Science Advances

www.advances.sciencemag.org/cgi/content/full/1/2/e1500019/DC1

Supplementary Materials for

Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats

Yosef Prat, Mor Taub, Yossi Yovel

Published 27 March 2015, *Sci. Adv.* **1**, e1500019 (2015)
DOI: 10.1126/sciadv.1500019

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.advances.sciencemag.org/cgi/content/full/1/2/e1500019/DC1)

Methods:

Animal retrieval and care. All adult bats (*Rousettus aegyptiacus*), females in their late pregnancy and males were caught in a natural roost near Herzliya, Israel. The roost is inhabited by a colony of 5,000 to 10,000 bats. The bats were kept in acoustic chambers, large enough to allow flight (Fig. S1), and fed with a variety of local fruit. Pups were separated from their mothers, and joined together, after the last pup was observed feeding on fruit by itself. All experiments were reviewed and approved by the Animal Care Committee of Tel Aviv University (Number L-13-016). The use of bats was approved by the Israeli National Park Authority.

Experimental design. Bat pups were reared together with their mothers until weaning. The members of the control group were raised in a colony chamber, which accommodated 5 females with their pups and one male (to more closely simulate the natural vocal diversity). All pups were born in the chambers and were monitored (video and audio) from birth. As soon as all 5 control pups were weaned, i.e. could fly and feed on fruit by themselves, they were moved into a single pups-only chamber, in which they were continuously monitored. In the isolated group each pup was reared only with its mother in a private chamber, which was also monitored. As in the control group, when the 5 isolated pups were weaned they were grouped into a single pups-only chamber. The recording continued for another 145 days, after which the pup groups were mixed and monitored for another 40 days. We could not retrieve audio recordings from one of the isolation chambers due to a technical issue, thus this pup was only recorded from the group assembly day onwards (covering most of the experiment duration). *R. aegyptiacus* bats do not emit social calls while kept alone, and our observations revealed that this is also the case when a mother is housed only with her own young pup (as in the isolation chambers). We manually scrutinized all of the recordings from the isolation chambers and found practically no social calls that were emitted by the mothers (only isolation calls and unripe calls emitted by the pups). In the four recorded isolation chambers we found 8, 1, 52, and 18 calls, whose source could not be definitely verified as the pup, although they were all emitted when the pups were old enough to produce such calls (when compared with the control pups). As a comparison, in the same period more than 100,000 adult calls were recorded in the control colony. The pups of the isolated group were 14 days younger, on average, than the control pups on the assemblage day of their group. Although this is a minor difference relative to the duration of the different stages of the experiment, we still verified that there was neither a significant physiological difference between the groups, nor a correlation between the physiological state and the measured acoustic parameters. The correlation was tested between the pups' weight (measured at day 94 after groups' assemblage, when pups were around the age of 6 months) and similarity to adults, using separate Spearmen-correlation tests for both groups and for all pups together (all tests produced p-value > 0.15). Furthermore, we emphasize that the isolated pups emitted plenty of calls with a fundamental frequency which is as low as in those emitted by the control pups (Fig. S3), indicating that the isolates had no physiological barrier in the production of the calls. The experiment was repeated twice (May – December 2012, and September 2012 – June 2013). Due to a low survival rate in the first experiment (only two pups of each group survived) the focus of this paper is the second experiment. A third experiment included playback of adult calls (June – October 2014). This experiment included one group which was similar in all aspects (except for the played vocalizations) to the isolated group in the previous experiment. During their isolation (when each pup was housed alone with its mother) a playback was played in their chambers. The playback consisted of sequences of low-frequency adult vocalizations which were recorded in previous experiments. 79 sequences of average fundamental frequency lower than 350 Hz were randomly chosen from our database (the average adult call fundamental is 700 Hz). The sequences were played in a random order. Playbacks were played continuously day and night, with intervals of random duration, in a way that kept their occurrence along the day in concordance with the distribution of the adult vocalizations in the control group, i.e., many calls at dawn and dusk, and less during the day than during the night. However, playbacks were played around 5 times more frequently than calls emitted in the 6-adults control colony, in order to enhance their hypothesized effect. Three pups were weaned, grouped (at the average age of 73 days), and monitored for one month after their group assembly.

Experimental chambers. The chambers were continuously monitored with IR-sensitive cameras and omnidirectional electret ultrasound microphone (Avisoft-Bioacoustics Knowles FG-O). The chambers were sealed to provide acoustic isolation and their walls were covered with foam to diminish echoes. To assure that no bat vocalization could be heard in the isolation chambers, we played two intense (110dB SPL re. 20µPa at 10cm) signals of 9 kHz (mean bat call's peak frequency) and 16 kHz (mean bat call's spectral centroid) in the colony chamber. We verified that the sounds arriving in the isolation chamber were attenuated by at least 75dB to below the noise floor recorded by our sensitive microphone (35dB SPL, Condenser ultrasound microphone Avisoft-Bioacoustics CM16/CMPA), which is similar to experiments conducted in songbirds(*16*). Playbacks were performed with a wideband speaker (Avisoft-Bioacoustics ultrasonic dynamic speaker ScanSpeak) connected to a D/A converter (Avisoft-Bioacoustics UltraSoundGate player 116). The light regime in the chambers reflected the natural conditions.

Data preprocessing. Audio recordings were conducted using Avisoft-Bioacoustics UltraSoundGate 1216H A/D converter with a sampling rate of 250 kHz. Raw audio files were segmented automatically into social calls which are separated by silence of at least 4 ms. To this end, an SVM-based classifier was used in a running-window over the audio file. The classifier was trained (based on a training set of 1080 manually classified segments) to classify time signals into 3 classes of acoustically dissimilar sounds: adult-like social calls, isolation-like calls, and background (including external and internal noises, echolocation and silence). Time points which were marked as non-background (i.e., classes 1-2) were grouped and selected as valid segments if they were at least 20 ms long. A post-processing classifier was trained to separate voiced segments from any (wrongly) identified background ones, and filtered out any remnant noises. This segmentation tool was written in Matlab 8, ad-hoc for this purpose. Its validity was assured by comparing its performance to that of an expert: recordings containing ~2000 classified segments were randomly sampled and validated by the expert, proving a sensitivity of 99.75%. Another 2000 segments were randomly sampled and assessed by an expert: only 2 were non-voiced (1 noise and 1 echolocation call), giving specificity of 99.9%. Video was synchronized with the audio recording, resulting in a short video footage accompanying each audio recording. Videos were then analyzed by trained students, which identified the circumstances of each call (source, addressee, and behavioral context). All in all 942,549 adult and pup calls were included in the current analysis, out of which 228,600 were annotated. In order to exclude behavioral differences between the groups, 5 behavioral contexts were also analyzed separately (Fig. S6).

Data analysis. For each call, a set of 10 acoustic features was extracted, including: fundamental frequency, first and second formants, Shannon entropy of the power spectrum, Wiener entropy, spectral centroid, frequency of 0.95 roll-off, zero crossing rate, energy entropy, and duration of the call (see details in 'Acoustic features measurement' below). The average feature for each call was measured. In order to calculate the adult-dissimilarity and the dispersion of the calls in the feature-space (see below) each of the features was first normalized across the entire dataset to yield its z-score. The difference from the adults (Fig. 1C) was measured as the mean Euclidean distance between each pup calls at every age and the closest adult calls. To this end the distance of each pup call was calculated from all adult calls, and only the 10 shortest distances were considered and averaged. The incorporation of only the closest adult calls (rather than reporting the average distance from all calls) minimizes biases which may occur if only a subset of the adult repertoire is used by the pup. The dispersion of the calls, representing the repertoire diversity, was measured as the median absolute 2-dimenssional deviation (from the median) of the data in two normalized features: the fundamental frequency and the Wiener entropy. The fundamental frequency was calculated using the YIN algorithm(*22*). To assess the statistical significance of the differences between the isolated and control groups, in their developmental trends, we used mixed model ANOVA. For both similarity to adults and fundamental frequency the interaction between the age and the group (control vs. isolated) was significant ($p<0.01$) as expected, and the result for the between-subject (i.e. group membership) analysis is presented in the text. This test was applied to the vocalizations recorded since the formation of the two pup groups. Acoustic analyses are presented in bins of 20 days along the ontogeny of the pups. The processing and the analysis of the data were performed using Matlab 8. The statistical analysis was performed using SPSS.

Acoustic features measurement. In order to measure the acoustic features of segmented calls, a sliding window was used, and the features calculated over each window were averaged. This averaging method reduces the amount of acoustic information representing the call. However, it only means that the differences we find could probably have been sharpened in a more refined analysis. Importantly, this process was identical for all calls, at every age, and in every group. The averaging was conducted for the sake of computability of comparisons between features of hundreds of thousands of calls. The duration of the sliding window was 20 ms and the overlap between consecutive windows was 19 ms.

The acoustic features that were calculated:

1. Fundamental frequency (F_0): This is the pitch of the call. It was calculated using the YIN algorithm (*22*).
2. (and 3.) 1st and 2nd formants (F_1, F_2): The two highest energy peaks in the power spectrum. Calculated using LPC method.
4. Shannon entropy of the power spectrum: The entropy is defined as –

$$-\sum_{i=1}^{n} p_i \log(p_i)$$

   where p_i is the energy concentrated in the frequencies of the i'th bin, and n = 256 bins.

5. Wiener entropy: This is the spectral flatness. It is defined as the ratio of the geometric mean to the arithmetic mean of the power spectrum, i.e.:

$$\frac{\sqrt[n]{\prod_{i=1}^{n} p_i}}{\frac{1}{n}\sum_{i=1}^{n} p_i}$$

   where p_i is the energy concentrated in the frequencies of the i'th bin, and n = 256 bins.
6. Spectral centroid: the power spectrum center of mass –

$$\sum_{i=1}^{n} f_i p_i$$

   where f_i is the frequency of the i'th bin and p_i is the (normalized) amount of energy in that bin.
7. 95% roll-off frequency: The frequency under which reside most (95%) of the energy in the power spectrum.
8. Zero crossing rate: amount of sign alterations per second in the raw audio data.
9. Energy entropy: the Shannon entropy of the energy distribution in time –

$$-\sum_{i=1}^{T} e_i \log(e_i)$$

   where e_i is the (normalized) amount of energy in bin i, T = 20 time bins.
10. The duration of the call in seconds.

The power spectrum was calculated by FFT over the entire window (5000 samples) with 8192 frequency points, using hamming window.

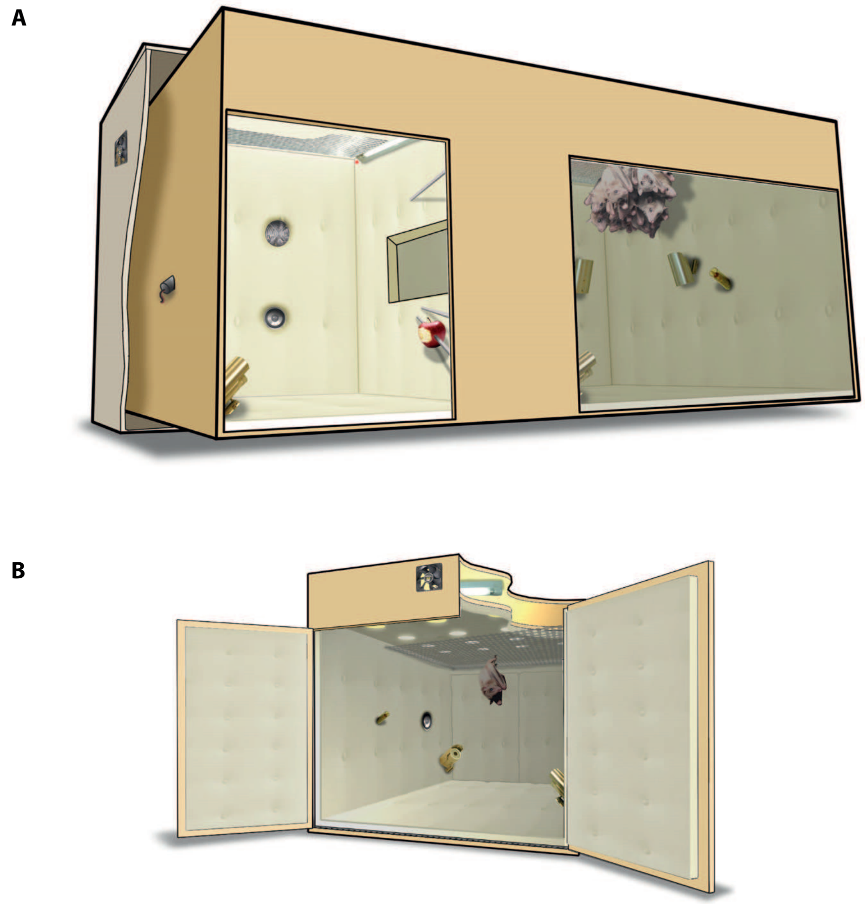

Fig. S1. Scheme of experimental acoustic chambers. These chambers were used for the rearing and recording of bats. The pups that participated in this study were born inside these chambers and only left them at the end of the experiment. Both chamber types were big enough to allow flight, and were continuously monitored with audio and video recording devices. (**A**) Colony chamber (length: 190cm; width: 90cm; height: 82cm) – this chamber was used to rear the control pups with their mothers and a male in the first phase of the experiment (pre-weaning phase). In the second phase – after the two pup-colonies were assembled, both pup groups were reared in chambers of this type. (**B**) Isolation chamber (length: 120cm; width: 70cm; height: 60cm) – this type of chamber was used to rear each of the isolated pups. Each isolate was reared in a single chamber together with its mother until it was weaned.

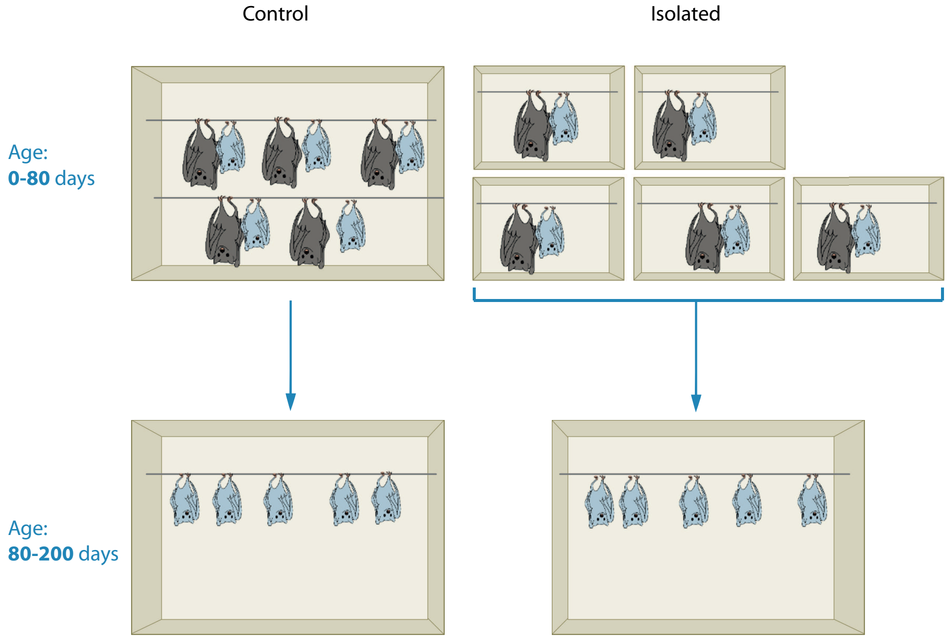

Fig. S2. Experiment outline. The experimental pup groups participated in this study. The pups of the control group (left) were reared all together with their mothers and a male in one chamber. After weaning, the pups were moved into a colony chamber of their own and the adults were excluded. The isolated pups were reared with their mothers, each pair of mother and pup in a separate chamber. After weaning the pups were grouped into a colony chamber of their own and the mothers were excluded. The manipulated group was similar to the isolated group, but playback was played in the isolation chambers.

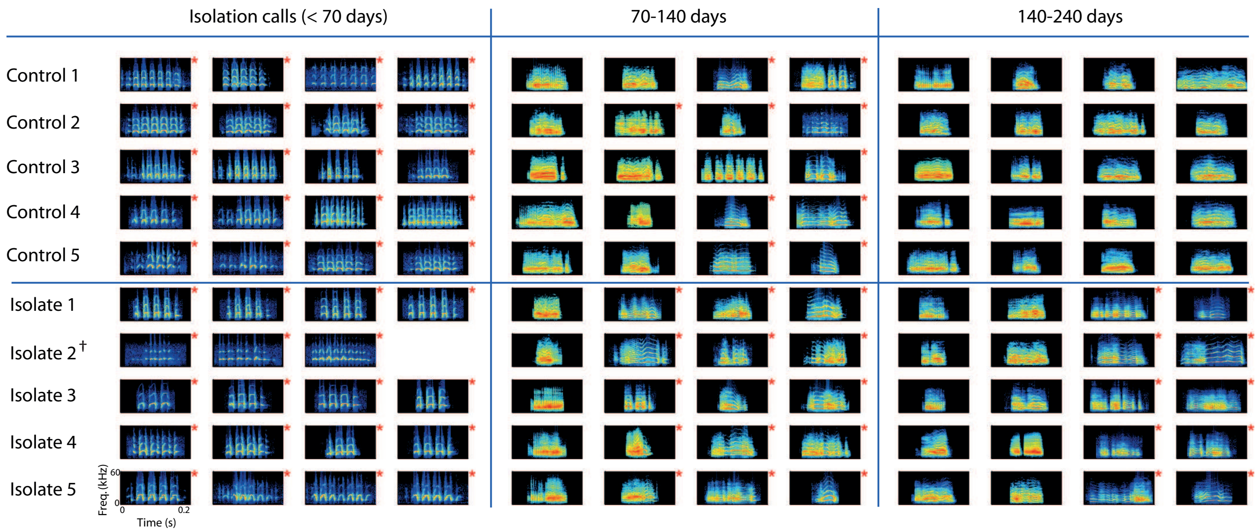

Fig. S3. The developing repertoire of bat calls. Examples of calls emitted by each pup during its vocal ontogeny. The left column includes only isolation calls, emitted until the age of around two months. During maturation (middle column) both groups produced unripe social calls, and increasingly, adult-like calls. At the age of 5-8 months (right column) the control pups produced only adult-like calls while the isolated pups continued to use unripe-calls alongside adult-like calls. All spectrograms present a time frame of 200 ms and frequencies of 0 to 60 kHz. Calls with average fundamental frequency of over 1.5 kHz are marked with a red asterisk. At later developmental stages, the occurrence of these calls was much higher in the isolation group. †Isolate pup 2 was not recorded during isolation, thus its presented isolation calls are from the age of 72 days and already contain basic modifications.

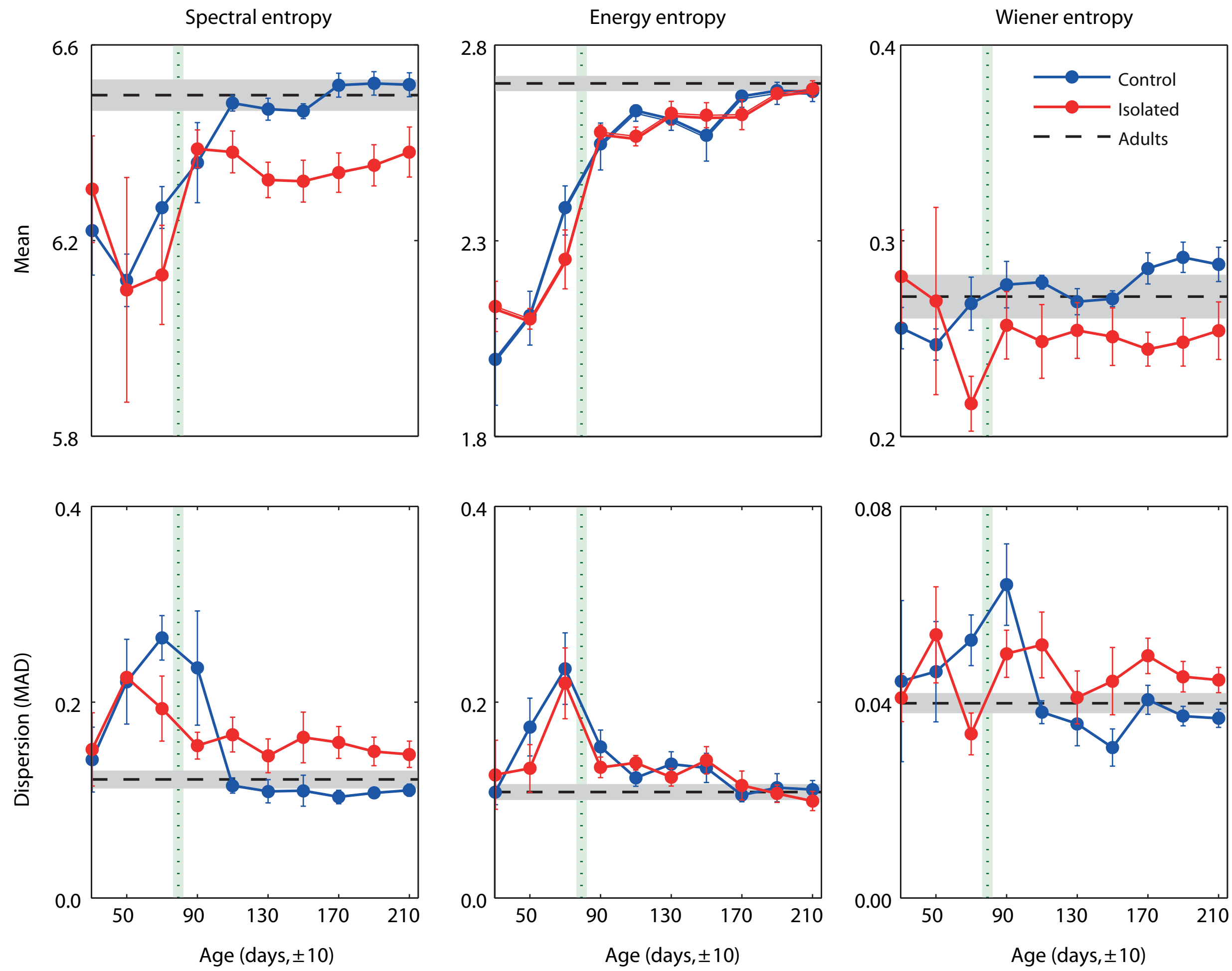

Fig. S4. Vocal ontogeny of different acoustic features. Two spectral features - spectral entropy (left) and Wiener entropy (right), and one temporal feature - energy entropy (middle), are presented (see Methods for calculation details). The plots depict the average of the five pups in each group. Upper panels show the mean of all of the calls of each pup at every age. Lower panels: the dispersion (or variability) of the entropies among the calls of each pup was computed as the median absolute deviation (MAD) of the calls at every age (see Methods). Quantities are measured in bins of 20 days. Blue – control group (n=5); red – isolated group (n=5; n=4 for the first 2 bins); dashed black – the same measurement when applied to adults (n=10); green dotted line – average of pups' age on the assemblage day of pups-only groups. Error bars and shades - SEM. Note that in the spectral entropy control pups converge to the adult baseline while the isolated pups lag behind (similar to the case of the fundamental frequency, Fig. 1D). A similar, but less profound, effect appears in the Wiener entropy. Also note that in both cases the dispersion stays higher in the isolated group at a later age, which means that isolates keep a highly diverse repertoire through their maturation. In contrast, the experimental condition had no effect on the energy entropy which developed simultaneously into adult-like in both groups.

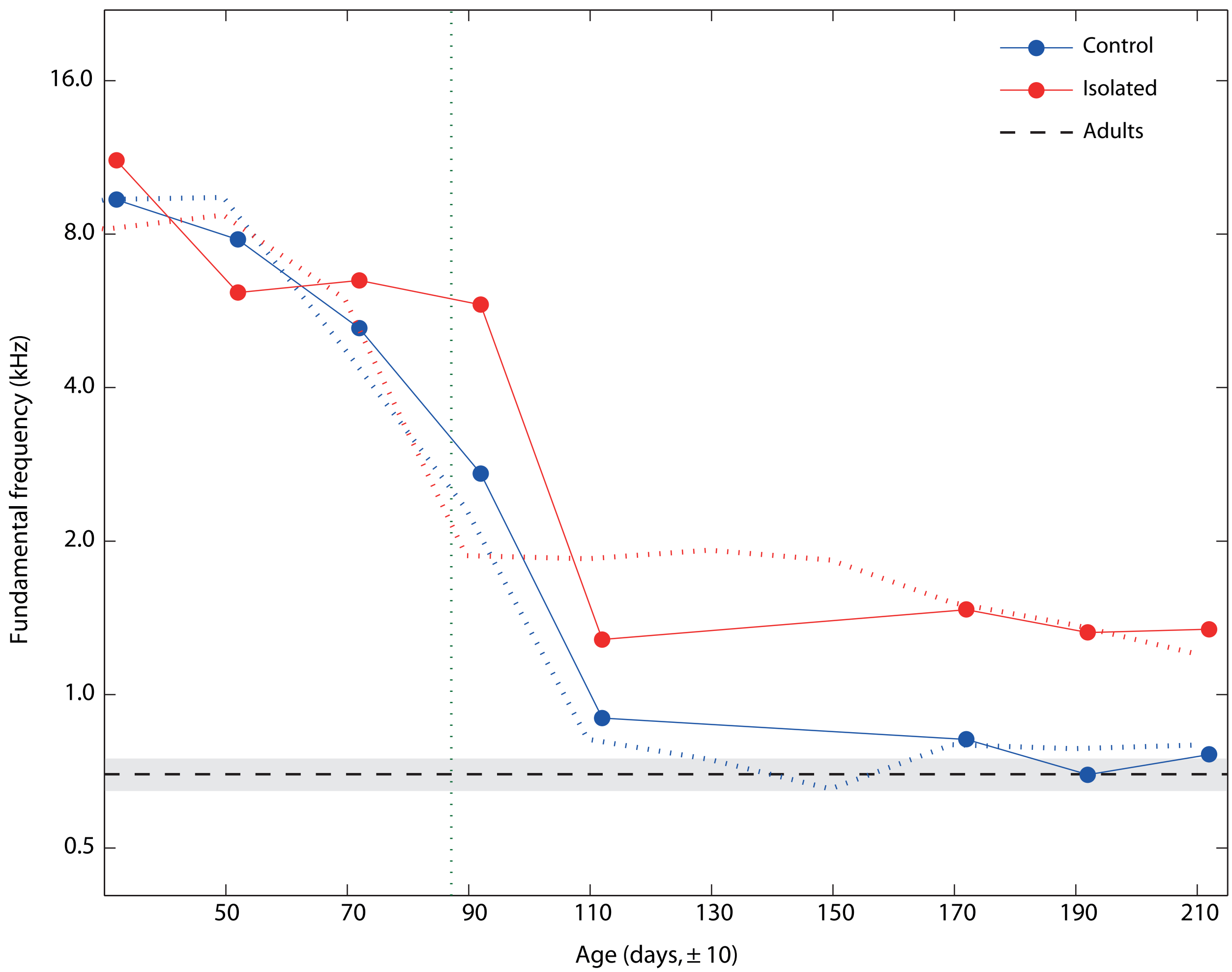

Fig. S5. The development of fundamental frequency among pups of the first experiment. The average fundamental frequency of the pup calls at different ages (measured in bins of 20 days). Blue – control group (*n*=2); red – isolated group (*n*=2); dashed black – the adult average (*n*=10); green dotted line – average pups' age on the assemblage day of pups-only groups. This data shows an identical pattern to the one observed in the second experiment which is described in the main text. However, due to high mortality rate we were left with only 2 pups in each group, and hence the focus of our study is the second (full-data) experiment, and these results are brought as an independent support. The blue and red dotted lines are the respective control and isolation lines from the full-data experiment (as presented in Fig. 1D).

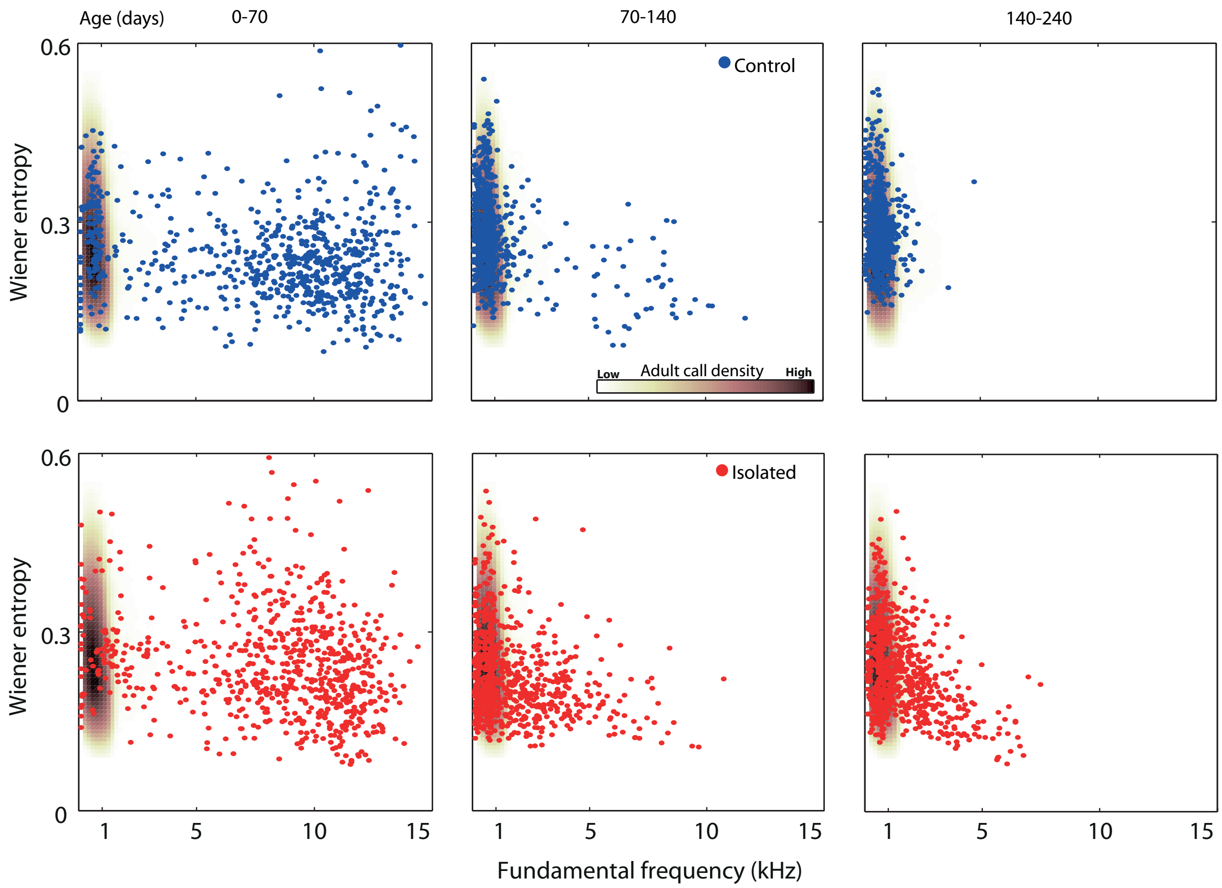

Fig. S6. Development of vocal diversity. Three developmental stages of the 5 pups from the control group (upper panels, blue) and the 5 pups from the isolated group (lower panels, red), presented as a scatter plot of two acoustic features. Brown shades indicate these features' distribution among adult calls. A random sample of 750 calls is presented in each panel for convenience of presentation and comparison.

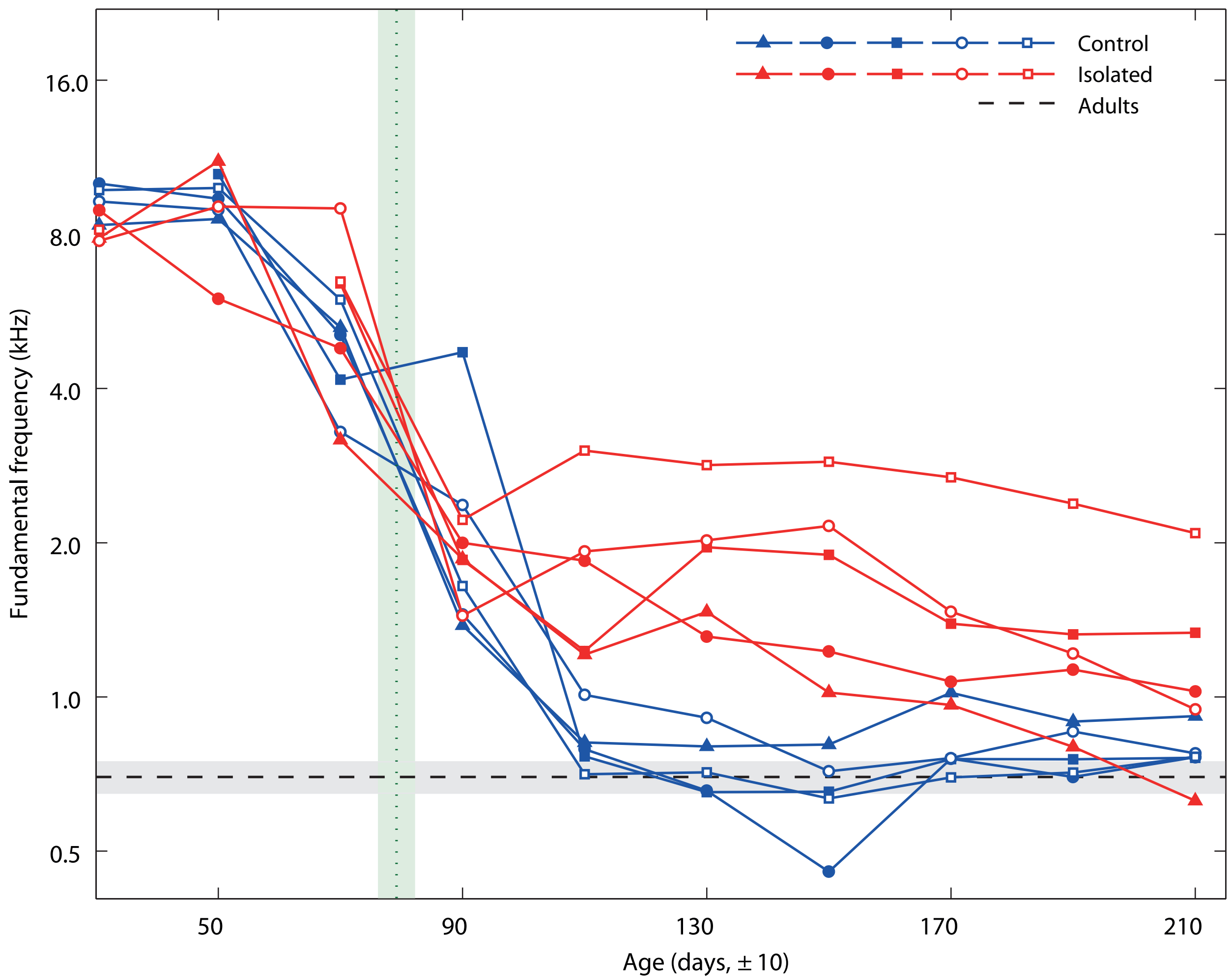

Fig. S7. The development of fundamental frequency of individual pups. Blue – control group; red – isolated group; dashed black – the adult average; green dotted line – average pups' age on the assemblage day of pups-only groups. Note how the isolated pups use a higher average fundamental than the control pups, which in turn use adult-like fundamentals. Even at the age of 7 months 4 isolated pups used fundamental frequencies which are higher than those used by the control pups, and one isolate used frequencies which were lower on average than all of those used by the control pups, pointing to the high variability among the isolates which is not present in the control group.

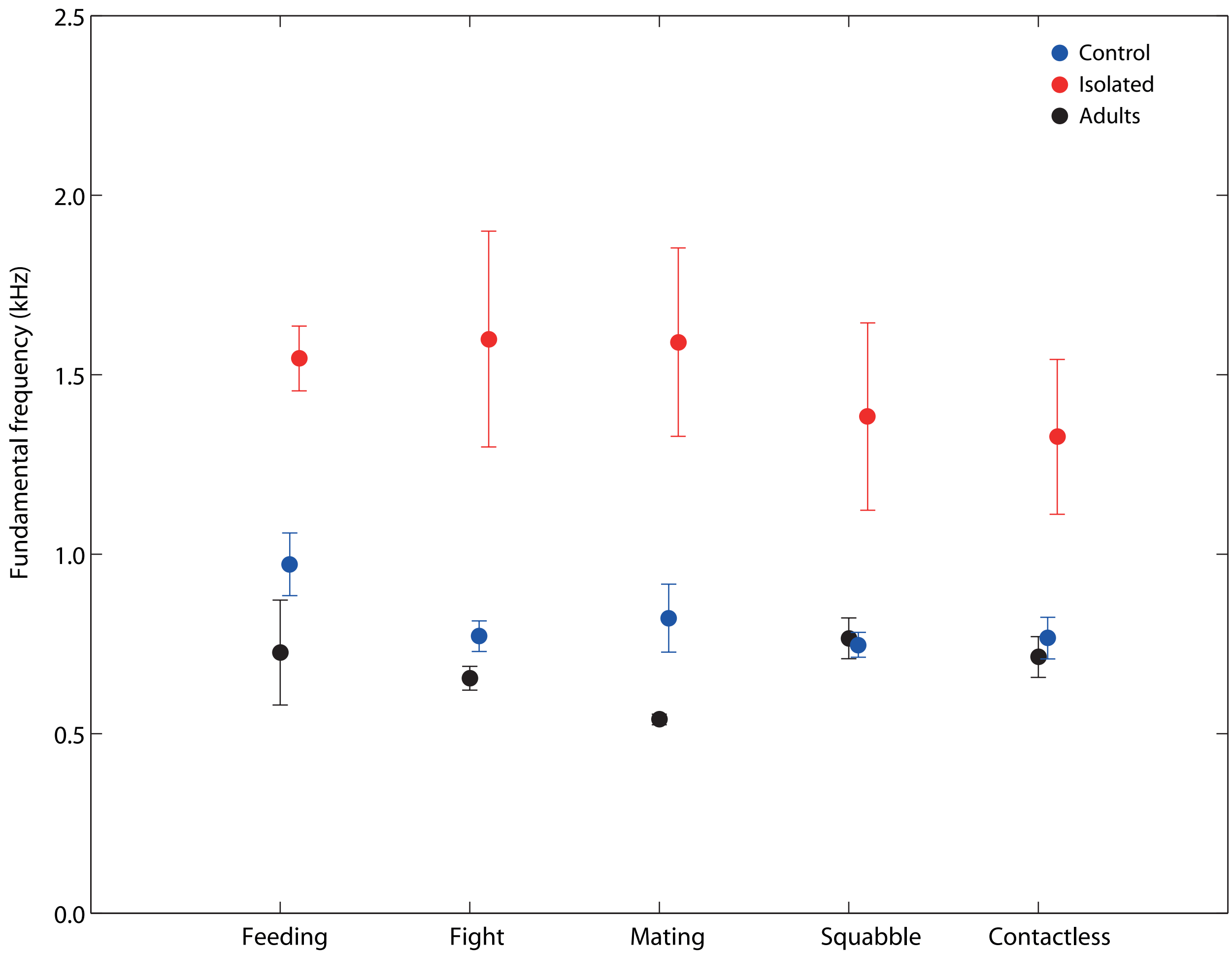

Fig. S8. The fundamental frequency of calls produced in different behavioral contexts. The average of the fundamental frequency used by each pup at the age of 120-220 days, in each behavioral context. Blue – control group (*n*=5); red – isolated group (*n*=5); black – the adult average (*n*=10). Error bars indicate the standard error of the mean. Behavioral contexts (left to right): Feeding related interactions, typical intense face-to-face fights, copulations (for pups: mating-play; note that the pups in the experiment were prepubertal, which might explain the difference between the control and adult groups), squabbles over positions in the day-sleeping cluster, other pairwise interactions lacking physical contact.

Movie S1. A typical intense face-to-face fight as displayed by the control group. A characteristic aggressive interaction which is common to all bats of this species. This movie shows such interaction between two pups of the control group. An attentive listener will notice the lower fundamental of the vocalizations, compared to that of the isolated pups (Movie 2). The spotlight marks the interacting pair.

Movie S2. A typical intense face-to-face fight as displayed by the isolation group. A characteristic aggressive interaction which is common to all bats of this species. This movie shows such interaction between two pups of the isolation group. An attentive listener will notice the higher fundamental of the vocalizations, compared to that of the control pups (Movie 1). The spotlight marks the interacting pair.