Sensory error drives fine motor adjustment

Huimin Wang[a,1], Yuxuan Zhou[b,1], Huanhuan Li[c], Cynthia F. Moss[d], Xingxing Li[b,2], and Jinhong Luo[a,2]

Fine audiovocal control is a hallmark of human speech production and depends on precisely coordinated muscle activity guided by sensory feedback. Little is known about shared audiovocal mechanisms between humans and other mammals. We hypothesized that real-time audiovocal control in bat echolocation uses the same computational principles as human speech. To test the prediction of this hypothesis, we applied state feedback control (SFC) theory to the analysis of call frequency adjustments in the echolocating bat, *Hipposideros armiger*. This model organism exhibits well-developed audiovocal control to sense its surroundings via echolocation. Our experimental paradigm was analogous to one implemented in human subjects. We measured the bats' vocal responses to spectrally altered echolocation calls. Individual bats exhibited highly distinct patterns of vocal compensation to these altered calls. Our findings mirror typical observations of speech control in humans listening to spectrally altered speech. Using mathematical modeling, we determined that the same computational principles of SFC apply to bat echolocation and human speech, confirming the prediction of our hypothesis.

auditory feedback | echolocation | vocal production control | Kalman filter | human speech

Significance

Audiovocal feedback control is essential to speech, a uniquely human product. Yet the foundations of our acoustic communication system share mechanisms with other mammals. We discovered that bat echolocation and human speech rely on common computational principles for audiovocal feedback control. This finding demonstrates that general mechanisms can be uncovered through comparative studies.

The use of sensory signals to adapt motor actions is fundamental to a wide array of human and animal behaviors. Sensorimotor adaptations require precise, continuous coordination of stimulus processing and muscle activity on a rapid time scale (1, 2). A salient example is the fine audiovocal control in human speech production (3). Although speech is a uniquely human product, sensorimotor speech control may share mechanisms with some animals (4, 5). Identifying such shared mechanisms requires suitable mammalian models of comparable vocal control capacity and experimental tractability.

Echolocating bats offer a tractable model to probe mechanisms of audiovocal control. They share basic brain anatomy with humans and demonstrate an exceptional capacity for rapid, precise vocal feedback control (6–8). Echolocating bats that use short-duration (low duty cycle) frequency-modulated calls adaptively adjust call features to changes in echo properties with a reaction time of merely a tenth of a second (9–11). Other bat species that use long-duration (high duty cycle) constant frequency (CF) calls dynamically control the frequency of their sonar vocalizations to compensate for Doppler shifts introduced by their own flight, and such audiovocal control is commonly known as Doppler shift compensation (DSC) (12–14). Only bat species in the families Rhinolophidae and Hipposideridae and two species of Mormoopidae exhibit DSC (13, 14). These bat species use DSC to adjust the frequency of echo returns with high precision (15–17) and benefit from the frequency separation of calls and echoes (18).

Humans, like bats, perform compensatory vocal adjustments to altered auditory feedback (19–22). Human subjects rapidly change the pitch of their speech when they hear a frequency-shifted copy of their own vocalizations. The state feedback control (SFC) theory explains human vocal frequency control in response to altered auditory feedback (23–25). In the SFC framework, sensory error, or feedback prediction error, drives vocal frequency adjustment. Sensory error is a neural estimation of the difference between predicted and actual feedback from the auditory and laryngeal somatosensory systems (3, 20, 23). When sensory errors occur, the vocal production system adjusts the signal parameter of subsequent vocalizations in the opposite direction. The SFC model can be implemented computationally as a Kalman filter that integrates the predicted state and sensory feedback based on relative uncertainty (1, 20).

Our study evaluates principles of the SFC model in the audiovocal behavior of the great roundleaf bat, *Hipposideros armiger*, a species that uses CF echolocation signals and shows high DSC precision of 0.15 to 0.17% (16, 17). This bat species' finely tuned sensorimotor behavior makes it a tractable model of audiovocal control. Given that bats share basic auditory and vocal motor control pathways with other mammals, including humans (26–28), we hypothesized that the computational principles of SFC

Author affiliations: [a]Hubei Key Laboratory of Genetic Regulation & Integrative Biology, School of Life Sciences, Central China Normal University, Wuhan 430079, China; [b]School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China; [c]Research and Development Department, Shanghai KeyGo Technology Company Limited, Shanghai 200090, China; and [d]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218

[1]H.W. and Y.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: jluo@ccnu.edu.cn or xxli@sgg.whu.edu.cn.

apply to *H. armiger*. We used an auditory feedback perturbation paradigm analogous to that for human subjects to test predictions of this hypothesis.

Results

We built a programmable auditory feedback perturbation system that delivers spectrally altered sounds to vocalizing bats in real time (Fig. 1*A*). The perturbation system consists of a microphone, a loudspeaker, and a digital signal processing (DSP) unit. It shifts the frequency of the microphone recording sample by sample up to a sampling rate of 1 MHz, without affecting the fine temporal structure of the echolocation calls (Fig. 1*B*). As shown in Fig. 1 *C–E*, postanalyses of the feedback stimuli confirmed the quality of the perturbation system across varying shift sizes and perturbation delays.

Compensatory Adjustment of Call Frequency to Spectrally Altered Feedback. Our research first investigated whether *H. armiger* shows compensatory adjustments of call frequency to altered auditory feedback. We found that there was little change in call frequency after the onset of 0-cent shifted feedback (Fig. 2 *A* and *D*). By contrast, the bat rapidly increased the call frequency after the onset of the −50-cent shifted feedback (Fig. 2 *B* and *E*), and it decreased the call frequency after the onset of the 50-cent shifted feedback (Fig. 2 *C* and *F*). Statistical analyses on the relative call frequency confirmed that the 0-cent shifted feedback did not affect the call frequency of the bat (Fig. 2*G*; signed-rank test, all *P* > 0.05 with effect sizes < 0.15 kHz). Conversely, the bat made compensatory adjustments of call frequency in the opposite direction of the shifted feedback as early as the first call after the perturbation onset (Fig. 2 *H* and *I*; signed-rank test, *P* < 0.05 with effect sizes > 0.15 kHz).

Next, we compared the frequency adjustments of *H. armiger* across different shift sizes. Fig. 3 *A* and *B* shows the time course of the relative call frequency of two individuals (Ha054 and Ha082) that received spectrally altered feedback of 0, ±20, and ±50 cents. When statistical analyses were conducted for all four *H. armiger* bats, we found that before the perturbations the maximum frequency adjustments were all centered around zero and did not differ across conditions (Fig. 3*C*; rank-sum test with post hoc pairwise comparisons, all *P* > 0.05). During the perturbations, *H. armiger* increased the call frequency after receiving the −50-cent shifted feedback and decreased the call frequency when receiving the 20- and 50-cent shifted feedback (Fig. 3*D*; rank-sum test with post hoc pairwise comparisons, all *P* < 0.05). *H. armiger* made a larger frequency adjustment in the 50-cent than in the 20-cent shifted feedback condition (rank-sum test with post hoc pairwise comparisons, *P* < 0.05). After the perturbations, apart from the 50-cent shifted feedback condition, the maximum frequency adjustment was similar across the shift sizes (Fig. 3*E*).

There is one notable difference between these two bats: Ha054 exhibited compensatory adjustments of call frequency to both positive (upward) and negative (downward) feedback shifts, while Ha082 only decreased the call frequency to the positive feedback shifts (Fig. 3 *A* and *B*). Similarly, Ha045 showed bidirectional frequency adjustments for shifted feedback, while Ha052 responded only to the positive feedback shifts (*SI Appendix*, Fig. S1 *A* and *B*). To explore the possibility that Ha082 and Ha052 failed to compensate for the negative feedback shifts because the maximum negative shift size of −50 cents was not large enough, we collected additional data with all four *H. armiger* bats receiving feedback stimuli of larger shift sizes (±100 and ±150 cents). We also repeated the ±50- and 0-cent shift size conditions. As in the first experiment, we found that Ha054 and Ha045 made compensatory adjustments in call frequency to the ±50-cent shifted feedback, Ha082 and Ha052 only decreased the call frequency to the 50-cent shifted feedback, and all *H. armiger* made no frequency adjustment to the 0-cent shifted feedback (Fig. 3 *F* and *G* and *SI Appendix*, Fig. S1 *C* and *D*). To our surprise, we did not observe an increase in call frequency in response to the larger negative feedback shifts of −100 or −150 cents in either Ha082 or Ha052 (Fig. 3*G* and *SI Appendix*, Fig. S1*D*). Notably, the ±50-cent feedback shifts were the conditions in which *H. armiger* made the
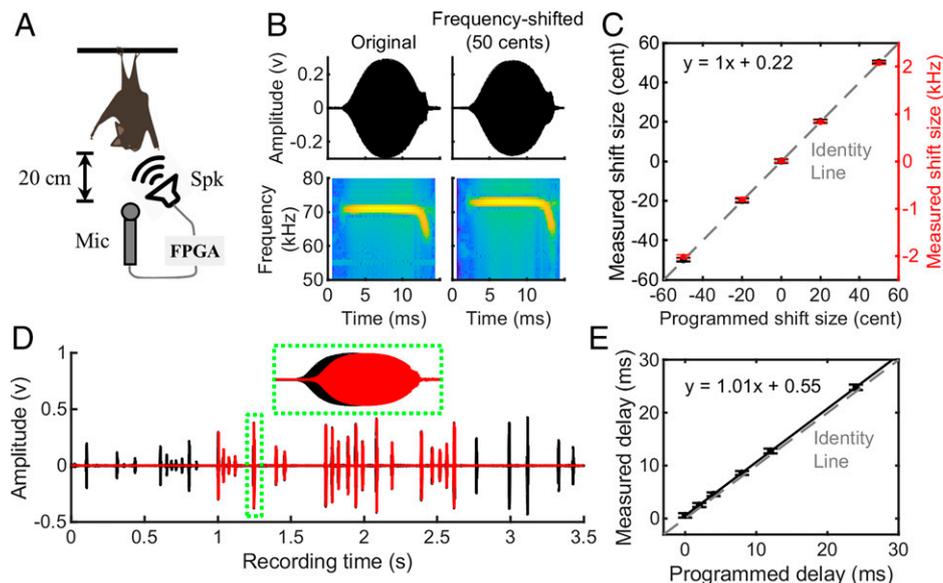
Fig. 1. The experimental setup and the real-time auditory feedback perturbation system. (*A*) An illustration of the relative spatial relationships of the components of the real-time auditory perturbation system. Mic, microphone; Spk, speaker. (*B*) The original and spectrally altered version of an echolocation call of *H. armiger*. (*C*) Theoretical (predicted) and estimated frequency shift sizes based on the real recordings of the feedback stimuli. (*D*) Synchronized plotting of a waveform of the original and spectrally altered feedback stimuli at the zero programmed delay reveals a minimum DSP delay of about 0.6 ms. (*E*) Theoretical (predicted) and estimated feedback delay based on the real recordings of the feedback stimuli. Equations in *C* and *E* describe the best linear fit.
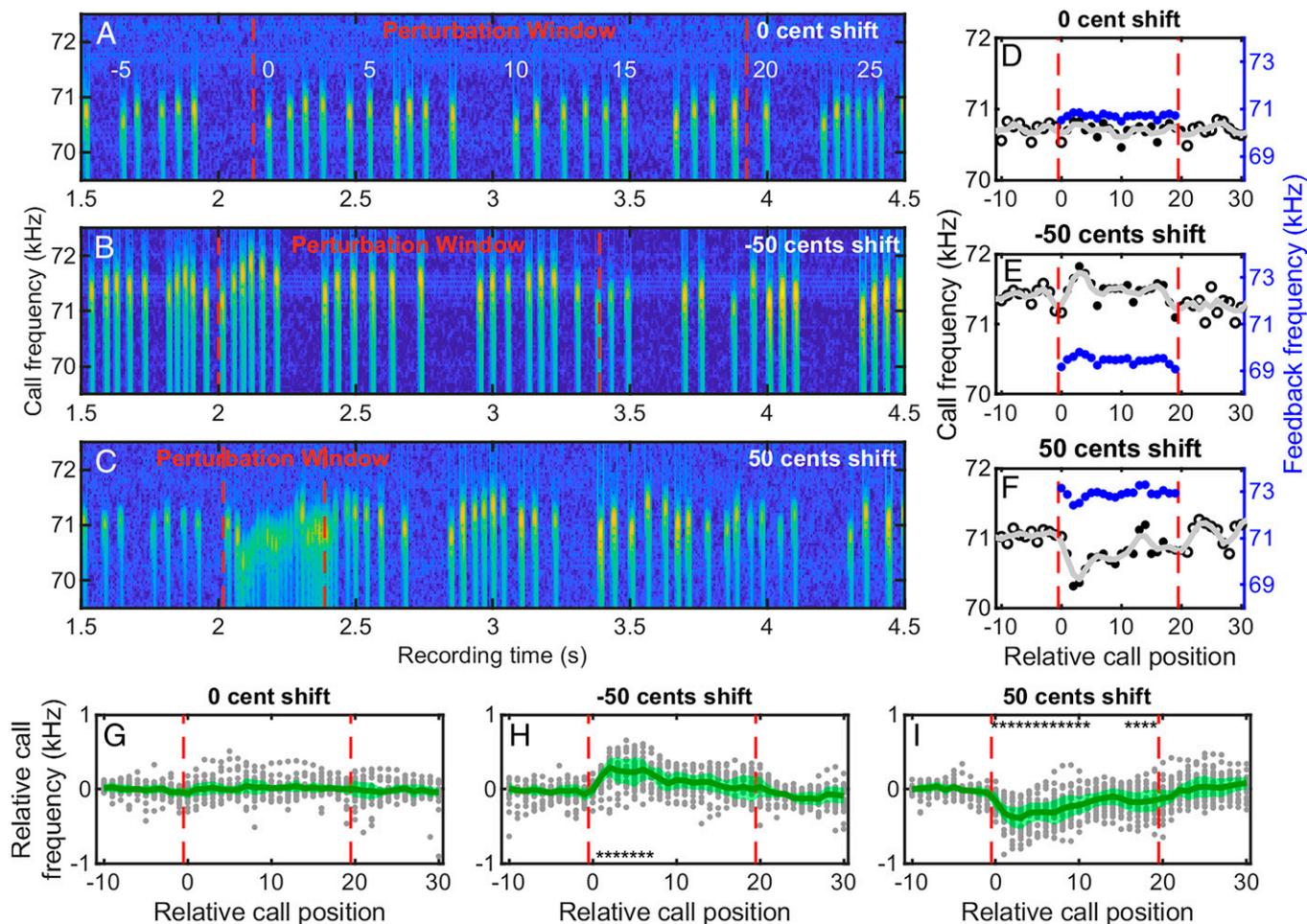
Fig. 2. A bat's compensatory frequency adjustment to altered auditory feedback of three shift sizes. (*A–C*) Spectrograms show bat CF adjustment across three trials when the bat received spectrally altered feedback of 0, −50, and 50 cents. (*D–F*) CF adjustment of 10 calls before, 20 calls during, and 10 calls after perturbation (black circles). CF of feedback stimuli is indicated with blue circles. The general trend of call frequency (3-point smoothed) is shown with the gray line. (*G–I*) Relative change in bat call frequency to perturbation events across all trials. The dark green line shows average relative frequency. Light green areas show the SD. Gray dots show relative call frequency of individual trials. Black asterisks indicate statistical significance for nonzero call frequencies at the relative call position (signed-rank test, *P < 0.05; mean difference, >0.15 kHz). For all panels, the perturbation window is between the red dashed lines.

largest frequency adjustments. Overall, there was little difference in the maximum frequency adjustment before and after the perturbations (Fig. 3 *H* and *J*). During both the negative and positive feedback shifts, the largest frequency adjustment occurred at the medium shift sizes of ±50 cents (Fig. 3*I*). *H. armiger* also compensated more for the positive feedback shifts than the negative shifts (Fig. 3*K*) and more for the early than for the later feedback perturbations (Fig. 3*L*). The asymmetric frequency adjustments between the negative and positive feedback shifts and the dynamic frequency adjustments across the time course of feedback perturbations cannot be explained by DSC.

Sensory Error Explains Distinct Patterns of Vocal Adjustment.
Next, we applied the SFC framework to bat vocal adjustment data. Specifically, we tested whether the highly distinct patterns of frequency adjustment in individual *H. armiger* can be simulated by a Kalman filter. As described next, we successfully built a simple Kalman filter that can accurately replicate the distinct patterns of frequency adjustment across shift size, time course of feedback perturbation, and individual animals.

Fig. 4*A* illustrates the conceptual structure of the SFC model. In this model, the production of an echolocation call is achieved by a SFC law and starts with the vocal motor system issuing control signals to the larynx that outputs a call of the frequency value x_k. Shortly before the call emission, the vocal

motor system sends a copy of the call frequency information ($\hat{x}_{k|k-1}$), coded as the efference copy, to the somatosensory and auditory systems. During the call emission, both the somatosensory and the auditory systems receive the actual feedback, $z_{som,k}$ and $z_{aud,k}$, after certain delays. Subsequently, the predicted sensory feedback is combined with the actual sensory feedback to update the state estimate ($\hat{x}_k$) in an optimal way. The stream of information of higher reliability is given a larger weight. The updated state estimation is sent back to the vocal motor control system to guide the production of a next call recursively.

As indicated in the system model of Eq. **1** in Fig. 4, we modeled the frequency value for call k (x_k) as a linear combination of the frequency value from the previous call $k − 1$ (x_{k-1}) and the required frequency adjustment (u_{k-1}). To keep the number of free parameters of the model to a minimum, we specified the required frequency adjustment as the reverse of the updated state estimate based on the previous call $k − 1$ ($\hat{x}_{k-1}$), if the bat were to make a full correction to the sensory prediction error at call $k − 1$ (Eq. **2** in Fig. 4). Since *H. armiger* receives the actual sensory feedback after some delay, the current state is extended using its copies (Eq. **3** in Fig. 4) and the system is updated with the most delayed state ($\hat{x}_{k-N}$). Eqs. **4–6** in Fig. 4 describe the observation models for the somatosensory and the auditory systems, which provide actual sensory feedback to
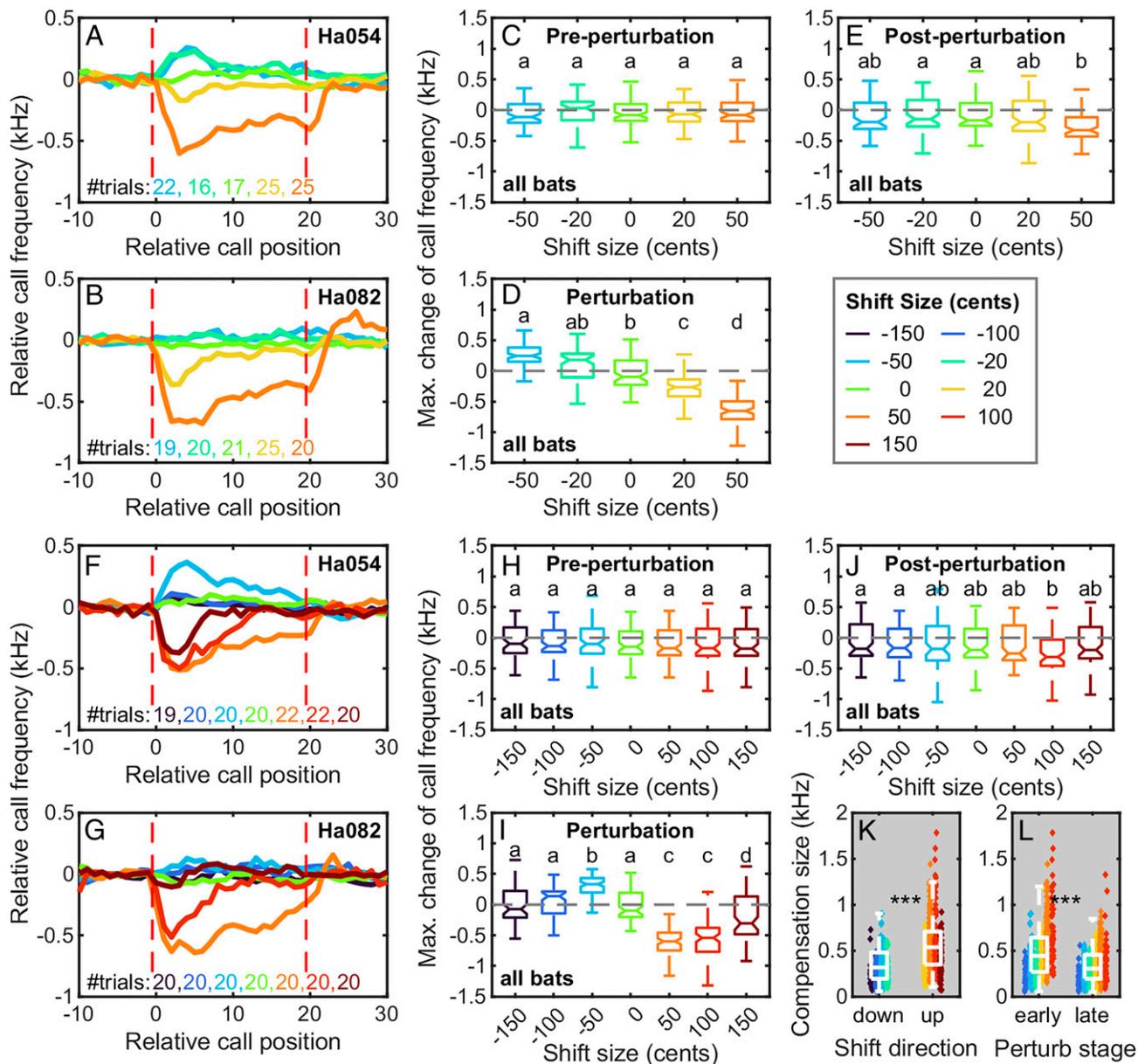
Fig. 3. Effects of shift size and direction on the frequency adjustments of individual bats. (*A* and *B*) Time course of the frequency adjustment for two representative *H. armiger* bats receiving feedback with small shift sizes (−50, −20, 0, 20, and 50 cents) and directions (up and down). (*C–E*) Boxplots of the maximum frequency adjustment of all *H. armiger* before (preperturbation), during (perturbation), and after (postperturbation) perturbations. (*F* and *G*) Time courses of frequency adjustment for two representative *H. armiger* bats receiving feedback of large shift sizes (−150, −100, −50, 0, 50, 100, and 150 cents) and directions (up and down). (*H–J*) Boxplots of maximum frequency adjustments with data from all *H. armiger* before (preperturbation), during (perturbation), and after (postperturbation) perturbations. (*K*) A statistical comparison of the frequency adjustment of all *H. armiger* between the down (negative) and up (positive) shifted feedback. (*L*) A statistical comparison of the frequency adjustment of all *H. armiger* between early (calls 1 to 10) and late perturbations (calls 11 to 20). Different lowercase letters associated with each box in panels *C*, *D*, and *H* to *J* indicate statistical significance between the two conditions. Asterisks in *K* and *L* indicate statistical significance (rank-sum test, ***P < 0.001). Boxplot in *C–E* and *H–L* shows the median and quartiles.

H. armiger after some delay. Thus, the optimization of the Kalman filter involves searching for only three free parameters: the ratio of the uncertainties between the somatosensory feedback and the prediction by the internal model ($\sigma_{som}/\sigma_{pred}$), the ratio of the uncertainties between the auditory feedback and the prediction by the internal model ($\sigma_{aud}/\sigma_{pred}$), and the baseline auditory feedback uncertainty before the perturbations [f(Δ)] for each condition. The baseline auditory feedback uncertainty reflects the estimated reliability of a bat for a given shift size of auditory feedback perturbation, which is thus individual and condition specific.

We found that the average difference per call between the predicted frequency adjustment during the perturbations by the

model and the actual frequency adjustment by *H. armiger* were 49, 49, 42, and 49 Hz in the four individuals (Fig. 4*C*). As a reference, the natural variations in call frequency, as measured by the SD, were 254 Hz (*n* = 1,544), 255 Hz (*n* = 1,684), 243 Hz (*n* = 1,185), and 299 Hz (*n* = 1,650) in the four bats, based on the data from the preperturbation window. The best model optimized for each individual independently revealed that the $\sigma_{aud}/\sigma_{som}$ ratios were all larger than 1, ranging from 1.15 and 1.53, and show that *H. armiger* relied slightly more on the somatosensory feedback than the auditory feedback for frequency control. By contrast, the much larger uncertainties of the sensory feedback systems relative to the internal model's prediction, with the $\sigma_{som}/\sigma_{pred}$ ratio between 2 and 5.2 and the
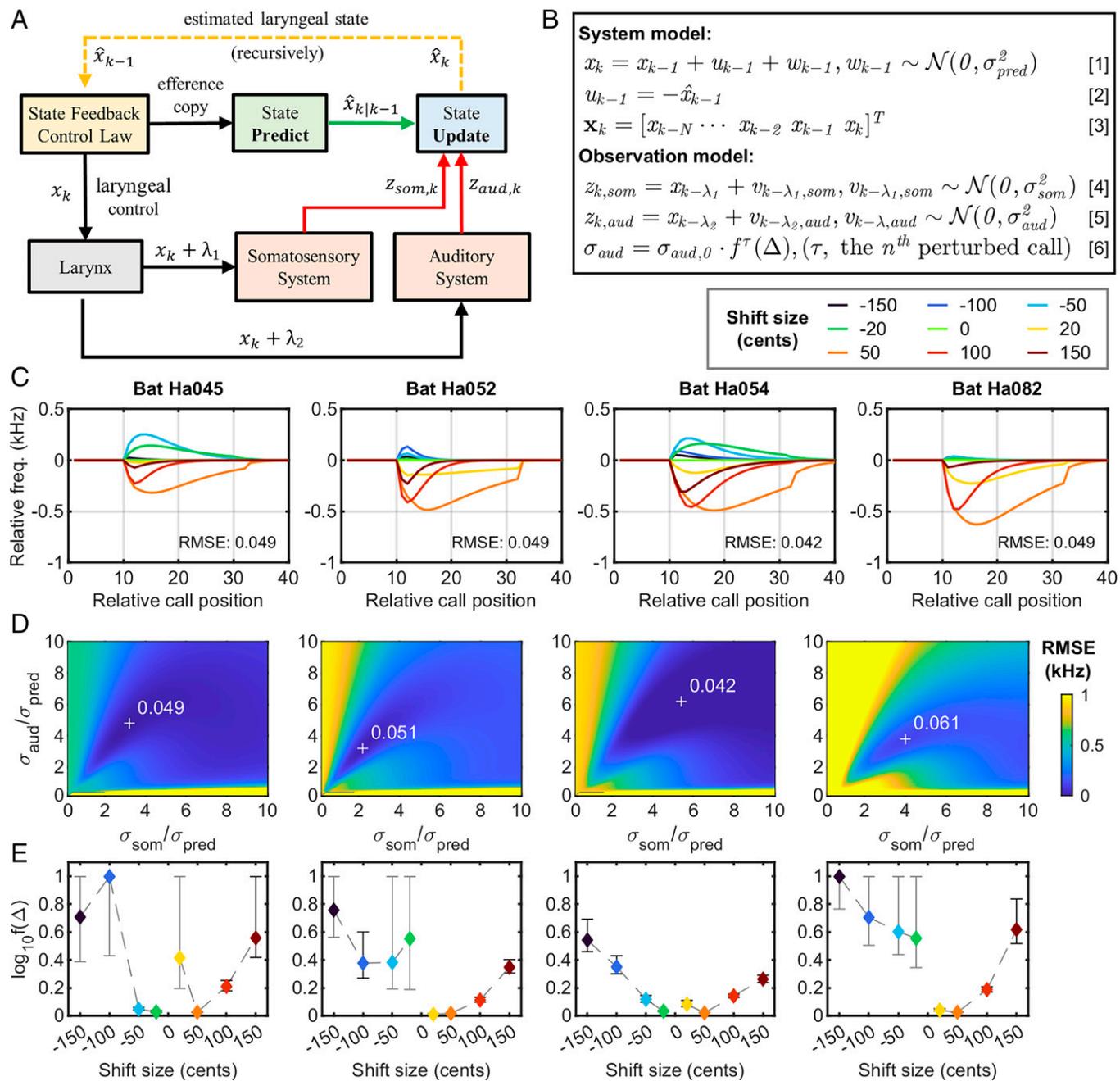
Fig. 4. An SFC model simulates the dynamic frequency adjustment of individual bats across various auditory feedback perturbation conditions. (*A*) An illustration of the structure of the SFC model. (*B*) The SFC model is computationally implemented as a Kalman filter, with the major components of the system model and observation model shown in Eqs. **1–6**. (*C*) The simulated trace of frequency adjustment for each perturbation condition and each bat with the best-optimized models. The root mean square error (RMSE) is the average frequency difference per call (in kilohertz) during the entire perturbation window across all perturbation conditions. (*D*) The relative uncertainties between the internal model prediction, the somatosensory feedback, and the auditory feedback systems. (*E*) The initial uncertainty of auditory feedback as a function of shift size. Error bar shows 1.05 times the optimized value. The somatosensory feedback delay and the auditory feedback delay were 30 and 60 ms, respectively, for the data shown in *C* to *E*.

$\sigma_{aud}/\sigma_{pred}$ ratio between 3 and 6, suggest that *H. armiger* relied strongly on the forward prediction for frequency control during the auditory feedback perturbations. Similarly, the model revealed a large difference in the uncertainty of the perturbed auditory feedback across shift sizes and individual bats (Fig. 4*E*). Lastly, we examined the effect of different feedback delays for the somatosensory system (0, 15, 30, and 45 ms) and the auditory system (0, 30, 60, and 90 ms) on the model's performance. As shown in *SI Appendix*, Figs. S2 and S3, the model can achieve similar prediction accuracy at varying combinations of somatosensory and auditory feedback delays. These findings underscore the robustness of the SFC framework.

Discussion

We implemented the human speech SFC framework to analyze audiovocal control of the echolocating bat, *H. armiger*, in response to auditory feedback perturbations. Astonishingly, the highly distinct vocal adjustments exhibited by individual bats mirrored the experimental observations of human speech control under similar perturbation conditions (21, 22). These findings reveal that bats and humans share similar computational principles for audiovocal control.

As is true for any SFC controller, our model assumed an internal model that can predict the sensory consequences of

vocal motor actions and the reliance on both somatosensory and auditory signals to update the state. Although auditory feedback is widely studied in echolocating bats, the roles of internal models and somatosensory feedback in these animals have not been well explored. One recent study by Salles and colleagues (29) offered experimental evidence for an internal model that allows echolocating bats to predict the trajectory of moving prey. Here, our model shows that the internal model prediction features a much lower uncertainty compared to that of sensory feedback, stressing a critical role of the internal model in call frequency control. At the neural level, predicted sensory feedback may be encoded by efference copy (30–33). At present, there is very limited information concerning efference copy in bats. A few studies have identified prevocal motor signals in the bat midbrain superior colliculus, a sensorimotor integration hub (34–36). Other potential efference copy mechanisms for encoding vocal motor commands in bats remain to be identified.

Our model also shows that the somatosensory feedback plays a key role in the call frequency control of echolocating bats. The observation that the baseline uncertainty of somatosensory feedback is slightly lower than the baseline uncertainty of the auditory feedback contrasts sharply with the general view that bats rely primarily on auditory feedback to perform echolocation tasks. The contribution of nonauditory sensory information to bat echolocation has been underestimated in general. Apart from a few studies that investigated the role of vision in modulating echolocation behavior of bats (37–39), there are scarce data on the potential role of other sensory modalities in bat sonar navigation. Although there is evidence for the contribution of somatosensory feedback to bat flight control (40, 41), the role of somatosensory feedback in bat vocal control has not been studied despite its long-suggested importance in mammals (26). One piece of indirect evidence comes from a recent study, which reports that horseshoe bats are capable of performing partial DSC when auditory feedback is eliminated (42). It is noteworthy that the indispensable role of somatosensory feedback in human speech control has been experimentally demonstrated repeatedly (43, 44). For a better understanding of the remarkable navigation and orientation behaviors of bats, researchers should fully explore the role of sensory modalities outside of hearing.

All four *H. armiger* bats showed relatively low auditory feedback uncertainty for most upward shifted sounds compared with the downward shifted sounds, with the lowest auditory feedback uncertainty occurring at the 50-cent condition. This may reflect biases that have evolved to support DSC behavior of *H. armiger* in flight, where they typically receive upward shifted echoes arising from their own movement. The flight speeds that result in Doppler shifts corresponding to the 20-, 50-, 100-, and 150-cent shift sizes are 2, 4.9, 9.8, and 14.7 m/s, respectively. Thus, the largest frequency adjustment at the 50-cent condition would be induced by a flight speed of 4.9 ms/s, which closely matches the maximum flight speed of *H. armiger* in the laboratory at ~4.7 m/s (16). *H. armiger* typically forages close to dense vegetation and flies slowly close to foliage (45). It is also known that bats tend to fly more slowly in the laboratory than in the field. Taken together, we believe that the 50-cent shift size represents more natural Doppler shifts experienced by *H. armiger* in the field and therefore gives rise to the most reliable (lowest uncertainty) auditory feedback.

Conclusions

The present study discovered that bat sonar calls and human speech, vocalizations serving very distinct functions, operate on similar computational principles of real-time audiovocal control. We identified sensory error as a fundamental mechanism driving dynamic audiovocal adjustments in the bat species *H. armiger*. It remains to be determined whether similar SFC models can account for vocal frequency adjustments in other species of bats that exhibit DSC, such as horseshoe bats and mustached bats. It is also unknown whether sensory error correction accounts for frequency control in bats without DSC capability. Notably, our research demonstrates that echolocating bats can serve as powerful mammalian models for understanding sensorimotor mechanisms of human speech control.

Materials and Methods

Animals. We tested four adult bats; all were female due to constraints of field collection. We caught the bats with a hand net during the daytime in September 2020, at the Baini Cave in Chongyang County, Xianning, Hubei Province, China. We housed the bats in social groups of two to four individuals. We placed their roosting cages in a room (4.4 m length × 3 m width × 2.2 m height) with a regulated air temperature of around 24 °C, relative humidity of about 60%, and a reversed light regime of 12 h of darkness and 12 h of light. Bats had free access to water and food. The Institutional Animal Care and Use Committee of the Central China Normal University approved the capture, housing, and behavioral studies.

Real-Time Auditory Feedback Perturbation System. We built a programmable auditory feedback perturbation system to deliver spectrally altered auditory feedback to vocalizing bats in real time (Fig. 1A). The setup included a measurement microphone, a loudspeaker, and a real-time DSP frequency-shifting module programmed in LabVIEW with field-programmable gate array (FPGA) chips (PXIe-7858R; National Instruments). The DSP unit shifts the frequency of a microphone recording sample by sample up to 1 MHz. The system performed frequency shifts without affecting the fine temporal structure of the echolocation calls (Fig. 1B). With the current program design, for a single trial, we can achieve a fixed frequency offset between the produced calls of the bat and the feedback stimuli at a programmable perturbation delay. We can also repeat the perturbation many times–that is, controlling the times of feedback perturbations. As shown in Fig. 1 C–E, postanalyses of the feedback stimuli confirmed the quality of the perturbation system for both the shift size and perturbation delay. Nevertheless, the minimum DSP delay increased slightly from 0.58 ± 0.4 ms for the 0-ms programmed delay to 0.81 ± 0.5 ms for the 24-ms programmed perturbation delay (Fig. 1E). Given the electronic delay of our playback system, combined with a 1 ms acoustic delay for sound travel from the bat to the microphone and from the loudspeaker to the bat, the minimum feedback delay is about 1.6 ms, which is approximately one-sixth of a typical echolocation call of *H. armiger*.

The frequency-shifting algorithm is a time-domain technique that compresses or expands the estimated period of the input signal and has been tested with FPGA hardware previously (46). The largest shift size that we can achieve is greater than 30 kHz for an input signal of around 70 kHz, which was experimentally tested with the echolocation calls of *H. armiger*. To facilitate the communications between the bat community and the speech community, we specified the shift size in cents, which describes frequency changes at a logarithmic scale. To convert the frequency change in cents into hertz, one can follow this equation: $Hz = (2^{\frac{cents}{1200}} - 1) \times F_{ref}$. Thus, when a 70-kHz signal is shifted up by 50 cents, the corresponding frequency increase is 2.05 kHz (Fig. 1C).

Recording, Playback, and Sound Analysis. We recorded calls produced by resting *H. armiger* individuals hanging from an elevated platform 2 m high. We sampled the input signal with a measurement microphone (7016, 1/4-inch condenser with protection grid on; ACO Pacific), amplified by 40 dB (OctaMic II; RME Audio Interfaces) before the analog-to-digital conversion. We filtered the samples with a 128-point finite impulse response digital filter with 20- and 80-kHz cutoff frequencies. Then we sent a copy of the digitized samples to the frequency-shifting algorithm, sample by sample.

Experiments took place in a large flight room (6.5 m length × 5 m width × 2.3 m height). We covered the walls and ceiling of the flight room with 8-cm-thick

acoustic foam and the floor with nylon blankets to reduce echo reflections. We divided each trial into three temporal windows (Fig. 2A): 1) preperturbation (1 or 2 s), 2) perturbation (variable time required for the bat to produce 20 echolocation calls), and 3) postperturbation (3 s). During the preperturbation and postperturbation windows, we set the amplitude of the playback system to zero.

We streamed a copy of the digitized recording and a copy of the output signals directly to the hard drive. The output from the frequency shifting was then amplitude scaled, digital-to-analog converted, and passed to the power amplifier (UltaSoundGate Player with the power adapter option; Avisoft Bioacoustics) and loudspeaker (Vifa; Avisoft Bioacoustics). We adjusted the digital gain of the input and output so the maximum peak amplitude of the playback signal was around 1 V, resulting in a 94-dB sound pressure level (relative to 20 μPa) amplitude at the bat's position. Both echolocation calls and playback signals were processed by customer-built sound analysis scripts in MATLAB that were described in a previous study (47).

Computational Simulation. We used a Kalman filter to simulate the highly distinct patterns of frequency adjustment in individual *H. armiger*. This computational

algorithm for the SFC framework simulated the frequency control of human speech (23, 25, 48). The Kalman filter used a two-stage process: 1) estimating the current state using an internal model and 2) updating the state estimation given sensory feedback. The goal was to identify the bats' computational principles for transforming spectrally altered feedback into call frequency control. Our Kalman filter accurately replicated distinct patterns of frequency adjustment across the shift sizes, multiple perturbations, and individual animals. Further details of the computational models are presented in *SI Appendix*.

1. R. Shadmehr, S. Mussa-Ivaldi, *Biological Learning and Control: How the Brain Builds Representations, Predicts Events, and Makes Decisions* (MIT Press, Cambridge, MA, 2012).
2. D. W. Franklin, D. M. Wolpert, Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442 (2011).
3. F. H. Guenther, *Neural Control of Speech* (MIT Press, 2016).
4. W. T. Fitch, The biology and evolution of speech: A comparative analysis. *Annu. Rev. Linguist.* **4**, 255–279 (2018).
5. E. D. Jarvis, Evolution of vocal learning and spoken language. *Science* **366**, 50–54 (2019).
6. C. F. Moss, A. Surlykke, Probing the natural scene by echolocation in bats. *Front. Behav. Neurosci.* **4**, 33 (2010).
7. C. F. Moss, C. Chiu, A. Surlykke, Adaptive vocal behavior drives perception by echolocation in bats. *Curr. Opin. Neurobiol.* **21**, 645–652 (2011).
8. T. K. Jones, K. M. Allen, C. F. Moss, Communication with self, friends and foes in active-sensing animals. *J. Exp. Biol.* **224**, jeb242637 (2021).
9. E. H. Gillam, N. Ulanovsky, G. F. McCracken, Rapid jamming avoidance in biosonar. *Proc. Biol. Sci.* **274**, 651–660 (2007).
10. J. Luo, C. F. Moss, Echolocating bats rely on audiovocal feedback to adapt sonar signal design. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10978–10983 (2017).
11. J. Luo, N. B. Kothari, C. F. Moss, Sensorimotor integration on a rapid time scale. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6605–6610 (2017).
12. H. U. Schnitzler, Control of doppler shift compensation in the greater horseshoe bat, *Rhinolophus ferrumequinum*. *J. Comp. Physiol.* **82**, 79–92 (1973).
13. H. U. Schnitzler, A. Denzinger, Auditory fovea and Doppler shift compensation: Adaptations for flutter detection in echolocating bats using CF-FM signals. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **197**, 541–559 (2011).
14. S. Hiryu, E. C. Mora, H. Riquimaroux, Bat bioacoustics in *Springer Handbook of Auditory Research*, M. B. Fenton, A. D. Grinnell, A. N. Popper, R. R. Fay, Eds. (Springer, New York, 2016), pp. 239–263.
15. H.-U. Schnitzler, Die Uhraschall-Ortungslaute der Hufeisen-Fledermaeuse (Chiroptera-Rhinolophidae) in verschiedenen Orientierungssituationen. *Z. Vgl. Physiol.* **57**, 376–408 (1968).
16. D. Schoeppler, H.-U. Schnitzler, A. Denzinger, Precise Doppler shift compensation in the hipposiderid bat, *Hipposideros armiger*. *Sci. Rep.* **8**, 4598 (2018).
17. Y. Zhang et al., Performance of Doppler shift compensation in bats varies with species rather than with environmental clutter. *Anim. Behav.* **158**, 109–120 (2019).
18. H. U. Schnitzler, E. K. V. Kalko, Echolocation by insect-eating bats. *Bioscience* **51**, 557–569 (2001).
19. J. F. Houde, M. I. Jordan, Sensorimotor adaptation in speech production. *Science* **279**, 1213–1216 (1998).
20. J. F. Houde, S. S. Nagarajan, Speech production as state feedback control. *Front. Hum. Neurosci.* **5**, 82 (2011).
21. T. A. Burnett, M. B. Freedland, C. R. Larson, T. C. Hain, Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* **103**, 3153–3161 (1998).
22. C. R. Larson, T. A. Burnett, S. Kiran, T. C. Hain, Effects of pitch-shift velocity on voice Fo responses. *J. Acoust. Soc. Am.* **107**, 559–564 (2000).
23. B. Parrell, J. Houde, Modeling the role of sensory feedback in speech motor control and learning. *J. Speech Lang. Hear. Res.* **62**, 2963–2985 (2019).
24. G. Hickok, J. Houde, F. Rong, Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* **69**, 407–422 (2011).
25. J. F. Houde, C. Niziolek, N. Kort, Z. Agnew, S. S. Nagarajan, "Simulating a state feedback model of speaking," in *Proceedings of the 10th International Seminar on Speech Production*, S. Fuchs, M. Grice, A. Hermes, L. Lancia, D. Mücke, Eds. (University of Cologne, 2014), pp. 202–205.
26. M. S. Smotherman, Sensory feedback control of mammalian vocalizations. *Behav. Brain Res.* **182**, 315–326 (2007).
27. M. Kössl, J. Hechavarria, C. Voss, M. Schaefer, M. Vater, Bat auditory cortex–Model for general mammalian auditory computation or special design solution for active time perception? *Eur. J. Neurosci.* **41**, 518–532 (2015).
28. M. J. Wohlgemuth, J. Luo, C. F. Moss, Three-dimensional auditory localization in the echolocating bat. *Curr. Opin. Neurobiol.* **41**, 78–86 (2016).
29. A. Salles, C. A. Diebold, C. F. Moss, Echolocating bats accumulate information from acoustic snapshots to predict auditory object motion. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 29229–29238 (2020).
30. H. Straka, J. Simmers, B. P. Chagnaud, A new perspective on predictive motor signaling. *Curr. Biol.* **28**, R232–R243 (2018).
31. M. A. Sommer, R. H. Wurtz, A pathway in primate brain for internal monitoring of movements. *Science* **296**, 1480–1482 (2002).
32. J. Luo, S. R. Hage, C. F. Moss, The Lombard effect: From acoustics to neural mechanisms. *Trends Neurosci.* **41**, 938–949 (2018).
33. S. J. Eliades, X. Wang, Corollary discharge mechanisms during vocal production in marmoset monkeys. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **4**, 805–812 (2019).
34. S. R. Sinha, C. F. Moss, Vocal premotor activity in the superior colliculus. *J. Neurosci.* **27**, 98–110 (2007).
35. N. B. Kothari, M. J. Wohlgemuth, C. F. Moss, Dynamic representation of 3D auditory space in the midbrain of the free-flying echolocating bat. *eLife* **7**, e29053 (2018).
36. M. J. Wohlgemuth, N. B. Kothari, C. F. Moss, Functional organization and dynamic activity in the superior colliculus of the echolocating bat, *Eptesicus fuscus*. *J. Neurosci.* **38**, 245–256 (2018).
37. S. Danilovich et al., Bats regulate biosonar based on the availability of visual information. *Curr. Biol.* **25**, R1124–R1125 (2015).
38. S. Danilovich, Y. Yovel, Integrating vision and echolocation for navigation and perception in bats. *Sci. Adv.* **5**, eaaw6503 (2019).
39. T. K. Jones, C. F. Moss, Visual cues enhance obstacle avoidance in echolocating bats. *J. Exp. Biol.* **224**, 241968 (2021).
40. S. Sterbing-D'Angelo et al., Bat wing sensors support flight control. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11291–11296 (2011).
41. C. Geberl, S. Brinkløv, L. Wiegrebe, A. Surlykke, Fast sensory-motor reactions in echolocating bats to sudden changes during the final buzz and prey intercept. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4122–4127 (2015).
42. A. Boonman et al., Echolocating bats can adjust sensory acquisition based on internal cues. *BMC Biol.* **18**, 166 (2020).
43. C. R. Larson, K. W. Altman, H. Liu, T. C. Hain, Interactions between auditory and somatosensory feedback for voice F0 control. *Exp. Brain Res.* **187**, 613–621 (2008).
44. F.-X. Brajot, D. Nguyen, J. DiGiovanni, V. L. Gracco, The impact of perilaryngeal vibration on the self-perception of loudness and the Lombard effect. *Exp. Brain Res.* **236**, 1713–1723 (2018).
45. U. M. Norberg, J. M. V. Rayner, Ecological morphology and flight in bats (Mammalia; Chiroptera): Wing adaptations, flight performance, foraging strategy and echolocation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **316**, 335–427 (1987).
46. H. Estephan, S. Sawyer, D. Wanninger, "Real-time speech pitch shifting on an FPGA" (Villanova University, 2006).
47. M. Lu, G. Zhang, J. Luo, Echolocating bats exhibit differential amplitude compensation for noise interference at a sub-call level. *J. Exp. Biol.* **223**, 225284 (2020).
48. B. Parrell, V. Ramanarayanan, S. Nagarajan, J. Houde, The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS Comput. Biol.* **15**, e1007321 (2019).
49. H. Wang et al., Sensory error drives fine motor adjustment. Zenodo. https://doi.org/10.5281/zenodo.6645444. Deposited 15 June 2022.