other). PCA frequency curvature: 36 frequency parameters, i.e. peak, maximum, and minimum frequency, and bandwidth at 9 points evenly distributed over each syllable part; PCA entropy curvature: 18 entropy and

harmonics-to-noise ratio parameters, i.e. entropy and harmonics-to-noise ratio at 9 points evenly distributed over each syllable part. All PCAs fulfilled Kaiser-Meyer-Olkin (KMO) and Bartlett´s test criteria, ensuring the appropriateness of our data for PCAs.

Subsequently, we conducted two separate discriminant function analyses (DFAs) for the two colonies, respectively. The DFAs were adjusted to the unequal number of analyzed syllable parts per pup and male by computing group sizes based on prior probabilities. We split the dataset into two phases, namely an early and late babbling phase (i.e. babbling phase duration divided by two). We calculated the multidimensional signal space with the acoustic data from the early babbling phase (i.e. entire babbling phase divided by two, first half: early phase, second half: late phase, Fig. S4B) and from the adult males and subsequently plotted the acoustic data from the late babbling phase into this signal space. This approach is like a test and training approach where a multidimensional space is calculated based on the training data and subsequently, the test data is analyzed based on the training set. We chose to integrate syllables from adult males to generate the signal space (DFA 1 = early babbling phase), since it is known that pups become more similar to their tutors during ontogeny(Knörnschild et al., 2010). Instead of including different adult males we summarized the syllables as a representative "generic" male, as we had unequal (and sometimes insufficient) number of syllables from individual males. For each pup we had the same number of syllable trains per babbling phase (range: 7-20 trains per babbling phase, N=8 pups). For both DFAs, we included the same parameters, namely, duration of the buzz part, duration of the tonal part, mean peak, minimum and maximum frequency of the buzzed respectively tonal part, and two derived parameters of the buzzed respectively tonal part, describing entropy curvatures (all with eigenvalues >1). Furthermore, we included the first two linear frequency cepstral coefficients (LFCC1, LFCC2). In total, we included 12 variables in both separate DFAs.

To assess the acoustic movement pups underwent during their babbling phase we calculated the Euclidean distances between individual centroids of babbling phase 1 and phase 2 based

on the DFAs. For each pup, we calculated the distance between itself in babbling phase 1 and babbling phase 2 (range distances: 0.46-2.96, with higher numbers depicting larger acoustic movement during ontogeny). Subsequently, we correlated the individual pup movements with the female activity index.

Acknowledgements

We thank the Smithsonian Tropical Research Institute and the Natural Reserve Curú for excellent research conditions. We thank S. Ripperger for valuable comments on the manuscript.

Funding sources

This work was supported by grants from the Elsa-Neumann Foundation to A.A.F. and from the European Research Council under the European Union's Horizon 2020 Programme (2014–2020)/ERC GA 804352 to M.K.

Author contributions

A.A.F. designed the study, A.A.F and N. L. S. collected the data, A.A.F., N. L. S. and S. C. F. analyzed the data, A.A.F. performed statistical analyses, including consultancy from the statistical office of Dr. F. Korner-Nievergelt, A.A.F. wrote the manuscript, A. A. F. & M. K. revised the manuscript, M.K. supervised the study.

Declaration of interests

The authors declare no competing interests.

References

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2017). Multiple Coordination Patterns in Infant and Adult Vocalizations. *Infancy*, *22*(4), 514-539. https://doi.org/10.1111/infa.12165

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. https://doi.org/10.1111/desc.12641

Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, *49*(3-4), 227-266. https://doi.org/10.1163/156853974X00534

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. https://doi.org/10.1007/s00265-008-0677-2

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. https://doi.org/10.1007/s00265-004-0768-7

Bistere, L., Gomez-Guzman, C. M., Xiong, Y., & Vallentin, D. (2024). Female calls promote song learning in male juvenile zebra finches. *Nature Communications*, *15*(1), 8938. https://doi.org/10.1038/s41467-024-53251-z

Bradbury, J. W., & Emmons, L. H. (1974). Social organization of some Trinidad bats I. Emballonuridae. *Zeitschrift für Tierpsychologie*, *36*(1-5), 137-183. https://doi.org/10.1111/j.1439-0310.1974.tb02130.x

Brainard, M. S., & Doupe, A. J. (2000). Auditory feedback in learning and maintenance of vocal behaviour. *Nature Reviews Neuroscience*, *1*(1), 31-40. https://doi.org/10.1038/35036205

Briefer, E. F., & McElligott, A. G. (2012). Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Animal Behaviour*, *83*(4), 991-1000. https://doi.org/10.1016/j.anbehav.2012.01.020

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. https://doi.org/10.1016/j.cub.2018.12.026

Carouso-Peck, S., Goldstein, M. H., & Fitch, W. T. (2021). The many functions of vocal learning. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200235. https://doi.org/10.1098/rstb.2020.0235

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. https://doi.org/doi:10.1073/pnas.1522306113

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

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

Fernandez, A. A., Burchardt, L. S., Nagy, M., & Knörnschild, M. (2021). Babbling in a vocal learning bat resembles human infant babbling. *Science*, *373*(6557), 923-926. https://doi.org/doi:10.1126/science.abf9279

Fernandez, A. A., & Knörnschild, M. (2020). Pup Directed Vocalizations of Adult Females and Males in a Vocal Learning Bat. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.00265

García, N. C. (2019). What have we recently learned about song learning and social interactions? *Behavioral Ecology*, *30*(5), 1193-1195. https://doi.org/10.1093/beheco/arz098

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

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. https://doi.org/doi:10.1073/pnas.1332441100

Goldstein, M. H., & Schwade, J. A. (2008). Social Feedback to Infants' Babbling Facilitates Rapid Phonological Learning. *Psychological Science*, *19*(5), 515-523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Developmental Psychology*, *47*(1), 67. https://doi.org/10.1037/a0020722

Gultekin, Y. B., & Hage, S. R. (2017). Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nature Communications*, *8*(1), 14046. https://doi.org/10.1038/ncomms14046

Janik, V. M., & Knörnschild, M. (2021). Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200244. https://doi.org/10.1098/rstb.2020.0244

Janik, V. M., & Slater, P. J. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, *26*(1), 59-100.

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften*, *93*(9), 451-454. https://doi.org/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. https://doi.org/10.1038/s41598-017-14434-5

Knörnschild, M., Eckenweber, M., Fernandez, A. A., & Nagy, M. (2016). Sexually Selected Vocalizations of Neotropical Bats. In J. Ortega (Ed.), *Sociality in Bats* (pp. 179-195). Springer International Publishing. https://doi.org/10.1007/978-3-319-38953-0_8

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. https://doi.org/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. https://doi.org/10.1016/j.anbehav.2008.05.018

Kuhl, P. K. (2007). Is speech learning ‘gated’ by the social brain? *Developmental Science*, *10*(1), 110-120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Marler, P. (1970). Birdsong and Speech Development - Could There Be Parallels. *American Scientist*, *58*(6), 669-673. http://www.jstor.org/stable/27829317

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. https://doi.org/10.1098/rspb.2007.1008

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*(3), 597-606. https://doi.org/10.1093/beheco/ars003

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., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science*, *4*(1), 49-60. https://doi.org/10.1111/1467-7687.00148

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. https://doi.org/10.1111/tops.12198

Strauss, M., von Helversen, O., & Knörnschild, M. (2010a). The ontogeny of courtship behaviours in bat pups (*Saccopteryx bilineata*). *Behaviour*, *147*(5/6), 661-676. http://www.jstor.org/stable/27822143

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

Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical*

Transactions of the Royal Society B Biological Sciences, *371*(1693), 20150370. https://doi.org/10.1098/rstb.2015.0370

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, *349*(6249), 734-738. https://doi.org/doi:10.1126/science.aab1058

Takahashi, D. Y., Liao, D. A., & Ghazanfar, A. A. (2017). Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Current Biology*, *27*(12), 1844-1852.e1846. https://doi.org/10.1016/j.cub.2017.05.004

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

Ter Haar, S. M., Fernandez, A. A., Gratier, M., Knörnschild, M., Levelt, C., Moore, R. K., Vellema, M., Wang, X., & Oller, D. K. (2021). Cross-species parallels in babbling: animals and algorithms. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200239. https://doi.org/10.1098/rstb.2020.0239

Theofanopoulou, C., Boeckx, C., & Jarvis, E. D. (2017). A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1861), 20170988. https://doi.org/10.1098/rspb.2017.0988

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. https://doi.org/10.1644/08-mamm-s-060.1

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A Social Feedback Loop for Speech Development and Its Reduction in Autism. *Psychological Science*, *25*(7), 1314-1324. https://doi.org/10.1177/0956797614531023

West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, *334*(6179), 244-246. https://doi.org/10.1038/334244a0

Yancey, F. D., Goetze, J. R., & Jones, C. (1998). Saccopteryx bilineata. *Mammalian Species*(581), 1-5. https://doi.org/10.2307/3504459