Wang et al. *Zool. Res.* 2026, 47(4): 1124−1133
https://doi.org/10.24272/j.issn.2095-8137.2025.350
https://cstr.cn/32248.14.zr.2025.350

ZR | Zoological Research

Article Open Access

Task learning drives adaptive vocal adjustments in echolocating bats

Qianyu Wang[1,#], Huan Ye[1,#], Jinhong Luo[1,*]

[1] *Key Laboratory of Pesticide & Chemical Biology of Ministry of Education*, *Hubei Key Laboratory of Genetic Regulation & Integrative Biology*, *School of Life Sciences*, *Central China Normal University*, *Wuhan*, *Hubei* 430079, *China*

ABSTRACT

Vocal production learning (VPL), a core ability underlying human speech, is rare among mammals. Apart from humans, two of the four identified VPL mammalian lineages (cetaceans and bats) are echolocators, exhibiting remarkable vocal flexibility for echolocation. It has been posited that enhanced vocal flexibility can facilitate the evolution of VPL, which operates at longer time scales. However, empirical data linking vocal flexibility to VPL is still lacking. Using a psychophysical reinforcement learning paradigm, we trained Pratt's roundleaf bats, *Hipposideros pratti*, to detect pure tones and virtual echoes, both with and without background noise. We report that bats learned to adapt their call amplitude to auditory tasks across various time scales, ranging from instantaneous to months. Surprisingly, bats modified the noise-induced call amplitude adaptation, a reflexive audio-vocal behavior known as the Lombard effect, in a task-specific manner. Moreover, over months of task learning and relearning, bats showed gradual, task-specific call amplitude adjustments that reflected their task performance. We posit that perceptual demands for echolocation efficiency over longer time scales may have facilitated the evolution of the rare VPL ability in echolocating mammals.

Keywords: Acoustic communication; Behavioral ecology; Evolution; Neuroethology; Vocal production control

INTRODUCTION

Vocal production learning (VPL), often defined as the ability of an animal to produce novel vocalization, is a core ability that enables humans to acquire both native and foreign languages, and to adapt to the local dialects (Janik & Slater, 1997; Jarvis, 2019; Vernes et al., 2021). Apart from humans, however, VPL is rare and has only been documented in four mammalian lineages: elephants, cetaceans, pinnipeds, and bats (Ancillotto et al., 2022; Fernandez et al., 2021; Janik, 2000; Janik & Knörnschild, 2021; Stansbury & Janik, 2019; Tyack, 2016). Interestingly, two of the four identified VPL non-human mammalian lineages (cetaceans and bats) are echolocating species, which use self-emitted sounds to probe the environment. One remarkable feature of the echolocation in both cetaceans and bats, on which intensive research has been carried out, is that it is highly flexible. These echolocating mammals can rapidly adapt their echolocation signals to the perceptual needs of the task at hand (Jakobsen et al., 2013; Madsen & Surlykke, 2013; Moss et al., 2011; Ratcliffe et al., 2013; Schnitzler & Denzinger, 2011; Thomas et al., 2004; Wang et al., 2025). For example, many echolocating bats modify acoustic features of the echolocation signals from searching to approaching and capturing insects (Griffin et al., 1960). Their echolocation is also highly sensitive to the habitat structure and background noise (Schnitzler & Kalko, 2001; Smotherman et al., 2021).

It has been repeatedly posited that enhanced vocal flexibility can facilitate the evolution of VPL (Jarvis, 2019; Tyack, 2016, 2019; Wirthlin et al., 2019). In this study, we formalize this idea as the Vocal Flexibility Facilitates Learning (VFFL) Hypothesis. However, VPL differs from vocal flexibility by a core feature: VPL is an essential learning process occurring on a relatively slow time scale of days, weeks, or even months. By contrast, many forms of vocal flexibility operate like a reflexive behavior on a rapid time scale, without involving a gradual learning process. A typical example of vocal flexibility is the Lombard effect, in which subjects instantly raise their call amplitude in response to background noise (Lombard, 1911). In bats, the Lombard effect is extremely rapid, taking merely 20–30 ms for its initiation (Luo et al., 2017a; Pedersen et al., 2024), emerges in infant bats of a few days old (Luo et al., 2017b), and can be induced by narrowband noise of a 1 kHz bandwidth (Liu et al., 2026). By contrast, the Lombard effect in humans appears to be mediated by cognitive processes: it varies with task engagement (Garnier et al., 2010) and can be inhibited with appropriate training (Pick et al., 1989).

In recent years, bats have received considerable interest

Received: 26 November 2025; Accepted: 22 December 2025; Online: 23 December 2025

Foundation items: This study was supported by the National Natural Science Foundation of China (32270535) and the Fundamental Research Funds for the Central Universities (JC2026TS-001) to Jinhong Luo.

[#]Authors contributed equally to this work

*Corresponding author, E-mail: jluo@ccnu.edu.cn

and are becoming an important mammalian model for VPL research (Genzel et al., 2019; Vernes & Wilkinson, 2020; Zhang et al., 2026). Compared with other VPL mammalian taxa (i.e., pinnipeds, cetaceans, and elephants), bats are more amenable to laboratory investigations at different levels from molecular underpins to brain mechanisms to behavior (Teeling et al., 2018; Vernes et al., 2022; Zhao et al., 2024). Although the number of VPL species of bats demonstrated by manipulation experiments remains limited (Esser, 1994; Knörnschild, 2014; Prat et al., 2015, 2017), there is a long list of bat species suggested to possess potential VPL based on their capacities for long-term vocal plasticity (Genzel et al., 2019; Vernes & Wilkinson, 2020; Ye & Luo, 2022). A major form of long-term vocal plasticity is the vocal convergence with conspecific bats (Boughman, 1998; Hiryu et al., 2006; Prat et al., 2017). When bats are translocated into another colony or population, new individuals can converge their vocalizations to the resident individuals. A key question about vocal convergence is whether it arises from goal-directed learning processes (as seen in classic vocal production learning) or immediate vocal adjustments driven by auditory feedback. In other words, do bats actively modify their vocalizations to reduce the acoustic differences between themselves and others (the goal)?

To test the VFFL Hypothesis, we investigated the effects of task learning on the vocal flexibility of Pratt's roundleaf bat, *Hipposideros pratti*, using a psychophysical reinforcement learning paradigm (Figure 1). The echolocation calls of Pratt's roundleaf bats typically were composed of constant-frequency (CF) and frequency-modulated (FM) components (Figure 1A). These CF-FM bats rely on the CF component for target detection and evaluation and the FM component for target ranging (Fenton et al., 2012; Riquimaroux et al., 1991; Simmons et al., 1979). They exhibit remarkable short-term vocal flexibility for echolocation signals (Luo et al., 2023; Ma et al., 2024, 2025; Xia et al., 2025). They can also adaptively modulate their call parameters over weeks, suggesting a VPL ability (Ye & Luo, 2022). In the present study, we specifically focused on call amplitude, which is a core signal parameter dynamically regulated by bats during echolocation tasks. Bats dramatically decrease call amplitude after entering a closed (cluttered) environment (Schnitzler & Kalko, 2001). Also, virtually all species of echolocating bats gradually decrease call amplitude during target approach, which helps to stabilize the received echo levels (Hiryu et al., 2007; Jakobsen et al., 2013; Luo et al., 2023; Stidsholt et al., 2020). A stronger Lombard effect on the FM component than the CF component was also found in a previous study for spontaneously vocalizing hipposiderid bats not engaged in tasks (Lu et al., 2020). We reasoned that although the difference in the Lombard effect between humans and bats or other animal species may be attributed to the advanced cognitive capacities of humans, it may also be due to a lack of research on the Lombard effect on animal subjects involved in different perceptual tasks. Thus, we hypothesized that these bats would adapt echolocation call amplitude not only to the ongoing perceptual task instantly, but also on a longer time scale. We further hypothesized that bats are more likely to show task-specific call amplitude adjustments for the CF component than the FM component in detection tasks across both quiet and noisy conditions.

We trained *H. pratti* to perform two auditory detection tasks: a tone detection (TD) task for pure tones, and an echo detection (ED) task for virtual echoes, both in the presence or absence of background noise. Theoretically, the TD and ED tasks likely require distinct amplitude adjustment strategies to optimize performance. Increasing call amplitude is adaptive in the ED task, as it enhances echo intensity. Conversely, decreasing amplitude is beneficial in the TD task, as it mitigates masking. Bats may also employ other strategies in the TD task, such as shifting peak frequency away from the tone, shortening call duration, or reducing calling rate. Furthermore, the forward masking effect from emitted calls is up to 40 dB stronger for virtual echoes than for pure tones, likely due to the structural similarity and short delay of echoes relative to the calls (Ye & Luo, 2022). Therefore, we predicted that if task learning effectively shapes vocal control, bats would increase their call amplitude when learning the ED task but decrease it when learning the TD task.

METHODS

Animals

Four adult male *H. pratti* were used in this psychophysical study. These bats were caught with a hand net during the day in a cave in Xianning County, Hubei Province, China. They were housed in custom-made metal meshed cages (40 cm long×40 cm wide×60 cm high) on shelves in a room with a regulated air temperature of approximately 24°C, a relative humidity of approximately 60%, and a reversed light regime of 12 h light: 12 h dark. Throughout the experimental period, bats received food only during experiments, and their minimum food amount on an experimental day was maintained at approximately 50% of their maximum food intake to ensure their motivation. Bats had access to drinking water *ad libitum* in the cage. During the study, bats typically completed about 50 trials daily, six days a week. All procedures followed the protocols approved by the Institutional Animal Care and Use Committee of the Central China Normal University (CCNU-IACUC202209200201).

Psychophysical setup

The psychophysical setup used to collect data was essentially the same as in a previous study (Ye & Luo, 2022). Experiments took place inside an acoustically shielded room (5 m long×4 m wide×2.2 m high), with the walls and the ceiling covered with 8-inch-thick acoustic foam, and the floor covered by nylon blankets to reduce reverberations and echoes. The setup was wrapped with velvet felt to reduce echoes. The isosceles trapezoid two-alternative forced-choice (2-AFC) setup (Figure 1B) had a 10 cm upper arm that served as a starting area for the hanging bat and two 30 cm arms that served as the decision-making area which were separated by approximately 45º. A 1-inch sized electrostatic loudspeaker (ES1, TDT, USA) and a 1/4-inch sized measurement microphone (7016, ACO, USA; or 46BF-1, GRAS, Denmark) were fixed to the end of both setup arms. The speakers and the microphones were about 15 cm and 17 cm from the bat's head, respectively (Figure 1B). An infrared (IR) camera (FDR-AX700, SONY, Japan), placed at about 3 m in front of the bat, was used to film the bat's behavior during the experiment, under the illumination of an external IR light source (3022HW, Maixiu, China). The video recording was also streamed to a smartphone in real time and was monitored by the experimenter. During the test, the bat was physically

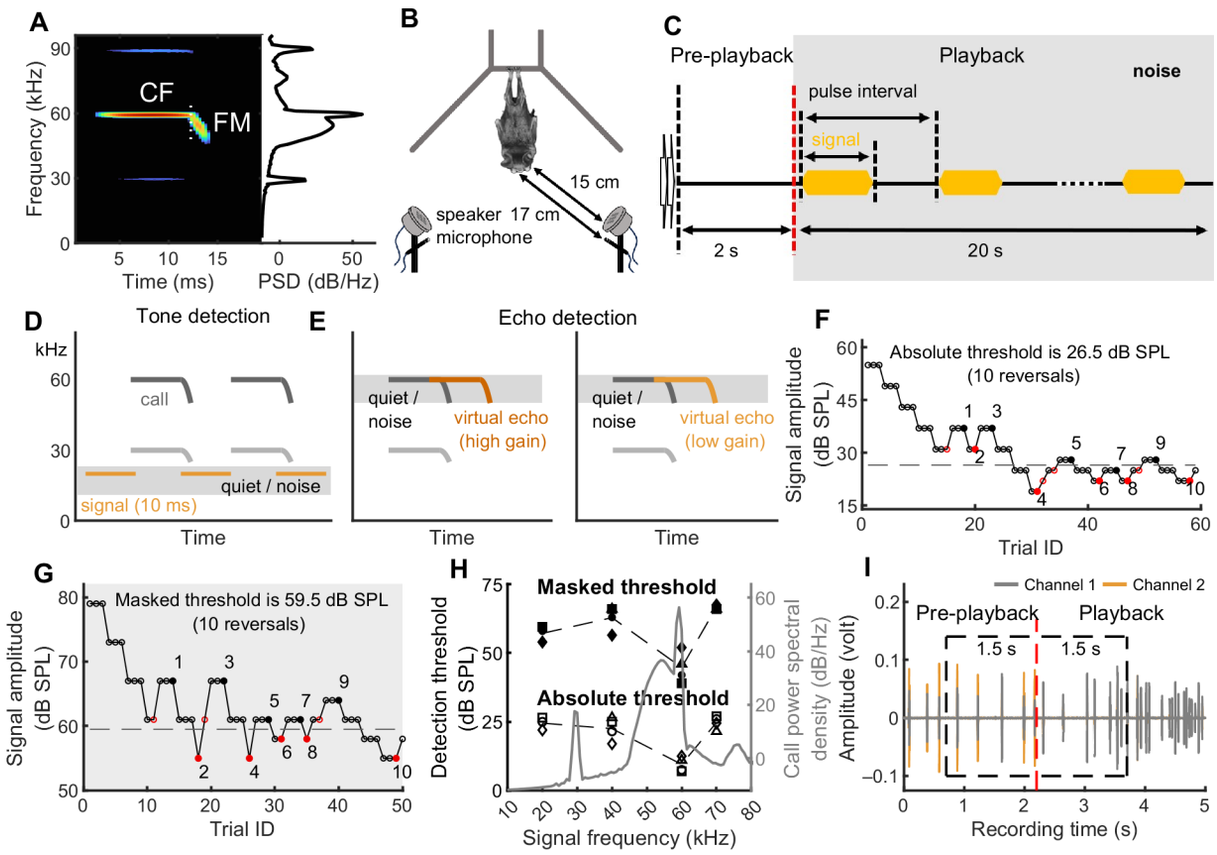

Figure 1 A psychophysical setup for studying the echolocation calls of roundleaf bats in auditory detection tasks

A: Spectrogram of a typical echolocation call for the Pratt's roundleaf bat, *Hipposideros pratti,* in the psychophysical experiment. B: Diagram of the setup used for the two-alternative forced choice (2-AFC) psychophysical experiments. C: Diagram of the procedures for presenting auditory stimuli. Each trial starts (marked by the red dashed line) with a pair of clicker sounds (0.2 s) to attract the attention of the bat, then a 2 s silence gap (pre-playback phase), and subsequently the target stimuli (referred to as the playback phase, 20 s long maximumly). For noise tasks, online-generated bandpass filtered (BFN) noise was broadcast in the background simultaneously in the playback phase (depicted as gray filling). D, E: Illustrations of the tone detection (TD) (D) tasks in quiet and noise, and echo detection (ED) (E) tasks in quiet and noise. F, G: Two example trials show the adaptive up-down method for measuring the absolute (F) and masked (G) auditory detection thresholds. SPL: sound pressure level. Black circles: correct performance; Red circles: incorrect performance; numbers associated with the circles indicate the n^{th} reversal in performance of the bat. H: Left Y-axis (data in black): the absolute and masked auditory detection thresholds of individual *H. pratti* detecting tones at four different frequencies, i.e., 20 kHz, 40 kHz, 60 kHz, and 70 kHz. Masked thresholds were measured at a noise level 30 dB higher than their absolute hearing thresholds. Right Y-axis (data in gray): the power spectral density of the echolocation calls of *H. pratti*. I: Echolocation call recordings of a *Hipposideros pratti* bat in an example trial. The dashed rectangle shows the time window for call analysis, i.e., the calls 1.5 s before and after the start of the playback window (marked by the red dashed line). The gray and orange waveforms represent echolocation calls recorded by two simultaneous microphones. For each echolocation call, the one of higher call amplitude from the two microphones was used for further analysis, and online generation of virtual echoes in the echo detection tasks. CF: constant frequency; FM: frequency-modulated.

separated from the experimenter by a black curtain hanging from the ceiling.

Sound recording and playback

Sound recording and playback were controlled through custom-written programs with the SoundMexPro toolbox (Hörzentrum Oldenburg, Germany) in MATLAB (R2019b, MathWorks, USA). Sound stimuli were generated online, converted into analog signals using a multichannel audio interface (Fireface 802, RME, Germany), and broadcast via the two speakers. The frequency response of the playback system was measured with a calibration microphone (46BF-1, GRAS, Denmark) at a distance of 15 cm (the same distance from bat to speaker) by pure tones of different frequencies from 5– 90 kHz, with a frequency resolution of 100 Hz for 55–65 kHz frequency range and a frequency resolution of 1 kHz for other frequencies. The uneven frequency response of the playback system was compensated by adjusting the digital voltage of the output signal to achieve a flat frequency response (±1 dB). The sensitivity of the calibration microphone was measured by a sound calibrator (Type 521, ACO, USA) that outputs a 1 kHz tone of 94 dB SPL level. The playback system was calibrated with white noise stimuli to design the compensatory finite impulse response (512 points) for outputting virtual echoes and bandpass-filtered white noise (BFN, see below). Echolocation calls were recorded simultaneously by both microphones at a sampling rate of 192 kHz. The bat's detection performance was logged at the end of each trial. Each trial consisted of three event windows, a pair of clicker sounds of 0.2 seconds long, followed by a two-second silence (the pre-playback phase), and then by a playback phase (up to 20 seconds) (Figure 1C). During the playback phase, pure tones or virtual echoes were broadcast. Virtual echoes were generated by filtering and attenuating the

emitted call of the bat in real time and playing it back to the bat with an 8 ms delay.

Animal training and auditory detection tasks

Animal training:: *H. pratti* was trained to learn and conduct two sound detection tasks: the tone detection (TD) task to detect pure tones, and the echo detection (ED) task to detect virtual echoes, either with or without background noise, resulting in four distinct experimental conditions for the same bat individual. The pure tones and virtual echoes are referred to as target signals hereinafter. Bats were first trained and tested for the TD tasks and then for the ED tasks. Bats learned to conduct these four tasks sequentially. For the ED tasks, bats were trained and tested for two echo levels differing by 10 dB, i.e., the –24 dB echo gain (high echo level, HEL) and –34 dB echo gain (low echo level, LEL), which correspond to echo levels of approximately 70 dB SPL and 60 dB SPL (call amplitude: Bat1, 97.0±2.6 dB SPL; Bat3, 92.5±3.6 dB SPL; Bat4, 100.9±1.4 dB SPL). Four bats participated in the two TD tasks and three also participated in the two ED tasks. One bat passed away accidentally during the learning phase of the ED task. In each experiment day, a bat typically participated in around 50 trials. When learning TD tasks, different individuals were assigned different training frequencies (Bat1, 15 kHz; Bat2, 55 kHz; Bat3, 25 kHz; Bat4, 45 kHz, Supplementary Table S1). The detailed procedures of training *H. pratti* for the TD and ED detection tasks were described in a recent study (Ye & Luo, 2022), and two bats of the present study learned tone detection tasks in that study.

Auditory detection tasks:: For each detection task, a bat was required to indicate its decision by moving toward the side where the target stimuli were broadcast. Before the target stimuli, there were a pair of clicker sounds (0.2 seconds) at the beginning, which was followed by a 2-second silence interval (Figure 1C). For the TD tasks, the target stimuli were a sequence of pure tones at a fixed frequency (Figure 1D). Each pure tone was 10 ms long and was presented at a rate of 5 Hz. Bats were tested by pure tones of four different frequencies: 20, 40, 60, and 70 kHz. The peak frequency of bat echolocation calls is around 60 kHz, at which the bats have enhanced hearing sensitivity (Figure 1H). The 20 kHz frequency falls within their social communication calls. The 40 kHz and 70 kHz frequencies are not directly related to their echolocation or social communication calls. For the ED tasks, the target stimuli were a copy of the attenuated calls emitted by the bat, which were digitally generated online, and played back to the bat at a short delay (Figure 1E). Thus, the bats can control different parameters of the calls at will. The virtual echoes, containing the intact second harmonic, were broadcast at two different echo levels differing by 10 dB, relative to the emitted calls, i.e., the HEL and LEL conditions (Figure 1E). In the ED tasks, the echo levels were directly related to the amplitude of the emitted calls and thus were under active control of the bats.

During the test, when the bat made a correct choice, the trial was terminated immediately and a reward (i.e., a piece of mealworm) was offered to the bat. When the bat made an incorrect choice or no decision within 20 seconds of the trial start, the bat received no reward. We tracked the echolocation calls of the bats during the task tests (Figure 1I) and throughout the learning period. It took individual bats 12–42 days to reach the 90% performance rate when learning the TD task. It took individual bats 47–59 days to reach the 85% performance rate when learning the ED task.

For each of the four pure tone frequencies tested in the TD task, we first measured the absolute auditory detection threshold of the bats using the adaptive up-down method (Leek, 2001), in which the amplitude was lowered by 6 dB or 3 dB (only after 30 trials) in case of three consecutive correct decisions and increased by 6 dB or 3 dB (only after 30 trials) in case of one incorrect decision (Figure 1F, G). The absolute auditory threshold during the ED task was not measured. To study the effect of masking noise on vocal behavior and auditory detection performance, we generated and broadcast bandpass-filtered white noise (BFN) of 10 kHz bandwidth online based on the pure tone frequency of the target stimuli. The BFNs were of 11–21 kHz for the 20 kHz pure tone, 35–45 kHz for the 40 kHz pure tone, 51–61 kHz for the 60 kHz pure tone, and 65–75 kHz for the 70 kHz pure tone. The 51–61 kHz BFN overlapped with the CF and the FM of the echolocation calls (Figure 1A), which thus allows us to examine the noise effects on different acoustic components. The level of each BFN was set to 30 dB above the absolute auditory detection threshold of each bat, i.e., the +30 dB BFN condition. For the 60-kHz TD task, we additionally tested the bats in the +40 dB BFN condition. During these TD tasks, we measured the masked auditory detection threshold of the bats and recorded their echolocation calls simultaneously. The value of one animal's absolute hearing at 20 kHz pure tone was so unusually high, beyond the dynamic range of our playback system, and thus was not measured.

To test whether echolocating bats adapt the echolocation calls to different auditory detection tasks, we trained and tested the same bats for the ED tasks after finishing the TD tasks. Again, in the ED task, the noise condition for individual bats was the same as that of the 51–61 kHz BFN. For the ED tasks, the bats were tested with two echo gain levels differing by 10 dB. As the detection performance of the bats was below 70% success in the low-echo-gain ED tasks after the 1st round of tests, we re-tested them with the same tasks (i.e., the 2nd round) for two weeks after the 1st round of ED tasks learning. This re-test served to verify whether the changes in call amplitude represented a learning-driven, adaptive vocal adjustment for the ED task. Specifically, if the same bats showed improved performance and concomitant adjustments in call amplitude between the two rounds, it would demonstrate an effect of task learning or familiarity on vocal control.

Data analysis

Echolocation calls were batch-processed with custom-written MATLAB scripts (R2019b, MathWorks, USA). Calls with a higher peak amplitude from the two recording channels were used for further spectro-temporal analyses. In the present study, we focused on the parameter of call amplitude, including both the CF amplitude and FM amplitude measured from the dominant second harmonic. This is because call amplitude plays a dominant role in affecting signal detectability when compared to other call parameters such as call duration (Luo et al., 2016). To compare the effect of the detection task *per se* on echolocation call control, we compared the echolocation call recorded 1.5 s before and after the start of the target stimuli broadcast, i.e., calls between the pre-playback window and the playback window (Figure 1C). The Lombard effect referred to the difference in call amplitude between noise and quiet conditions in different tasks. The magnitude of this increase represented the degree of call compensation, with a larger difference indicating a stronger

compensation. In total, we analyzed 71 206 echolocation calls from the TD tasks, 66 054 echolocation calls from the ED tasks, 216 717 echolocation calls from the learning phase of the TD task, and additionally 283 768 echolocation calls from the learning phase of the ED task. Details are shown in Supplementary Table S1.

All statistical analyses were performed with the Statistical and Machine Learning Toolbox in MATLAB (R2021a, MathWorks, USA). We used the linear mixed model by setting call amplitude (CF amplitude or FM amplitude) as the dependent variable, the experimental condition as the fixed effect (categorical), and the identity of the individual as the random effect (categorical), to examine whether call amplitude of the bats statistically differed between different noise conditions versus silence control. The general formula of the mixed model is 'amplitude ~1+condition+(1+condition| individual)'. The data used for plotting are the same as those used in the mixed model. We used the estimated values from the mixed model for data comparison, after controlling for the effects of other factors in the model. The statistical significance of comparisons between conditions depends on the overlap of the upper and lower limits predicted by the mixed model. We also calculated the population effect size, defined as 0.8 times the standard deviation for each pairwise comparison. We used linear regression and partial correlation to analyze the changes in amplitude during ED tasks learning. A significance level of α=0.05 was adopted in all tests. Due to the small sample size (number of bats), which is yet typical for psychophysical experiments on animals due to the lengthy training required (Bates et al., 2011; Geberl et al., 2019; Yovel et al., 2009), all analyses were also conducted for each individual animal (Supplementary Table S2–S18). Wilcoxon rank sum test was used for the statistical analyses for individual animals.

RESULTS

Task-specific vocal flexibility in bats performing sound detection tasks

To investigate whether bats instantly modify the amplitude of their emitted calls to the playback of target signals in quiet, for each trial, we calculated the amplitude difference for the calls between the playback window and the pre-playback window (Figure 1I). For each call, we analyzed the amplitude for the CF and FM components separately. We used the Linear Mixed Models (LMM) to assess the statistical significance. We found that in the four TD tasks for different target signal frequencies, there were little or only minor changes in the amplitude of the CF or the FM components (Supplementary Figure S1; Supplementary Table S10–S15). The maximum amplitude decrease for the CF component was 2.1 dB when detecting the 40 kHz tones. The maximum amplitude increase for the FM component was 2.0 dB when detecting the 20 kHz tones. By contrast, the bats significantly decreased the amplitude of the CF component when detecting virtual echoes (Figure 2A; Supplementary Table S2; LMMs, all t<–3.74, all P<0.05). Specifically, when detecting HEL virtual echoes in both the 1st and 2nd rounds the bats decreased the median amplitude of the CF component by a similar amount of –4.1 dB and –4.5 dB respectively (Supplementary Table S2; LMMs, t<–9.08, P<0.05). When detecting LEL virtual echoes in the 1st round, the bats decreased the median amplitude of the CF component by –4.0 dB, but only by –1.6 dB in 2nd round, evidencing an effect of task relearning or familiarity on vocal flexibility, that is, bat repeated exposure to LEL task led to adaptation. Moreover, the bats did not modify the amplitude of the FM component in any ED tasks (Figure 2B; Supplementary Table S3; LMMs, t<0.44, all P>0.05). Together, these data suggest that the instant vocal flexibility exhibited by the Pratt's roundleaf bats is task-specific.

To investigate whether bats also modify the amplitude of their emitted calls in a task-specific way in the presence of masking noise, we compared the magnitude of the Lombard effect between the TD task and the ED tasks. The magnitude of the Lombard effect was measured as the amplitude differences between the noise condition and the silence control, using the calls in the playback window (Figure 1I**)**. We found that bats exhibited a typical Lombard effect in noise in all conditions, except for the CF component in the 20-kHz and 40-kHz TD tasks in which the noise did not overlap in frequency with echolocation calls (Figure 2C, D; Supplementary Figure S1; Supplementary Table S5–S7). The Lombard effect was stronger for the FM component than for the CF component by 2.4 to 9.2 dB (medians) across different conditions (Rank-sum tests, all z<–5.49, all P<0.05), except for the 2nd round of tests with HEL virtual echoes (Supplementary Figure S2A).

Next, we compared the magnitude of the Lombard effect between the 60-kHz TD task and the ED tasks. In these conditions, bats were exposed to the same noise stimuli, i.e., the 51–61 kHz bandpass filtered white noise that overlapped with the dominant second harmonic of the echolocation calls (see insets above Figure 2C). We found that in both the +30 dB and +40 dB noise conditions, the bats exhibited a Lombard effect in the TD and ED tasks for both the CF and FM components (Figure 2C, D; Supplementary Table S5–S7; LMMs, all t>3.22, all P<0.05). We compared the probability distributions of the Lombard effect difference for each call between the two detection tasks, analyzing the CF and FM components separately. However, the magnitude of the Lombard effect differed between the TD and ED tasks and depended on the call component. Specifically, the Lombard effect of the CF component was stronger in the ED tasks than in the TD task (Figure 2E; median, 2.1 dB; range: –7.5 dB and 12.1 dB; Figure 2F; median, 2.6 dB; range: –9.2 dB and 12.8 dB; Supplementary Figure S2B; Rank-sum tests, all z>5.81, all P<0.05), while the Lombard effect of the FM component was weaker in the ED tasks than in the TD task (Figure 2E, median, –1.8 dB; range: –11.5 dB and 4.5 dB; Figure 2F, median, –1.9 dB; range: –12.4 dB and 5.0 dB; Supplementary Figure S2C; Rank-sum tests, all z<–5.79, P<0.05). By contrast, the effect of task relearning on the Lombard effect was complex and depended on the noise level, echo level, and call component (Figure 2C, D). We found that in the low-echo-level ED task, the bats increased the Lombard effect for both the CF and FM components in the +40 dB noise condition after task relearning (Figure 2C, D; Rank-sum tests, z=3.50 and 5.05, both P<0.05). Together, these data showed that the Lombard effect in bats is modulated by detection tasks and can be affected by task relearning.

Long-term task learning drives adaptive vocal adjustments in bats

As in the test phase, the bats were actively vocalizing during the learning phase of both the TD and ED tasks. When learning the TD task, the bats emitted 9.8±3.5 and 10.4±4.0 echolocation calls per second in the pre-playback and playback windows respectively (n=4 357 trials from 4 bats).

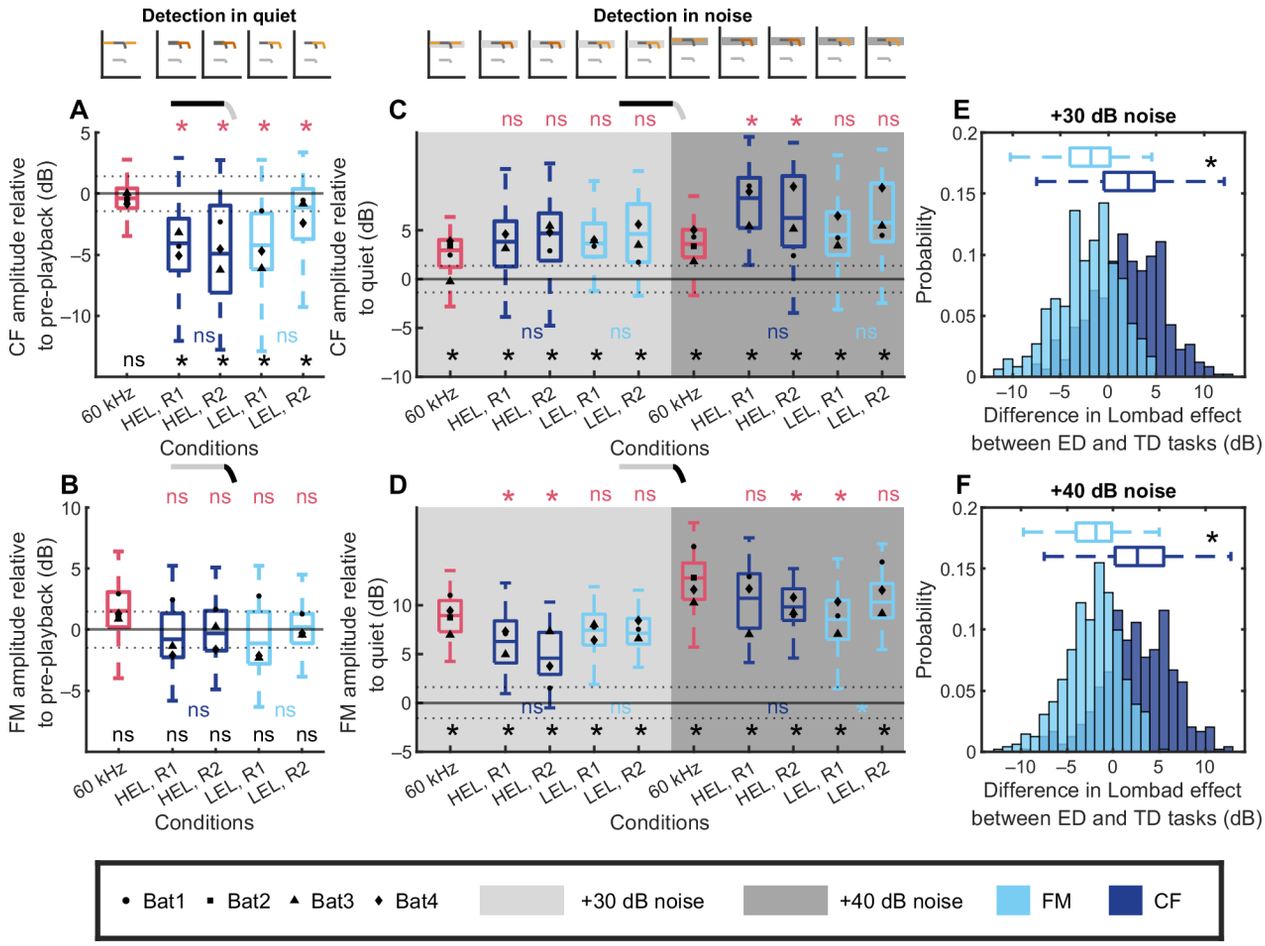

Figure 2 Echolocation call amplitude of *Hipposideros pratti* in two detection tasks

A, B: Amplitude changes of the CF (A) and FM (B) components of echolocation calls when bats were performing the tone detection task and echo detection task in quiet. Calls in the pre-playback window (1.5 seconds before the start of the playback window) were used as the control. C, D: Amplitude changes of the CF (C) and FM (D) components of echolocation calls when bats were performing the tone detection task and echo detection task in noise. Calls in the playback window in quiet were used as the control separately. In both tasks, analyzed calls were from the 1.5 seconds after the start of the playback window. E, F: Differences in the Lombard effect between the two detection tasks for the CF (E) and FM (F) components respectively. Boxplots show the median, the lower and higher quartiles, and whiskers (multiplier of 1.5) respectively. In panel A–D, medians of the data for individual bats are shown as filled markers in black. Bat1, circles; Bat2, triangles; Bat3, squares; and Bat4, diamonds. The statistical significance of comparisons between a condition and the control is indicated by an asterisk (*) or "ns" in black; between the two detection tasks in the same noise level by an asterisk (*) or "ns" in pink; between the two round ED tasks in the same noise level by an asterisk (*) or "ns" in the same color with the color of boxplot; and between the FM component and CF component by an asterisk (*) in black in panel (E, F). For panels E and F, the data represent the difference in Lombard effect between the ED and TD conditions for each individual. This was calculated by subtracting the median Lombard effect in the TD condition (the reference) from the value in the ED condition, under identical noise levels (+30 dB in (E), +40 dB in (F)). *: P<0.05; ns: Not significant. Gray dotted lines show the median effect size of the population calculated as 0.8 times the standard deviation of the pooled data. All boxplots were based on the medians at the trial level. For example, each data point of the boxplot in panel A was the median difference of call amplitudes during the test and before the test in a trial. CF: constant frequency; FM: frequency-modulated; HEL: high-echo-level; LEL: low-echo-level; R1: the 1st round echo detection task; R2: the 2nd round echo detection task; TD: tone detection; ED: echo detection.

When learning the ED task, the bats emitted 8.7±3.4 and 11.0±4.9 echolocation calls per second in the pre-playback and playback windows respectively (n=6 928 trials from 3 bats). Since the emitted echolocation calls can act as maskers interfering with the detection of target signals, and determine the echo amplitude in the ED tasks, we were keen to know whether bats can learn to adapt their echolocation calls, specifically call amplitude, to the detection tasks. We focused on echolocation calls during the pre-playback window (Figure 1I), as these calls are not directly related to the instant perceptual needs for detecting the target signals broadcast in the subsequent playback window, and thus reflect the long-term learning effect.

We analyzed the correlations between the call amplitude, task learning days, and daily task performance for individual bats. Like in the analysis above, we examined the amplitudes of the CF and FM components separately, as they are implicated in different functions of echolocation. We found that bats adapted their call amplitude in a task-specific manner throughout the learning period. Bats generally decreased the amplitudes of both the CF and FM components while learning the TD task (Figure 3A, B; Linear Partial Correlation: r=–0.44, –0.47, P<0.05), whereas they increased their amplitudes while learning the ED task (Figure 3C, D; Linear Partial Correlation: r=0.51, 0.27, P<0.05 for all bats). The amplitudes of the CF and FM components decreased by 6.3 dB and 15.4 dB, respectively, from the beginning to the end of the TD task learning. By comparison, the amplitudes of the CF and FM

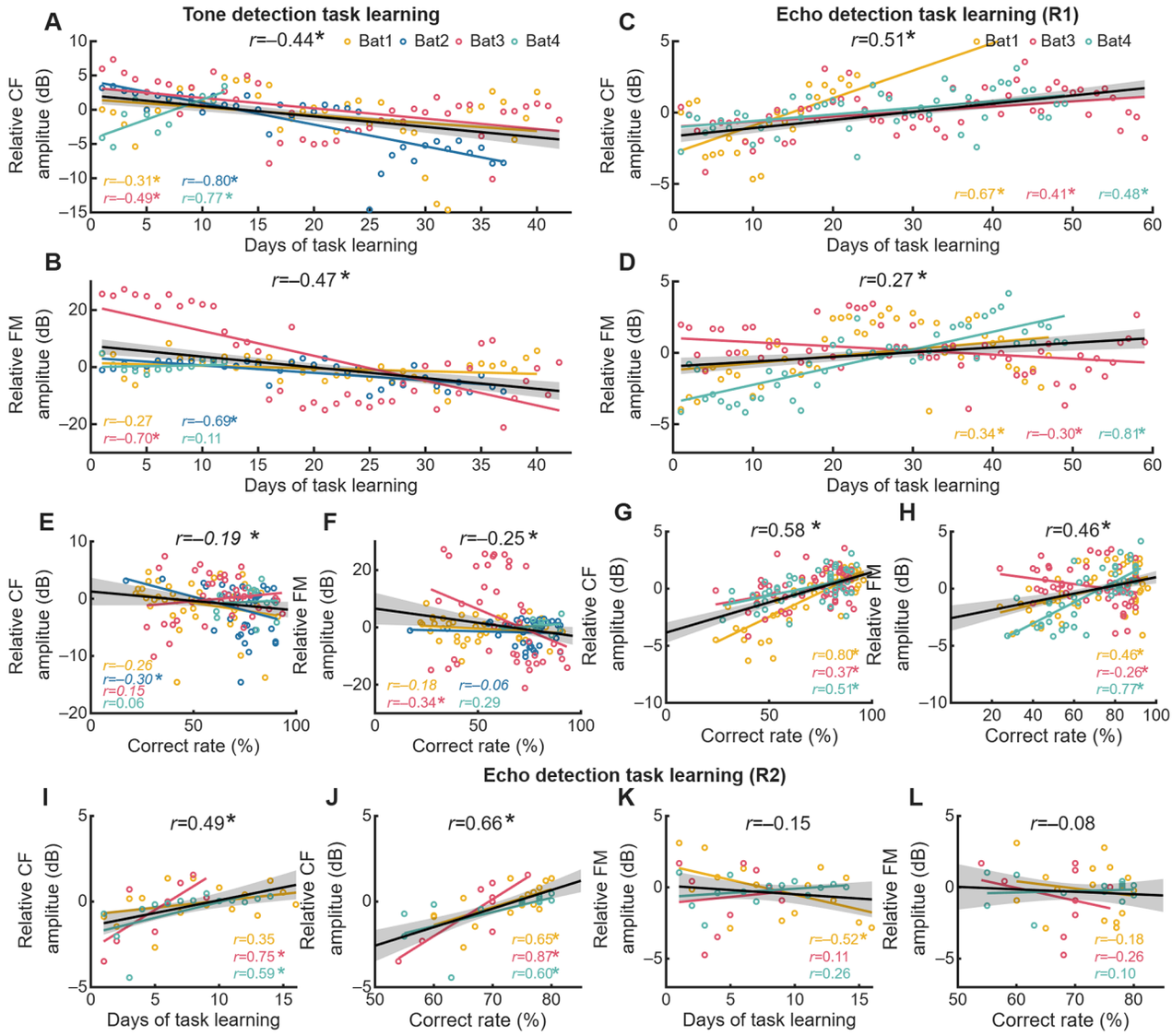

Figure 3 Task learning shapes the call amplitude of bats

A, B: Correlations between call component amplitude (CF (A) and FM (B)) of bats and the days of the tone detection task learning. C, D: Correlations between call component amplitude (CF (C) and FM (D)) of bat and the days of the echo detection task learning (with high-echo-level) across the 1st round learning phase. E, F: Correlations between call component amplitude (CF (E) and FM (F)) of bats and task performance in the tone detection task learning. G, H: Correlations between call component amplitude (CF (G) and FM (H)) and task performance (with high-echo-level) across the 1st round learning phase. I–L: As in C, D, G, H for *Hipposideros pratti* during the two-week extra task learning phase, i.e., the 2nd round of task learning. In all panels, the solid black lines show the fit from data of all bats by linear regression and the gray area shows the 95% confidence interval (*: $P<0.05$). Each data point, represented by a circle, was the average of the call amplitudes across all trials for a bat. CF: constant frequency; FM: frequency-modulated; R1: the 1st round echo detection task; R2: the 2nd round echo detection task.

components were 3.3 dB and 1.9 dB higher at the end than at the beginning of the ED task learning. Thus, the bats reversed the direction of amplitude adjustments from the TD task learning to the ED task learning.

At the individual level, only in the ED task did all bats consistently adjust the amplitude of the CF component (Figure 3C; Pearson Correlation: Bat1, r=0.67, P<0.05; Bat3, r=0.41, P<0.05; Bat4, r=0.48, P<0.05). When learning the TD task, three bats significantly decreased the amplitude of the CF component, but one bat significantly increased it (Figure 3A; Pearson Correlation: Bat1, r=–0.31, P<0.05; Bat2, r=–0.80, P<0.05; Bat3, r=–0.49, P<0.05; Bat4, r=0.77, P<0.05); two bats significantly decreased the amplitude of the FM component, but two other bats did not (Figure 3B; Pearson Correlation: Bat1, r=–0.27, P=0.09; Bat2, r=–0.69, P<0.05; Bat3, r=–0.70, P<0.05; Bat4, r=0.11, P=0.73). When learning the ED task, two bats significantly increased the amplitude of the FM component, but one bat significantly decreased it (Figure 3D; Pearson Correlation: Bat1, r=0.34, P<0.05; Bat3, r=–0.3, P<0.05; Bat4, r=0.81, P<0.05). These overall and individual relationships between call amplitude and days of task learning are reflected by the relationships between call amplitude and performance rate (Figure 3E–H). Again, only in the ED task did all bats exhibit a consistent, positive correlation between the amplitude of the CF component and performance rate (Figure 3G; Linear Partial Correlation: r=0.58, P<0.05 for all bats; Pearson Correlation: Bat1, r=0.80, P<0.05; Bat3, r=0.37, P<0.05; Bat4, r=0.51, P<0.05).

To confirm that an increase in the CF amplitude represents a learning-driven, adaptive vocal adjustment for the ED task,

we retrained the bats for two weeks after completing the first round of tests (Figure 3I–L). We found that, even for this relatively short retraining period, the bats again increased the amplitude of the CF component (Figure 3I; Linear Partial Correlation: r=0.49, P<0.05 for all bats), which led to improved task performance (Figure 3J; Linear Partial Correlation: r=0.66, P<0.05 for all bats). By contrast, there was no significant relationship between the amplitude of the FM component and task learning days (Figure 3K; Linear Partial Correlation: r=–0.15, P=0.35 for all bats) or task performance (Figure 3L; Linear Partial Correlation: r=–0.08, P=0.95 for all bats).

DISCUSSION

This study provided three novel insights into the vocal adaptation capacity of echolocating bats. First, the instant vocal plasticity, i.e., vocal flexibility, is task-specific and differs between the TD and ED tasks. Second, the Lombard effect, a reflexive form of vocal flexibility, is also task-specific and can be modulated by task relearning. Third, long-term task learning drives adaptive amplitude adjustments in echolocating bats. Therefore, our results provide support for the Vocal Flexibility Facilitates Learning (VFFL) Hypothesis.

Importantly, we demonstrate that bats engage in long-term vocal adaptation to modify their vocalizations in a goal-directed manner. We found that over two months, the Pratt's roundleaf bats gradually increased the CF amplitude of their echolocation calls to improve task performance in the ED tasks. Conversely, they gradually decreased the CF amplitude in the TD tasks. In both detection tasks, the observed call amplitude adjustments are adaptive. In the ED task, the raised amplitude of the emitted calls led to higher echo levels and thus made echoes more detectable. In the TD tasks, the decreased amplitude of the emitted calls can reduce their potential masking effect on detecting the tones. Although these findings involve the modification of existing vocal parameters rather than novel vocalization, the results from these reinforcement learning experiments are consistent with the premise that *H. pratti* is a candidate species for VPL.

This long-term vocal adaptation contrasts sharply with the short-term amplitude adjustments observed at the beginning of the tasks. While bats exhibited an immediate decrease in call amplitude in the ED task, they showed minimal change in the TD task. The rapid adjustment in the ED task is likely mediated by the echo feedback loop, consistent with the known phenomenon of echo-level compensation, where bats adaptively lower call amplitude in response to increasing echo levels during target approach (Koblitz et al., 2011; Luo et al., 2023; Stidsholt et al., 2020). The absence of this adjustment in the TD task suggests that pure tones are insufficient to trigger echo-level compensation behavior of bats.

Moreover, this study has shown that Pratt's roundleaf bats possess an advanced ability for vocal flexibility. They showed gradual, task-specific call amplitude adjustments over months of task learning. We found that these bats expressed the Lombard effect differently in the TD and ED tasks, with a stronger compensation for the CF component but a weaker compensation for the FM component in the ED tasks compared to the TD tasks. These data were consistent with a previous study that reported CF-FM bats can differentially modulate the amplitude of the CF and FM components in noise (Lu et al., 2020). So far, the Lombard effect is primarily regarded as an adaptive reflex. Virtually all differences in the expression of the Lombard effect in animal subjects can be attributed to the physical properties of the background noise *per se*, being either the level, the bandwidth, or the duration of the noise (Hage et al., 2013; Luo et al., 2016, 2018; Tressler & Smotherman, 2009; Zollinger & Brumm, 2011). Thus, the task-specific Lombard effect in bats may parallel those in human subjects (Garnier et al., 2010; Villegas et al., 2021). Surprisingly, the bats showed a stronger Lombard effect after task relearning. These data show that the Lombard effect, despite its reflexive nature, is more dynamic than previously thought, and thus may involve more distributed brain networks, e.g., associated with learning or reward.

In short, our findings bridge the knowledge gap between vocal flexibility and VPL by showing that both short-term (instant) and long-term vocal adaptations can result from task learning in echolocating bats. We propose that demands for perceptual efficiency during long-term echolocation tasks may have driven the evolution of this rare VPL ability in bats, as well as in echolocators like cetaceans and pinnipeds. Future work involving a larger number of bats engaged in perceptual tasks, which is feasible with automated animal training setups (Lattenkamp et al., 2018), is essential to fully address the link between long-term vocal adaptation and VPL.

DATA AVAILABILITY

All data and codes have been deposited at Science Data Bank and are publicly available as of the date of publication (https://doi.org/10.57760/sciencedb.42137)

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

J.L. designed the study, Q.W. and H.Y. collected and analyzed the data, and J.L. and Q.W. wrote the paper. All authors read and approved the final version of the manuscript.

REFERENCES

Ancillotto L, Pafundi D, Cappa F, et al. 2022. Bats mimic hymenopteran insect sounds to deter predators. *Current Biology*, **32**(9): R408–R409.

Bates ME, Simmons JA, Zorikov TV. 2011. Bats use echo harmonic structure to distinguish their targets from background clutter. *Science*, **333**(6042): 627–630.

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

Esser KH. 1994. Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *NeuroReport*, **5**(14): 1718–1720.

Fenton MB, Faure PA, Ratcliffe JM. 2012. Evolution of high duty cycle echolocation in bats. *Journal of Experimental Biology*, **215**(17): 2935–2944.

Fernandez AA, Burchardt LS, Nagy M, et al. 2021. Babbling in a vocal learning bat resembles human infant babbling. *Science*, **373**(6557): 923–926.

Garnier M, Henrich N, Dubois D. 2010. Influence of sound immersion and communicative interaction on the Lombard effect. *Journal of Speech, Language, and Hearing Research*, **53**(3): 588–608.

Geberl C, Kugler K, Wiegrebe L. 2019. The spatial resolution of bat biosonar quantified with a visual-resolution paradigm. *Current Biology*, **29**(11): 1842–1846. e3.

Genzel D, Desai J, Paras E, et al. 2019. Long-term and persistent vocal plasticity in adult bats. *Nature Communications*, **10**(1): 3372.

Griffin DR, Webster FA, Michael CR. 1960. The echolocation of flying insects by bats. *Animal Behaviour*, **8**(3-4): 141–154.

Hage SR, Jiang TL, Berquist SW, et al. 2013. Ambient noise induces independent shifts in call frequency and amplitude within the Lombard effect in echolocating bats. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(10): 4063–4068.

Hiryu S, Hagino T, Riquimaroux H, et al. 2007. Echo-intensity compensation in echolocating bats (*Pipistrellus abramus*) during flight measured by a telemetry microphone. *The Journal of the Acoustical Society of America*, **121**(3): 1749–1757.

Hiryu S, Katsura K, Nagato T, et al. 2006. Intra-individual variation in the vocalized frequency of the Taiwanese leaf-nosed bat, *Hipposideros terasensis*, influenced by conspecific colony members. *Journal of Comparative Physiology A*, **192**(8): 807–815.

Jakobsen L, Brinkløv S, Surlykke A. 2013. Intensity and directionality of bat echolocation signals. *Frontiers in Physiology*, **4**: 89.

Janik VM. 2000. Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, **289**(5483): 1355–1357.

Janik VM, Knörnschild M. 2021. Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **376**(1836): 20200244.

Janik VM, Slater PJB. 1997. Vocal learning in mammals. *Advances in the Study of Behavior*, **26**: 59–99.

Jarvis ED. 2019. Evolution of vocal learning and spoken language. *Science*, **366**(6461): 50–54.

Knörnschild M. 2014. Vocal production learning in bats. *Current Opinion in Neurobiology*, **28**: 80–85.

Koblitz JC, Stilz P, Pflästerer W, et al. 2011. Source level reduction and sonar beam aiming in landing big brown bats (*Eptesicus fuscus*). *The Journal of the Acoustical Society of America*, **130**(5): 3090–3099.

Lattenkamp EZ, Vernes SC, Wiegrebe L. 2018. Volitional control of social vocalisations and vocal usage learning in bats. *Journal of Experimental Biology*, **221**(14): jeb180729.

Leek MR. 2001. Adaptive procedures in psychophysical research. *Perception & Psychophysics*, **63**(8): 1279–1292.

Liu ZX, Cao QH, Lu MM, et al. 2026. Echolocation behaviour is highly sensitive to noise interference in flying Pratt's roundleaf bats. *Animal Behaviour*, **233**: 123465.

Lombard E. 1911. Le signe de l'elevation de la voix. *Annales des Maladies de l'Oreille, du Larynx, du Nez et du Pharynx*, **37**: 101–119.

Lu MM, Zhang GM, Luo JH. 2020. Echolocating bats exhibit differential amplitude compensation for noise interference at a sub-call level. *Journal of Experimental Biology*, **223**(19): 225284.

Luo JH, Goerlitz HR, Brumm H, et al. 2016. Linking the sender to the receiver: vocal adjustments by bats to maintain signal detection in noise. *Scientific Reports*, **5**: 18556.

Luo JH, Hage SR, Moss CF. 2018. The Lombard effect: from acoustics to neural mechanisms. *Trends in Neurosciences*, **41**(12): 938–949.

Luo JH, Kothari NB, Moss CF. 2017a. Sensorimotor integration on a rapid time scale. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(25): 6605–6610.

Luo JH, Lingner A, Firzlaff U, et al. 2017b. The Lombard effect emerges early in young bats: implications for the development of audio-vocal integration. *Journal of Experimental Biology*, **220**(6): 1032–1037.

Luo JH, Lu MM, Luo J, et al. 2023. Echo feedback mediates noise-induced vocal modifications in flying bats. *Journal of Comparative Physiology A*, **209**(1): 203–214.

Ma NN, Xia HJ, Yu C, et al. 2024. Effects of insect pursuit on the Doppler shift compensation in a hipposiderid bat. *Journal of Experimental Biology*, **227**(6): jeb246355.

Ma NN, Xia HJ, Zheng HJ, et al. 2025. Prey evasiveness and masking noise jointly promote the ultrahigh call rate in echolocating bats. *BMC Biology*, **23**(1): 323.

Madsen PT, Surlykke A. 2013. Functional convergence in bat and toothed whale biosonars. *Physiology*, **28**(5): 276–283.

Moss CF, Chiu C, Surlykke A. 2011. Adaptive vocal behavior drives perception by echolocation in bats. *Current Opinion in Neurobiology*, **21**(4): 645–652.

Pedersen MB, Egenhardt M, Beedholm K, et al. 2024. Superfast Lombard response in free-flying, echolocating bats. *Current Biology*, **34**(11): 2509–2516. e3.

Pick Jr HL, Siegel GM, Fox PW, et al. 1989. Inhibiting the Lombard effect. *The Journal of the Acoustical Society of America*, **85**(2): 894–900.

Prat Y, Azoulay L, Dor R, et al. 2017. Crowd vocal learning induces vocal dialects in bats: playback of conspecifics shapes fundamental frequency usage by pups. *PLoS Biology*, **15**(10): e2002556.

Prat Y, Taub M, Yovel Y. 2015. Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats. *Science Advances*, **1**(2): e1500019.

Ratcliffe JM, Elemans CPH, Jakobsen L, et al. 2013. How the bat got its buzz. *Biology Letters*, **9**(2): 20121031.

Riquimaroux H, Gaioni SJ, Suga N. 1991. Cortical computational maps control auditory perception. *Science*, **251**(4993): 565–568.

Schnitzler HU, Denzinger A. 2011. Auditory fovea and Doppler shift compensation: adaptations for flutter detection in echolocating bats using CF-FM signals. *Journal of Comparative Physiology A*, **197**(5): 541–559.

Schnitzler HU, Kalko EKV. 2001. Echolocation by insect-eating bats: we define four distinct functional groups of bats and find differences in signal structure that correlate with the typical echolocation tasks faced by each group. *BioScience*, **51**(7): 557–569.

Simmons JA, Fenton MB, O'Farrell MJ. 1979. Echolocation and pursuit of prey by bats. *Science*, **203**(4375): 16–21.

Smotherman MS, Simmons AM, Simmons JA. 2021. How noise affects bats and what it reveals about their biosonar systems. *In*: Lim BK, Fenton MB, Brigham RM, et al. 50 Years of Bat Research: Foundations and New Frontiers. Cham: Springer, 61–76.

Stansbury AL, Janik VM. 2019. Formant modification through vocal production learning in gray seals. *Current Biology*, **29**(13): 2244–2249. e4.

Stidsholt L, Müller R, Beedholm K, et al. 2020. Energy compensation and received echo level dynamics in constant-frequency bats during active target approaches. *Journal of Experimental Biology*, **223**(2): jeb217109.

Teeling EC, Vernes SC, Dávalos LM, et al. 2018. Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. *Annual Review of Animal Biosciences*, **6**: 23–46.

Thomas JA, Moss CF, Vater M. 2004. Echolocation in Bats and Dolphins. Chicago: University of Chicago Press.

Tressler J, Smotherman MS. 2009. Context-dependent effects of noise on echolocation pulse characteristics in free-tailed bats. *Journal of Comparative Physiology A*, **195**(10): 923–934.

Tyack PL. 2016. Vocal learning and auditory-vocal feedback. *In*: Suthers RA, Fitch WT, Fay RR, et al. Vertebrate Sound Production and Acoustic Communication. Chem: Springer, 261–295.

Tyack PL. 2019. A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **375**(1789): 20180406.

Vernes SC, Devanna P, Hörpel SG, et al. 2022. The pale spear‐nosed bat: a neuromolecular and transgenic model for vocal learning. *Annals of the New York Academy of Sciences*, **1517**(1): 125–142.

Vernes SC, Kriengwatana BP, Beeck VC, et al. 2021. The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **376**(1836): 20200236.

Vernes SC, Wilkinson GS. 2020. Behaviour, biology and evolution of vocal learning in bats. *Philosophical Transactions of the Royal Society B*, **375**(1789): 20190061.

Villegas J, Perkins J, Wilson I. 2021. Effects of task and language

nativeness on the Lombard effect and on its onset and offset timing. *The Journal of the Acoustical Society of America*, **149**(3): 1855–1865.

Wang ZQ, Gong LX, Si M, et al. 2025. Seasonal changes in 3D space use and foraging strategies in the avivorous bat *Ia io*. *Zoological Research*, **46**(6): 1465–1476.

Wirthlin M, Chang EF, Knörnschild M, et al. 2019. A modular approach to vocal learning: disentangling the diversity of a complex behavioral trait. *Neuron*, **104**(1): 87–99.

Xia HJ, Ma NN, Li AQ, et al. 2025. Call production and wingbeat coupling is flexible and species-specific in echolocating bats. *Annals of the New York Academy of Sciences*, **1547**(1): 105–115.

Ye H, Luo JH. 2022. Perceptual hearing sensitivity during vocal production. *iScience*, **25**(11): 105435.

Yovel Y, Melcon ML, Franz MO, et al. 2009. The voice of bats: how greater mouse-eared bats recognize individuals based on their echolocation calls. *PLoS Computational Biology*, **5**(6): e1000400.

Zhang KK, Wang HW, Lin HJ, et al. 2026. Geographic and genetic factors shape acoustic divergence and dialect-like variation in communication calls in the greater horseshoe bats. *Zoological Research*, **47**(2): 414–428.

Zhao L, Yuan JQ, Wang GQ, et al. 2024. Chromosome-level genome and population genomics of the intermediate horseshoe bat (*Rhinolophus affinis*) reveal the molecular basis of virus tolerance in Rhinolophus and echolocation call frequency variation. *Zoological Research*, **45**(5): 1147–1160.

Zollinger SA, Brumm H. 2011. The Lombard effect. *Current Biology*, **21**(16): R614–R615.