Behavioral Ecology

The official journal of the
ISBE
International Society for Behavioral Ecology

Behavioral Ecology (2022), XX(XX), 1–8. https://doi.org/10.1093/beheco/arac053

Original Article

Multisensory integration facilitates perceptual restoration of an interrupted call in a species of frog

Bicheng Zhu,[a] Yue Yang,[a] Ya Zhou,[a] Ke Deng,[a,©] Tongliang Wang,[b] Jichao Wang,[b] Yezhong Tang,[a] Michael J. Ryan,[c,©] and Jianguo Cui[a,©]

[a]CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization and Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, Sichuan, China, [b]Ministry of Education Key Laboratory for Ecology of Tropical Islands, Key Laboratory of Tropical Animal and Plant Ecology of Hainan Province, College of Life Sciences, Hainan Normal University, Haikou 571158, Hainan, China, and [c]Department of Integrative Biology, University of Texas, University Station C0930, Austin, TX 78712, USA

Received 15 October 2021; revised 22 April 2022; editorial decision 1 May 2022; accepted 6 May 2022

Communication signals by both human and non-human animals are often interrupted in nature. One advantage of multimodal cues is to maintain the salience of interrupted signals. We studied a frog that naturally can have silent gaps within its call. Using video/audio-playbacks, we presented females with interrupted mating calls with or without a simultaneous dynamic (i.e., inflating and deflating) vocal sac and tested whether multisensory cues (noise and/or dynamic vocal sac) inserted into the gap can compensate an interrupted call. We found that neither inserting white noise into the silent gap of an interrupted call nor displaying the dynamic vocal sac in that same gap restored the attraction of the call equivalent to that of a complete call. Simultaneously presenting a dynamic vocal sac along with noise in the gap, however, compensated the interrupted call, making it as attractive as a complete call. Our results demonstrate that the dynamic visual sac compensates for noise interference. Such novel multisensory integration suggests that multimodal cues can provide insurance against imperfect sender coding in a noisy environment, and the communication benefits to the receiver from multisensory integration may be an important selective force favoring multimodal signal evolution.

Key words: **acoustic composition, audio-visual integration, frog, multimodal signal, mate choice.**

INTRODUCTION

In recent years, multimodal communication has become a compelling interest, especially in behavioral ecology (Hebets and Papaj 2005; Caldart et al. 2021). Multimodal displays are composed of signals and cues in more than one sensory modality and are widespread in humans and other animals (Halfwerk et al. 2019). An effective method to probe possible functions for a multimodal signal is addressing questions about signal content, efficacy, and inter-signal interactions (Hebets and Papaj 2005). While previous studies of multimodal communication have often tested whether different components serve as back-up messages or provide multiple meanings (Partan and Marler 2005), few studies focus on the inter-signal interactions of multimodal communication and how specific interactions across multimodal components influence the receivers' behavioral responses (Halfwerk et al. 2019). Different signal components of multisensory signals often interact so that one signal can change the receiver's response to another signal (Bahrick et al. 2004; Caldart et al. 2021). One impressive example is the McGurk effect, which shows that mismatching lip movements and auditory cues during speech can change our acoustic perceptions (McGurk and Macdonald 1976).

Acoustic signals produced by most frogs are accompanied by the inflation and deflation of a conspicuous vocal sac (Narins et al. 2003; Taylor et al. 2008; Zhu et al. 2021). The acoustic component of a frog's call is its dominant characteristic, yet visual components (e.g., dynamic vocal sac) can also influence signal attraction (Stange

Address correspondence to J. Cui. E-mail: cuijg@cib.ac.cn

et al. 2017; Kaiser et al. 2018; Zhu et al. 2021). Many frogs maintain inflation of their vocal sacs in between calls when they pause to call, suggesting a communication function. Although most frogs are nocturnal, previous studies demonstrated that some frogs have excellent behavioral and retinal visual sensitivities under nocturnal conditions and can perceive the vocal sac (Cummings et al. 2008; Leslie et al. 2020). We now know that this visual cue can function as a component of a male's multimodal sexual display in some species (Taylor et al. 2008; Starnberger et al. 2014; Zhu et al. 2021). Similar to the movement of lips when speaking, which can modify our perception of auditory information (e.g., the McGurk effect), dynamic (inflating-deflating) vocal sacs can also influence frogs' behavioral responses (Narins et al. 2003; Taylor and Ryan 2013). One study of frogs has shown the potential of multimodal cues to rescue saliency in a frog's call (the multisensory rescue hypothesis), although in a different context than we are studying here. Female túngara frogs (*Physalaemus pustulosus*) perceived two relatively unattractive signals (an acoustic signal and a visual signal) as an attractive compound signal when the two unattractive signals were combined (Taylor and Ryan 2013). As vocal communication in frogs often takes place in intraspecific and/or interspecific choruses, the signal received by receivers is often incomplete (Hödl and Amézquita 2001; Bee and Micheyl 2008). Thus, how integrating audio-visual cues to compensate for an interrupted call becomes a common challenge faced by frogs and other animals relying on vocal communication.

In addition, it is well-known that calling rate (the number of calls in a call bout per unit time) and specific call parameters (e.g., dominant frequency and duration) are key assessment indicators in mate selection and often determine female preference (Schwartz et al. 2011; Henderson and Gerhardt 2013; Tanner and Bee 2020). However, these conclusions are obtained only in the case of unimodal communication. Considering the potential inter-signal interactions between multimodal components, the roles of calling rate and call parameters in mate choice may change in the presence of a visual cue.

As with most frogs, male serrate-legged small treefrogs (*Kurixalus odontotarsus*) produce multimodal courtship calls by vocalizing and presenting their inflating and deflating vocal sac as a visual cue. Male frogs of this species usually call at night, but a few occasionally call during the day (Zhu et al. 2021). Female frogs base their mate choices on male advertisement calls, which usually consist of five wideband frequency notes (Zhu et al. 2017a, 2017b). Interestingly, we found that males produce advertisement calls with silent gaps in the wild (Supplementary Figure S1). Male frogs produce more interrupted calls when another male is calling nearby (217 interrupted calls in 306 min of recordings from 51 male frogs), even though a single male rarely produces the interrupted calls when there are no competitors around (25 calls in 288 minutes of recordings from 67 male frogs). Meanwhile, our field observation found that some male serrate-legged small treefrogs occasionally inflate the vocal sac without vocalization, which suggests that the role of the vocal sac itself could be important in visual communication. However, whether adding a dynamic vocal sac with an interrupted call can restore the attractiveness of the interrupted call remains unknown.

Video/audio-playbacks have proven to be an effective method to study the perception of multimodal signals in some nocturnal anurans (Rosenthal et al. 2004; Zhu et al. 2021). Using video as well as audio playbacks, we presented female serrate-legged small treefrogs with manipulated acoustic stimuli and/or a simultaneous dynamic (i.e., inflating and deflating) vocal sac and tested whether the attractiveness of interrupted calls can be restored to that of a complete call by inserting into the gap of the call: 1) an acoustic cue, 2) a visual cue, and 3) a simultaneous multimodal cue. In addition, we tested whether the roles of calling rate and call parameters in mate choice change in the presence of a visual cue.

MATERIALS AND METHODS

Study site

We completed all experiments in Hainan, China at the Mt. Diaoluo National Nature Reserve during the breeding season of *K. odontotarsus* in 2015, 2016, and 2018; we collected female frogs from the same metapopulation. Gravid females were captured by hand in the field and placed in individual light-safe boxes which contained water and branches with foliage. We kept female frogs in the tanks for an average of one hour between collecting and testing.

Stimulus preparation

We filmed calling males at night (temperature 21.8 ± 0.76°C) to make stimuli using a digital camera as we did in our previous study (Zhu et al. 2021). Males are often observed vocalizing near streetlights or beside buildings that have artificial light. Our previous data revealed that the light intensity of male calling sites under artificial light varied from 6.98 to 55.15 lux (Deng et al. 2019). The light intensity measured at the male's calling site in the present study was about 4.40 lux.

We chose a five-note advertisement call as the basic stimulus, which is common in male calls (Zhu et al. 2017a, 2017b). The acoustic stimuli tested included the complete five-note advertisement call, "complete call," five-note advertisement call with the gap of silence (dropping the second note), "silent gap," and the advertisement call with the gap of silence filled with a broadband white noise (1–10 kHz), "noise gap" (Figure 1). The "fast" root-mean-square amplitude of acoustic stimuli was 80 dB SPL (re: 20 μPa, Z-weighted) measured with a sound level meter (AWA 6291, Hangzhou Aihua, China) at the release point of the female. All stimuli have an equivalent call duration, 2 s. To display multimodal signals, we placed an LCD monitor with a large visual angle (170°) above each speaker. The audio and the video were from the same frog. We used six audio-visual stimuli derived from six different calling males to minimize pseudoreplication. More details are shown in Supplementary Figure S2.

Our experiment includes five tests: complete call vs. silent gap (test a); complete call vs. noise gap (test b); silent gap plus dynamic vocal sac vs. silent gap (test c); complete call plus dynamic vocal sac vs. silent gap plus dynamic vocal sac (test d); and complete call plus dynamic vocal sac vs. noise gap plus dynamic vocal sac (test e; see Figure 3). Each stimulus pair was presented in a loop, and the intervals between two stimuli were 5 s. We displayed a background video of a male frog's calling site (as if they were blocked by leaves in the tropical rainforest) on the monitor as a control when the acoustic-only stimuli were playing to eliminate the possibility that females were merely attracted by light/images from the screens (phototaxis). There are many situations in the wild when the female can hear a male frog but not see it, because variations in foliage or geography may obscure male callers visually. Hence, our control design was consistent with natural conditions.

We adjusted the parameters (including gamma, brightness, contrast ratio, and color balance) of the two screens using the color calibration function of Windows to ensure the parameters were the same. Meanwhile, we adjusted the color and brightness of the screens using

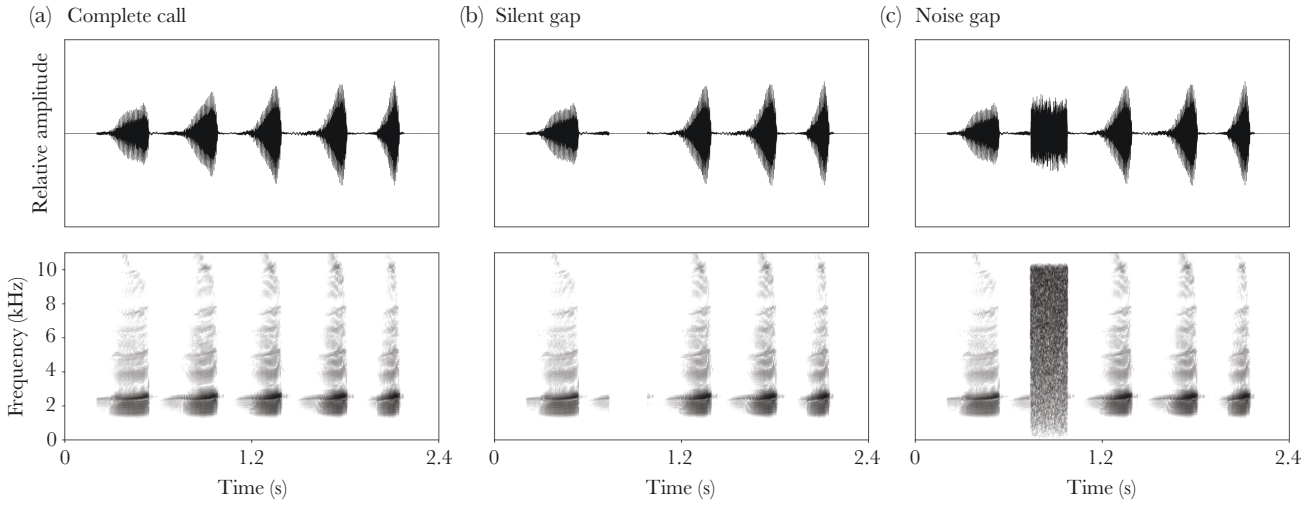

Figure 1
Amplitude-modulated waveforms (top) and spectrograms (bottom) of three acoustic stimuli. (a) Complete five-note advertisement call, "complete call"; (b) advertisement call with the gap of silence, "silent gap"; (c) advertisement call with the gap of silence filled with white noise, "noise gap." The FFT (fast Fourier transform) frame is 1024.

a colorimeter (Spyder 5, Datacolor, China) and a screen luminance meter (SM208, Sanpometer, China) respectively. The brightness of one screen was 75.5 cd/m^2, and that of the other screen was 77.1 cd/m^2. The light intensity measured at the release point (1 m away from the screen) was 0.06 lux, and 8.59 lux at 10 cm away from both screens, which coincides with the intensity of nocturnal light in the wild (range: 2.1–55.15 lux) where serrate-legged small treefrogs occur, especially considering that few male frogs call during daylight.

As mentioned in previous studies (Reichert and Höbel 2015; Zhu et al. 2021), it is difficult to present stimuli that replicate the visual features of objects in nature. The spectra of nocturnal environments vary widely and treefrogs experience variable spectra depending on different breeding habitats. Our field observation also found that the body color of male *K. odontotarsus* varies greatly. We measured the reflectance of the back and vocal sac from six individuals using a Jaz spectrometer (Ocean Optics Inc., Dunedin, FL) with a PX2 light source. Four parts of the male back and three parts of the vocal sac were measured in each frog. The reflectance of the vocal sac is higher than that of the male back (Supplementary Figure S3). The reflectance of the male vocal sac does not vary much in the wavelength range of 400–700 nm, even though it has the maximum reflectance at 670 nm (Supplementary Figure S3). Previous data obtained from video-playback experiments in túngara frogs (Rosenthal et al. 2004) were statistically indistinguishable from the results of analogous experiments using robotic frogs (Taylor et al. 2008). Our previous study also demonstrates that the video-playback approach is effective in *K. odontotarsus* (Zhu et al. 2021).

Female testing

All experiments were conducted indoors, allowing control of ambient noise, light, and weather conditions. Each frog completed the binary phonotaxis tests in a sound-attenuating chamber (150 × 150 × 120 cm) under infrared illumination (84H10P, Woshida, China) between 20:00 and 02:00 h (temperature: 22.7 ± 1.1°C, relative humidity: 86.3 ± 4.9%). The testing arena was a 1-m equilateral triangle with two audiovisual devices occupying two apexes, in which the female was released at the third apex (Figure 2).

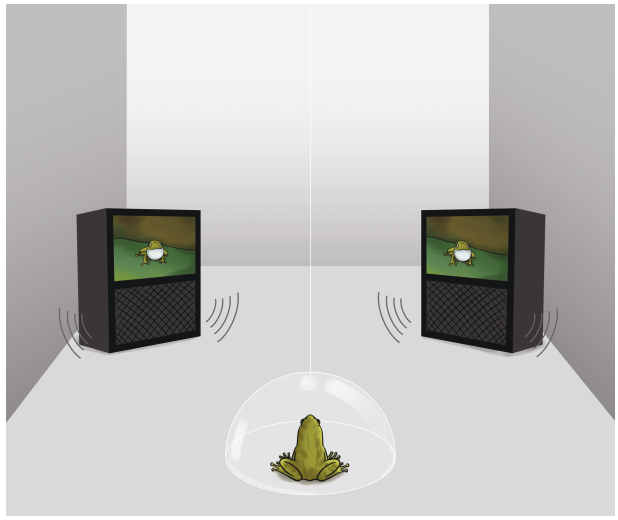

Figure 2
The diagram of the phonotaxis test arena. We restrained a female frog under a transparent dome that was placed between two audiovisual devices, which were separated by 100 cm and formed a triangle with a 60° separation relative to the female frog's release point in a 150 × 150 cm sound-attenuating chamber. The speakers play sounds, and the monitors above the speakers play videos of dynamic vocal sacs. Each audiovisual device can present a unimodal (acoustic or visual) or a synchronized multimodal (acoustic and visual) stimulus independently.

Before the test, each subject was allowed to remain under a visually and acoustically transparent dome (12.8 cm in diameter) for 2 min to acclimate, during which time both speakers played complete advertisement calls. Then, the dome was raised to allow the female to choose within 10 min. The test stimuli were played antiphonally. The speaker/monitor that broadcast the multimodal stimulus was determined randomly to avoid a potential order effect. Randomized orders were generated using the Randbetween function in Excel. We defined a choice being made when a frog entered the choice zone (within 10 cm of the speaker) and remained in the

zone for 3 s, or if they touched the speaker/monitor. Female frogs that did not enter into the zone within 10 min were later retested after enough time to rest (usually 15–30 min). Females that failed to respond twice were removed from the experiment. These frogs failed to respond presumably due to a lack of motivation. For each test, we also recorded the time female frogs took to make a choice (latency).

Each female was tested only once in a specific stimulus pair. For trials a, b, and e, females were tested in only one trial; for instance, frogs tested in trial e were not used in other trials. That was the same for frogs participated in trial a or b. Some females (i.e., 26) were tested in both trials c and d. However, there is no inter-group comparison between trials c and d. Each trial was designed to address a different question. To avoid fatigue, each frog was allowed a 3-min break after each test. To eliminate possible chemical odors, we mopped the ground between the tests to keep the arena clean and moist.

After completing all tests, we uniformly clipped part of the soft tissue (i.e., sucker) from the frog's smallest fingertip to prevent the recapture of the same subjects. Meanwhile, we sprayed disinfectant (Vetericyn Plus $^{+}$ Wound and Skin Care; Innovacyn Inc.) on the slight wound to avoid infection and relieve pain after the sucker-clipping. All operations followed the Guidelines for the Use of Live Amphibians and Reptiles in Field Research (Beaupre 2004). Our constant observations found that such treatment did not affect female activities (e.g., climbing and ovulation) and survivorship in the field. Many frogs were recaptured after 1 or 2 months. The regrown sucker is smaller than the original and easy to distinguish after about 2–3 months. We put the original paired female and male frogs together after the tests and they reformed into pairs very quickly. We released all female frogs immediately in their home location after testing.

Statistical analysis

We verified female choice responses, the number of visits as well as the latency to choose, by scoring the surveillance videos. The sonograms of male advertisement calls were drawn using PRAAT (Boersma 2002). The statistical graphs were visualized using Origin 2017 software (OriginLab Corp.). All statistical analyses were conducted using SPSS 21.0 software (SPSS Inc.). The exact binomial test was used to analyze female choice within each test and Fisher's exact test (two-tailed) was used to compare the results between different tests; we did not compare results between tests c and d. The generalized linear mixed model (GLMM) was utilized to analyze the difference in female latency to choose. We created a model with a Poisson error structure and log-link function. Specifically, we set the stimulus pair type (complete call vs. silent gap, complete call vs. noise gap, silent gap plus visual sac vs. silent gap, complete call plus visual sac vs. silent gap plus visual sac, complete call plus visual sac vs. noise gap plus visual sac) as a fixed effect. The female ID term provides information on how repeatable and individual-specific the behavior is. Thus, we modeled the female ID as a random term in the model. Pairwise comparisons with the estimated marginal means contrasts were used to complete multiple comparisons (adjust for multiple comparisons using the least significant difference). All data were examined for assumptions of normality and homogeneity of variance, using the Shapiro–Wilk and Levene tests, respectively. Data are expressed as Mean ± SD, and $p < 0.05$ was considered to be statistically significant.

Ethics

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were approved by the Animal Care and Use Committee of Chengdu Institute of Biology, CAS (CIB 2016008).

RESULTS

Preference responses

A total of 234 female frogs participated in this series of experiments. Female frogs preferred a complete advertisement call over an advertisement call with a gap of silence (complete call vs. silent gap, 68 vs. 39, $P = 0.0065$; Figure 3a).

Inserting broadband white noise in the gap did not restore the interrupted call's attractiveness to that of a complete call (complete call vs. noise gap, 44 vs. 23, $P = 0.0139$; Figure 3b).

We presented females with an interrupted call in which the vocal sac was synchronized with all the call notes and also occurred in the gap of the interrupted call. This call was more attractive than the same interrupted call that was not accompanied by the dynamic vocal sac (silent gap plus vocal sac vs. silent gap, 23 vs. 7, $P = 0.0052$, Figure 3c). However, the dynamic vocal sac did not restore the interrupted call as attractive as the complete call with the dynamic vocal sac (complete call plus vocal sac vs. silent gap plus vocal sac, 22 vs. 4, $P = 0.0005$, Figure 3d).

When the complete call and the interrupted call were played back alternately, 36.4% of the females chose the interrupted call, but the ratio of the females chose the interrupted call dropped to 15.4% when presenting the vocal sac in both monitors, although the difference between the two groups was not statistically significant (36.4% vs. 15.4% choices, $P = 0.08$, Figure 3a and d).

We gave females a choice between the complete call accompanied by a dynamic vocal sac versus the interrupted call accompanied by a dynamic vocal sac, as in the last experiment, but with the addition of white noise in the gap. Half of the female frogs chose the interrupted call with a gap of silence filled with noise over the complete call (complete call plus vocal sac vs. noise gap plus vocal sac, 15 vs. 15, $P = 1.0$, Figure 3e). Thus, the attractiveness of the interrupted call is restored to that of the complete call when white noise occurs in the gap and a dynamic vocal sac accompanies the call.

Latency to choose

We analyzed and compared female latency to choose from five tests. Overall, there are significant differences in female latency between different tests (GLMM: $F_{4, 255} = 19.301$, $P < 0.001$; Figure 4, Table 1). Pairwise comparisons showed that females exhibited significantly shorter latencies in the tests containing a multimodal stimulus than in tests that only contained acoustic stimuli ($P < 0.001$, Figure 4, Table 2).

DISCUSSION

In this study, we examine how multimodal cues might rescue the saliency of signals in a noisy environment. Specifically, we determined how the presence of white noise and the visual cue of an inflating-deflating vocal sac can restore the attractiveness of interrupted mating calls.

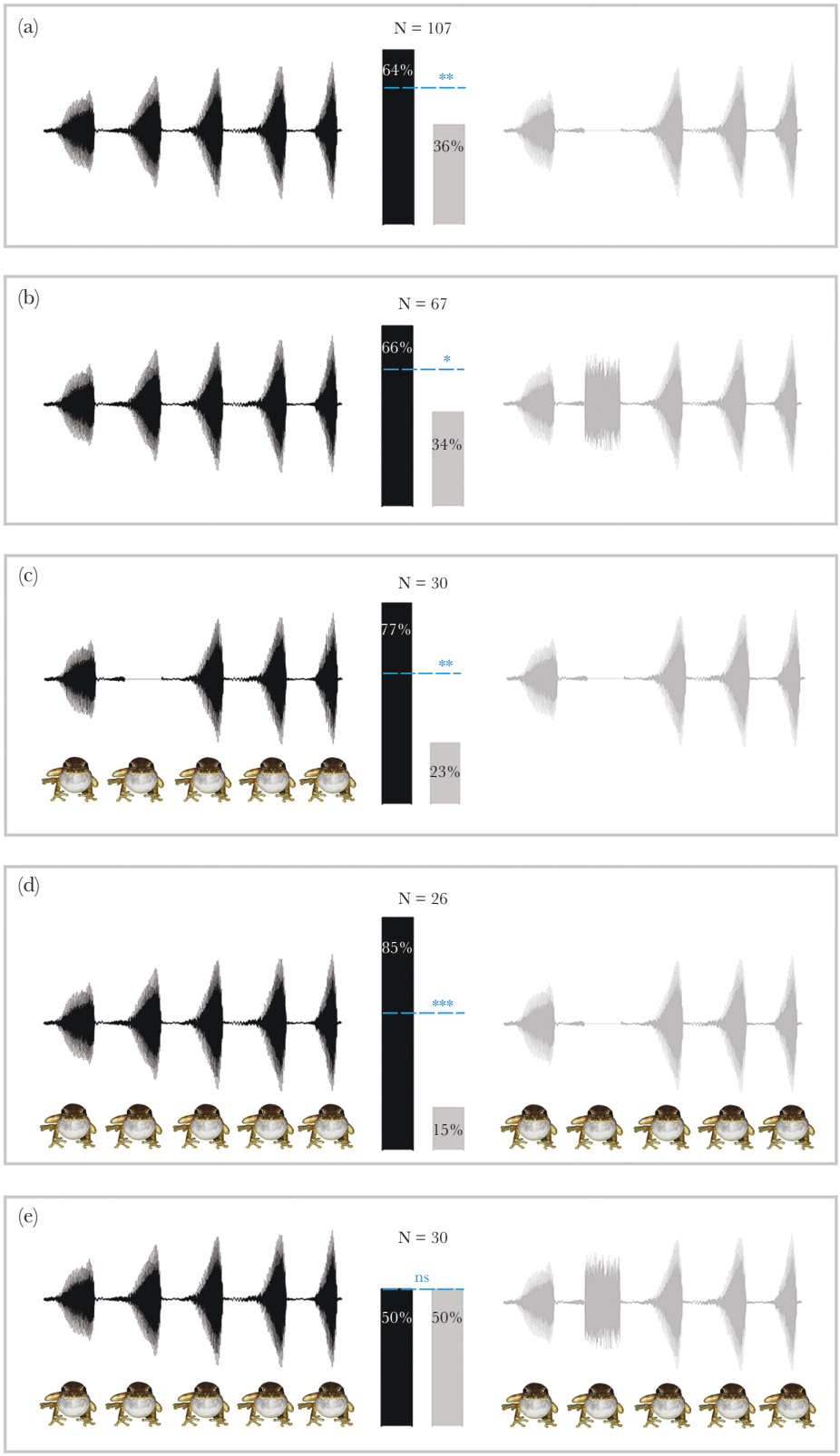

Figure 3

Female preference responses. Each portion of the figure illustrates the acoustic or/and visual components of the serrate-legged small treefrogs' mating signals: a complete five-note advertisement call, "complete call" (a-b, d-e, left black); advertisement call with the gap of silence, "silent gap" (a, c-d, right gray, c, left black); advertisement call with the gap of silence filled with noise, "noise gap" (b, e, right gray). The calling frogs under the waveform represent the inflation-deflation cycles of the vocal sac shown by video playback (c-e, left black, d-e, right gray); the inflation and deflation of the vocal sac are accompanied by male frogs producing each note. The x-axis represents 2 s. The vertical black and gray bars and the ratio in the bars represent the proportion of females choosing the respective signal, and the blue dashed horizontal lines represent the null hypothesis of equal preference (50%). The results of binomial tests (two tails) are noted as ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns (not significant) $P > 0.05$. The exact P values for each test are as follows: (a) $P = 0.0065$, $N = 107$; (b) $P = 0.0139$, $N = 67$; (c) $P = 0.0052$, $N = 30$; (d) $P = 0.0005$, $N = 26$; (e) $P = 1.0$, $N = 30$.

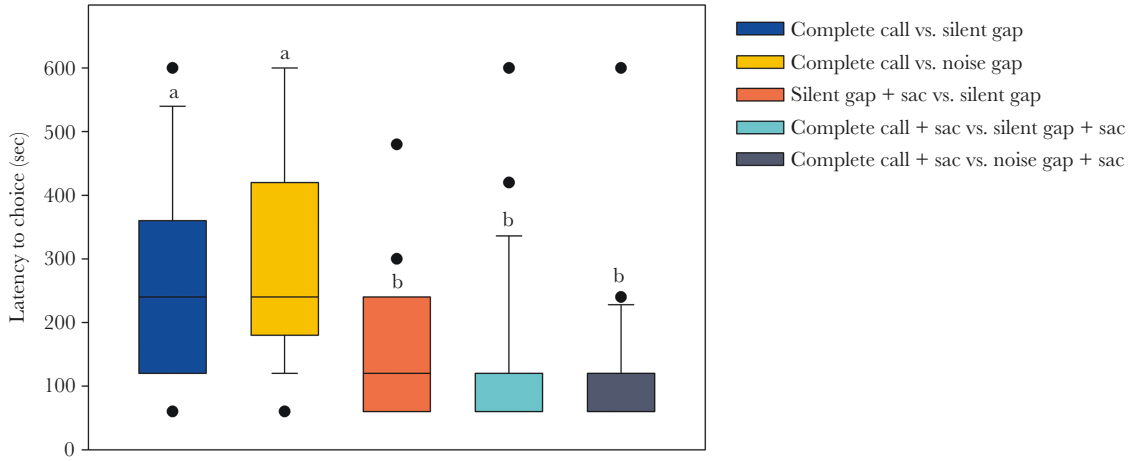

Figure 4

Female latency to choose in five tests. The latencies for each test are as follows: complete call vs. silent gap, 263.0 ± 161.4 s, n = 107; complete call vs. noise gap, 310.7 ± 162.5 s, n = 67; silent gap + visual sac vs. silent gap, 138 ± 101.1 s, n = 30; complete call + visual sac vs. silent gap + visual sac, 131.5 ± 129.3 s, n = 26; complete call + visual sac vs. noise gap + visual sac, 112 ± 138.0 s, n = 30. In the generalized linear mixed model (GLMM), pairwise comparisons with the estimated marginal means contrasts were used to compare female latency to choose between five tests. Data were expressed as Mean ± SD. Different superscript letters indicate significant differences among different tests.

Table 1

Parameter estimates of the generalized linear mixed model (GLMM)

					95% Confidence Interval	
Model Term	Coefficient	Std. Error	t	Sig.	Lower	Upper
Intercept	112.000	25.200	4.444	0.000	62.374	161.626
test = a	150.991	29.641	5.094	0.000	92.619	209.363
test = b	198.746	32.077	6.196	0.000	135.576	261.916
test = c	26.000	31.234	0.832	0.406	−35.510	87.510
test = d	19.538	35.745	0.547	0.585	−50.855	89.932
test = e	0[a]					

[a]This coefficient is set to zero because it is redundant.

Our results demonstrate that: 1) inserting white noise in the gap did not restore the attractiveness of interrupted calls; 2) presentation of a dynamic vocal sac resulted in more females preferring the interrupted calls, but they were still less preferred than the complete call with a dynamic vocal sac; and 3) combining the visual cue of a dynamic vocal sac along with noise in the gap, however, compensated the interrupted call and made it as attractive as the complete call. In addition, our results also demonstrate that: 4) acoustic calling rate, in this case, a non-interrupted call vs. an interrupted call, matters whether or not the display is unimodal or multimodal; 5) when calling rate is held constant multimodal displays are more attractive than unimodal displays; and 6) acoustic signal composition is less important in the presence of visual cues.

Revealing the function of multimodal signals is critical to understanding their evolution (Hebets and Papaj 2005). Here we show that the combined call was more attractive than the same interrupted call that was not accompanied by a dynamic vocal sac in serrate-legged small treefrogs. Our previous study revealed that female frogs preferred male advertisement calls over isolated visual sacs (Zhu et al. 2021). Thus, an isolated vocal sac can improve male attractiveness when combined with an acoustic call but is not sufficient to elicit a female response. Although we are not sure whether an advertisement call and the visual sac convey the same information about signaller quality, they certainly play different roles in female mate choice. The multimodal displays are more attractive than the unimodal acoustic display; thus, in general, our results support the inter-signal interaction hypothesis (Hebets and Papaj 2005; Caldart et al. 2021).

The presence of the dynamic vocal sac also reduces female latency to choose. Analogous results in response to multimodal stimulation are classically observed in neurophysics and psychophysics (Rowe 1999), but rarely in the context of mate choice (Hebets and Papaj 2005). Our results might be explained by the fact that visual cues offer more accurate location information than acoustic cues (Körding et al. 2007).

The vocal sac probably did not evolve as a visual signal but instead, as a mechanism to shuttle air back to the lungs during calling (Bucher et al. 1982; Pauly et al. 2006). There are however many species with adornments to the vocal sac, such as striking shapes and bright colors, and even behaviors in which the vocal sac is displayed in the absence of calling. All of this suggests that selection has favored the evolution of traits that enhance the conspicuousness of vocal sacs and thus promote the role and communication. In these cases, we can think of the vocal sac as a signal even though that was not its original function (Starnberger et al. 2014).

Table 2
Pairwise comparisons of female latency to choose between five tests

Test Pairwise Contrasts	Contrast Estimate	SE	*t*	*df*	*Adj. Sig.*	95% CI	
						Lower	Upper
a vs. b	−47.756	25.248	−1.891	255	0.060	−97.476	1.965
a vs. c	124.991	24.168	5.172	255	0.000	77.396	172.585
a vs. d	131.452	29.769	4.416	255	0.000	72.827	190.077
a vs. e	150.991	29.641	5.094	255	0.000	92.619	209.363
b vs. c	172.746	27.101	6.374	255	0.000	119.376	226.116
b vs. d	179.208	32.196	5.566	255	0.000	115.804	242.612
b vs. e	198.746	32.077	6.196	255	0.000	135.576	261.916
c vs. d	6.462	31.356	0.206	255	0.837	−55.289	68.212
c vs. e	26.000	31.234	0.832	255	0.406	−35.510	87.510
d vs. e	19.538	35.745	0.547	255	0.585	−50.855	89.932

The lowercase letters a–e represent the tests a–e, respectively. These pairwise comparisons are based on the estimated marginal means contrasts and adjusted using the least significant difference. $P < 0.05$ was considered to be statistically significant.

In general, multimodal signals may have evolved because they can be efficiently transmitted, detected, or remembered (Hogan and Stoddard 2018; Caldart et al. 2021; Zhu et al. 2021). Our results indicate that the visual sac not only provides enhancement of call attractiveness, but it also reduces female time to choose mates, and might potentially reduce search costs due to predation in some frogs (Ryan et al. 1982).

Noise interference often reduces the efficiency of communication (Wiley 2015; Gomes et al. 2016; Halfwerk et al. 2017; Jong et al. 2018). Similar to the cocktail party problem in humans, female frogs are challenged by noise interference as vocal communication often takes place in dense choruses (Bee and Micheyl 2008). Some frogs switch from the auditory channel to the visual channel (e.g., foot flagging) to cope with ambient noise (Hödl and Amézquita 2001). This tactic is referred to as multimodal shift (Partan 2017). In the present study, we show that the attractiveness of an interrupted call masked with noise was restored to that of a complete call when, and only when, exhibiting a dynamic vocal sac. This is similar to human beings who can read lip motions to enhance speech recognition when the speech is temporarily masked by loud noise (Xu et al. 2004). Our results demonstrate that dynamic visual cues (i.e., vocal sac) can compensate for noise interference in vocal communication. Such multisensory interactions can help frogs render an interrupted call masked with noise as attractive, and provide insurance against imperfect sender coding under a complex sound environment instead of a multimodal shift. Our results suggest that the process of multisensory integration in serrate-legged small treefrogs is not a simple additive operation of each multimodal component; the interaction between multimodal components may involve a unique behavioral response different from that involved by acoustic or visual component (Taylor and Ryan 2013; Stein et al. 2014; Halfwerk et al. 2019).

The intensity of a courtship display usually influences the mating success of signallers. Calling rate is an important indicator for female assessment of male quality (Schwartz et al. 2011; Henderson andGerhardt 2013; Tanner and Bee 2020). Our results demonstrate that acoustic calling rate is of primary importance regardless of whether the display is unimodal or multimodal. The presence of the dynamic vocal sac without noise compensates for some but not all of the salience of the interrupted call, unlike the complete multisensory rescue shown by túngara frogs (Taylor and Ryan 2013). An alternative explanation is that the complete call with the dynamic vocal sac might be substantially more attractive than the interrupted call with the same vocal sac. More interestingly, the proportion of female frogs choosing the interrupted call versus the complete call decreased from 36.4% to 15.4% in the presence of the visual cue, which indicates that the acoustic calling rate probably matters more in the presence of the vocal sac than in its absence. Our results suggest that the effect of multimodal integration (i.e., an auditory signal plus a dynamic visual cue) probably depends on the integrity of the auditory signal.

Broadcasting an interrupted call with only noise in the gap does not render the attractiveness of the call similar to that of a complete call, but this is not the case if the noise is accompanied by the visual cue of the vocal sac. These results demonstrate that acoustic signal composition is less important in the presence of visual cues, either because there is reduced attention to acoustic cues in the presence of visual cues or perception of the multimodal cue is more salient (Gomez et al. 2011). Regardless, it suggests that the selection pressure on acoustic signals may be lessened in the presence of visual cues, and this has important implications not only for current function but also for evolutionary patterns.

CONCLUSION

Our study revealed a novel function of multisensory integration in courtship communication. Combining a dynamic vocal sac along with noise compensated the interrupted call and made it as attractive as the complete call. This study helps us understand how animals integrate information from multisensory cues to rescue signal salience, especially in a noisy environment (e.g., the cocktail party problem). Such communication benefits to the receiver from multisensory integration may be an important selective force favoring multimodal signal evolution. Meanwhile, we found that acoustic signal composition is less important in the multimodal display, which suggests that the selection on acoustic signals may be lessened in the presence of visual cues.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at *Behavioral Ecology* online.

FUNDING

This work was supported by National Natural Science Foundation of China (31772464, 32001095), Sichuan Science and Technology Program (2022JDTD0026), Youth Innovation Promotion Association CAS (2012274),

and the Biodiversity Survey and Assessment Project of the Ministry of Ecology and Environment, China (2019HJ2096001006).

We thank Zhixin Sun for assistance with data collection, Haodi Zhang for assistance with data analysis, Xia Qiu for assistance with reflectance measurement, and the Mt. Diaoluo National Nature Reserve Administration for the support to perform this experiment.

Conflict of Interest Statement: The authors declare no competing financial interests.

Data Accessibility: Analyses reported in this article can be reproduced using the data provided by Zhu et al. (2022).

Handling Editor: Michael D. Jennions

REFERENCES

Bahrick LE, Lickliter R, Flom R. 2004. Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. Curr Dir Psychol Sci. 13:99–102.

Beaupre SJ., Jacobson ER, Lillywhite HB, Zamudio K. (2004). Guidelines for the use of live amphibians and reptiles in field and laboratory research. Available from: http://www.asih.org/sites/default/files/documents/resources/guide.

Bee MA, Micheyl C. 2008. The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol. 122:235–251.

Boersma PPG. 2002. Praat, a system for doing phonetics by computer. Glot Int. 10:341–345.

Bucher TL, Ryan MJ, Bartholomew GW. 1982. Oxygen consumption during resting, calling and nest building in the frog *Physalaemus pustulosus*. Physiol Zool. 55:10–22.

Caldart VM, dos Santos MB, Machado G. 2021. Function of a multimodal signal: a multiple hypothesis test using a robot frog. J Anim Ecol. 00:1–14.

Cummings ME, Bernal XE, Reynaga R, Rand AS, Ryan MJ. 2008. Visual sensitivity to a conspicuous male cue varies by reproductive state in *Physalaemus pustulosus* females. J Exp Biol. 211:1203–1210.

Deng K, Zhu B, Zhou Y, Chen Q, Wang T, Wang J, Cui J. 2019. Mate choice decisions of female serrate-legged small treefrogs are affected by ambient light under natural, but not enhanced artificial nocturnal light conditions. Behav Process. 169:103997.

Gomes DGE, Page RA, Geipel I, Taylor RC, Halfwerk W. 2016. Bats perceptually weight prey cues across sensory systems when hunting in noise. Science 353:1277–1280.

Gomez D, Théry M, Gauthier A, Lengagne T. 2011. Costly help of audiovisual bimodality for female mate choice in a nocturnal anuran (*Hyla arborea*). Behav Ecol. 22:889–898.

Halfwerk W, Smit J, Loning H, Lea AM, Geipel I, Ellers J, Ryan MJ. 2017. Environmental conditions limit attractiveness of a complex sexual signal in the túngara frog. Nat Commun. 8:1891.

Halfwerk W, Varkevisser J, Simon R, Mendoza E, Riebel KW. 2019. Toward testing for multimodal perception of mating signals. Front Ecol Evol. 7:124.

Hebets EA, Papaj DR. 2005. Complex signal function: developing a framework of testable hypotheses. Behav Ecol Sociobiol. 57:197–214.

Henderson JJ, Gerhardt HC. 2013. Restoration of call attractiveness by novel acoustic appendages in grey treefrogs. Anim Behav. 86:537–543.

Hödl W, Amézquita A. 2001. Visual signaling in anuran amphibians. In: Anuran Communication, M. J. Ryan, editor. Washington: Smithsonian Institution Press. p. 121–141.

Hogan BG, Stoddard MC. 2018. Synchronization of speed, sound and iridescent color in a hummingbird aerial courtship dive. Nat Commun. 9:5260.

Jong KD, Amorim MCP, Fonseca PJ, Heubel KU. 2018. Noise affects multimodal communication during courtship in a marine fish. Front Ecol Evol. 6:113.

Kaiser K, Boehlke C, Edauri N-P, Vega A, Robertson JM. 2018. Local preference encoded by complex signaling: mechanisms of mate preference in the red-eyed treefrog (*Agalychnis callidryas*). Behav Ecol Sociobiol. 72:182.

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. 2007. Causal inference in multisensory perception. PLoS One. 2:e943.

Leslie CE, Rosencrans RF, Walkowski W, Gordon WC, Bazan NG, Ryan MJ, Farris H. 2020. Reproductive state modulates retinal sensitivity to light in female túngara frogs. Front Behav Neurosci. 13:293.

McGurk H, Macdonald J. 1976. Hearing lips and seeing voices. Nature 264:746–748.

Narins PM, Hödl W, Grabul DS. 2003. Bimodal signal requisite for agonistic behavior in a dart-poison frog, *Epipedobates femoralis*. Proc Natl Acad Sci USA. 100:577–580.

Partan SR. 2017. Multimodal shifts in noise: switching channels to communicate through rapid environmental change. Anim Behav. 124:325–337.

Partan SR, Marler P. 2005. Issues in the classification of multimodal communication signals. Am Nat. 166:231–245.

Pauly GB, Bernal XE, Rand AS, Ryan MJ. 2006. The vocal sac increases call rate in the túngara frog, *Physalaemus pustulosus*. Comp Physiol Biochem. 79:708–719.

Reichert MS, Höbel G. 2015. Modality interactions alter the shape of acoustic mate preference functions in gray treefrogs. Evolution 69:2384–2398.

Rosenthal GG, Rand AS, Ryan MJ. 2004. The vocal sac as a visual cue in anuran communication: an experimental analysis using video playback. Anim Behav. 68:55–58.

Rowe C. 1999. Receiver psychology and the evolution of multicomponent signals. Anim Behav. 58:921–931.

Ryan MJ, Tuttle MD, Rand AS. 1982. Bat predation and sexual advertisement in a neotropical anuran. Am Nat. 119:136–139.

Schwartz JJ, Huth K, Jones SH, Brown R, Marks J, Yang X. 2011. Tests for call restoration in the gray treefrog H*yla versicolor*. Bioacoustics 20:59–86.

Stange N, Page RA, Ryan MJ, Taylor RC. 2017. Interactions between complex multisensory signal components result in unexpected mate choice responses. Anim Behav. 134:239–247.

Starnberger I, Preininger D, Hödl W. 2014. The anuran vocal sac: a tool for multimodal signalling. Anim Behav. 97:281–288.

Stein BE, Stanford TR, Rowland BA. 2014. Development of multisensory integration from the perspective of the individual neuron. Nat Rev Neurosci. 15, 520–535.

Tanner JC, Bee MA. 2020. Inconsistent sexual signaling degrades optimal mating decisions in animals. Sci Adv. 6, eaax3957.

Taylor RC, Klein BA, Stein J, Ryan MJ. 2008. Faux frogs: multimodal signalling and the value of robotics in animal behaviour. Anim Behav. 76:1089–1097.

Taylor RC, Ryan MJ. 2013. Interactions of multisensory components perceptually rescue túngara frog mating signals. Science 341:273–274.

Wiley, R. H. 2015. Noise matters: the evolution of communication. Cambridge: Harvard University Press.

Xu R, Ren Z, Dai W, Lao D, Kwan C. 2004. Multimodal speech enhancement in noisy environment. International Symposium on Intelligent Multimedia, Video and Speech Processing. IEEE 21:24.

Zhu B, Wang J, Brauth SE, Tang Y, Cui J. 2017a. The spectral structure of vocalizations match hearing sensitivity but imprecisely in *Kurixalus odontotarsus*. Bioacoustics 26:121–134.

Zhu B, Wang J, Zhao L, Chen Q, Sun Z, Yang Y, Brauth SE, Tang Y, Cui J. 2017b. Male-male competition and female choice are differentially affected by male call acoustics in the serrate-legged small treefrog, *Kurixalus odontotarsus*. Peer J. 5, e3980.

Zhu B, Yang Y, Zhou Y, Deng K, Wang T, Wang J, Tang Y, Ryan MJ, Cui J. 2022. Multisensory integration facilitates perceptual restoration of an interrupted call in a species of frog. Behav Ecol. doi:10.5061/dryad.f7m0cfxrs.

Zhu B, Zhou Y, Yang Y, Deng K, Wang T, Wang J, Tang Y, Ryan MJ, Cui J. 2021. Multisensory modalities increase working memory for mating signals in a treefrog. J Anim Ecol. 90:1455–1465.