

Cross-modal integration in a dart-poison frog

Peter M. Narins*[†], Daniela S. Grabul[‡], Kiran K. Soma*[§], Philippe Gaucher[¶], and Walter Hödl[‡]

*Department of Physiological Science, University of California, Los Angeles, CA 90095-1606; [†]Institute of Zoology, University of Vienna, Althanstrasse 14, A 1090 Vienna, Austria; and [‡]Mission pour la Création du Parc de la Guyane, 81 Rue Christophe Colomb, 97300 Cayenne, Guyane, France

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved December 22, 2004 (received for review August 31, 2004)

The mechanisms by which the brain binds together inputs from separate sensory modalities to effect a unified percept of events are poorly understood. This phenomenon was studied in males of the dart-poison frog *Epiplatys femoralis*. These animals physically and vigorously defend their territories against conspecific calling intruders. In prior field studies with an electromechanical model frog, we were able to experimentally evoke this aggressive behavior only when an auditory cue (advertisement call) was presented simultaneously with a visual cue (vocal-sac pulsations). In the present field experiments, we used a modified version of the electromechanical model frog to present territorial males with visual and auditory cues separated by experimentally introduced temporal delays or spatial disparities to probe temporal and spatial integration in this animal. In temporal integration experiments, bimodal stimuli with temporal overlap during calling bouts consistently evoked aggressive behavior; stimuli lacking bimodal temporal overlap were relatively ineffective at the same task. In spatial integration studies, despite presenting the components of the bimodal stimulus with an initial spatial disparity of up to 12 cm, fighting behavior persisted. These results demonstrate that temporal and spatial integration may be reliably estimated in a freely behaving animal in its natural habitat and that we can use aggressive behavior in this species as an index of cross-modal integration in the field.

animal communication | territorial defense | anuran | amphibian | Dendrobatidae

In many mammalian species, auditory signals are often produced with a synchronously generated and conspicuous set of visual signals (1). Both cues are known to be important in the perception of human speech, and experiments in which the timing or spatial location of these cues are intentionally separated can have profound effects on speech comprehension (2, 3). Nevertheless, ventriloquism, the art of making one's voice appear to emanate from a source different from the actual source, deliberately exploits our ability to "bind" spatially separate auditory and visual cues over a limited range in a process called cross-modal spatial integration (SI) (4,5,||). Increasing the spatial disparity between a voice's actual source and its apparent source by $>30^\circ$ or delaying the auditory cue relative to the visual cue by >200 ms reduces the realism of the illusion (6)

Many anuran amphibians also produce simultaneous bimodal advertisement signals in which the call is accompanied by a conspicuous visual cue: vocal-sac inflation.** For example, it has been shown that fighting behavior in the diurnal dart-poison frog *Epiplatys femoralis* can be experimentally evoked with an electromechanical model frog (EMF) broadcasting bimodal cues consisting of the synthetic advertisement call coupled with simultaneous vocal-sac pulsations (7). As long as the species-specific advertisement call is broadcast at a level >68 dB sound pressure level measured at the position of the focal male (12) and the vocal-sac pulsations are synchronous with the acoustic stimulus, fighting behavior ensues (7). Thus, a protocol has been established in which simultaneous and colocalized bimodal cues evoke aggressive behavior in territorial male *E. femoralis* (7). Taking advantage of this robust aggressive response to simulated intruders into their territory, we tested two new related hypoth-

eses concerning aggressive behavior in this species. The first hypothesis states that the introduction of sufficient discrete temporal disparities between the bimodal cues substantially reduces or extinguishes integration and, hence, evoked aggression. This possibility was tested by using the EMF's internal loudspeaker to present the natural call either in synchrony with or preceding the EMF's vocal-sac inflation and deflation with varying degrees of asynchrony (temporal delays). The second hypothesis states that the introduction of sufficient spatial disparity between the bimodal cues substantially reduces or extinguishes integration and, hence, evoked aggression. This possibility was tested directly by broadcasting the natural call from a second loudspeaker physically identical to the EMF's speaker but housed in a baffle and placed at various distances (spatial disparities) from the EMF. In these experiments, the natural call was broadcast in perfect synchrony with the EMF's vocal-sac inflations and deflations. Aggressive behavior was quantified to provide a biological measure of stimulus effectiveness. A corollary that follows from these two hypotheses is that there is a range of non-zero temporal and spatial disparities between the bimodal cues for which aggressive behavior persists. This corollary was also directly tested by providing small temporal delays or small spatial separations between the acoustic and visual cues and quantifying the resulting investigatory (approach) and aggressive (fighting) behavior. Experimental evidence in favor of either hypothesis would support the existence of the "ventriloquism illusion" in frogs (13, 14).

Methods

Study Animals and Field Site. We studied one of the >192 presently known species of dart-poison frogs, *E. femoralis* (15). Dart-poison frogs are endemic to Central America and northern South America, where they inhabit the wet tropical forests of the Amazonian lowlands and the Guiana Shelf (16). Males of *E. femoralis* are small (mean snout-to-urostyle length in our study population, 25.5 mm), diurnal, and produce a repetitive, four-note advertisement call lasting ≈ 0.4 s in which each note is swept upward in frequency in the range from ≈ 2.6 to 4.1 kHz (7, 17). The vocalizations are organized into bouts of up to 40 calls separated by ≈ 0.5 s, followed by a silent period of ≈ 8 s, on average. Males are territorial, and they physically and vigorously defend these territories against conspecific intruders (7, 12, 18). All experiments were done in January 2003 during the rainy season with free-ranging males of *E. femoralis* calling in primary

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: EMF, electromechanical model frog; IS, in sync; OL, overlap; IL, interleave; GA, gross alternate; SI, spatial integration; EL, external loudspeaker.

See Commentary on page 2267.

[†]To whom correspondence should be addressed. E-mail: pnarins@ucla.edu.

[§]Present address: Department of Psychology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4.

^{||}Grant, K. W. & Van Wassenhove, V. (2004) *J. Acoust. Soc. Am.* **115**, 2402 (abstr.).

**The importance of the visual component of the anuran advertisement call generally has been downplayed because most anuran amphibians are nocturnally active. Nevertheless, some evidence exists that visual signals may play an important role in diurnal species (7–9) and in nocturnal species under limited light regimes (10, 11).

© 2005 by The National Academy of Sciences of the USA

forest near the field station in Arataï, French Guiana (3°59'N, 52°35'W). The study site is in a lowland wet tropical forest (elevation 23 m) in which mean annual temperature and rainfall are 26°C and 3,000–3,250 mm, respectively.

Stimulus Presentation and Delivery. After a calling male was located and before each playback experiment, two nearly identical artificial logs equipped with frog models were placed on the forest floor 2 m from each other and from the male to be tested. One stimulus was chosen randomly from a group of four (see below) and presented in its entirety by using the first, “active” model, whereas the second log remained temporarily “inactive.” At the conclusion of each trial, the test male was usually on or in the vicinity of the active log. Before the subsequent trial, the active and inactive models were electrically switched (*i*) to ensure that, at the start of each new trial, the male under test was ≈2 m from the new active log and (*ii*) to minimize disturbance of the frog caused by the researchers. Each stimulus was presented a maximum of once to each male, with the exception of the advertisement call mimic stimulus (see below), which was occasionally repeated as a control to verify the aggressive state of the male being tested, or if rain or frog movement prevented completion of the trial. Not all stimuli were presented to all males tested.

Bimodal (i.e., visual and acoustic) stimuli were delivered by an EMF made from silicone rubber and painted to mimic an adult male of *E. femoralis*. The specifications of the model have been reported in ref. 7, so only a summary of its properties and updated features will be presented here. The EMF sits on an artificial log outfitted with hardware to inflate and deflate the male’s vocal sac and to broadcast the advertisement call. The principal new features added to the previous version of the EMF are (*i*) the ability to systematically regulate under computer control (MAC PowerBook G3, Apple) the timing between call broadcast and vocal-sac inflation, i.e., the degree of cross-modality synchrony, and (*ii*) the option to divert the call stimulus to the external loudspeaker (EL).

Behavioral Recording. All trials were filmed with a tripod-mounted digital video camcorder (Canon GL1). An approach was deemed successful if the test male reached an invisible 30-cm perimeter around the log during the trial period. For the analysis of the animals’ spatial location and its attraction to stimuli during the trials, only successful approaches were included; for the analysis of calling behavior, all approaches were included. For each trial, the time spent by the subject outside and inside the 30-cm perimeter was noted, as was the time spent on the log during which the subject vocalized, contacted, or attacked the model. The time during a trial spent outside the 30-cm perimeter is equivalent, in most cases, to the time for a “successful approach,” because males make direct, one-way approaches to the sound source (12). A contact was defined as touching the model in passing, whereas an attack usually involved prolonged bouts of wrestling, pushing, jumping on, and/or thrusting a forelimb at the model.

Temporal Integration Experiments. For these experiments, the call was presented from a loudspeaker (FNX140X tweeter, Rockford Fosgate, Tempe, AZ) located inside an artificial branch emerging from the upper surface of the log, with the loudspeaker cone located 2 cm behind the frog model. To the male being tested, we presented bimodal stimuli from the EMF consisting of vocal-sac pulsations that were delivered synchronously or after a given delay with respect to the advertisement call. To conduct this test, we recorded the “average” call of *E. femoralis* (7) on one channel of SOUNDMAKER 1.0.4b2 (MicroMat Computer Systems, Santa Rosa, CA) and a series of vocal-sac control pulses on a second channel (Fig. 1). The call stimulus parameters represent

the mean values for 15 males from the Arataï population and were as follows. The number of notes per call was four. The note duration and frequency-sweep range of notes 1, 2, 3, and 4 were 32.4 ms, 3,011–3,450 Hz; 66.1 ms, 2,985–3,846 Hz; 50.8 ms, 3,004–3,767 Hz; and 64.0 ms, 3,026–3,932 Hz, respectively. Internote intervals between notes 1 and 2, 2 and 3, and 3 and 4 were 50.2, 96.2, and 43.9 ms, respectively. The intercall interval was 458 ms, and the interbout interval was 8.2 s. The number of calls per bout was chosen arbitrarily to be 10. Prerecorded calls were broadcast at sound pressure levels between 80.8 and 84.1 dB (impulse time constant) measured in after-trial controls at a distance of 2 m. These values fall within the range of the natural calling levels of this species (12).

The bimodal stimuli used in these experiments (Fig. 1) were as follows.

1. In Sync (IS). The onset and offset of each call note was accompanied by a synchronous vocal-sac inflation and deflation, respectively; the sac was deflated between calls but remained partially inflated between bouts. This stimulus was designed to mimic most closely the natural advertisement call behavior of the male.
2. Overlap (OL). The last two of the four call notes were accompanied by vocal-sac pulsations. The time delay between the onset of the first call note and the onset of the first vocal-sac control signal was 248 ms.
3. Interleave (IL). The four call notes were immediately followed by four vocal-sac pulsations in the 458-ms intercall interval after the fourth call note. The time delay between the onset of the first call note and the onset of the first vocal-sac control signal was 434 ms.
4. Gross Alternate (GA). Ten calls (one bout) were produced without vocal-sac movements, and, after a delay of 458 ms, these were followed by 10 vocal-sac pulsations with no calls. The time delay between the onset of the first call note and the onset of the first vocal-sac control signal was 8.7 s.

Each of the stimuli above was repeated 25 times during one trial, except GA, which was repeated 16 times such that all experimental trials were of equal duration (402 s).

SI Experiments. For the SI experimental trials, simultaneous acoustic and visual cues were presented to the test male, but the call was broadcast from an EL (FNX2401x tweeter, Rockford Fosgate) housed in a Sony SRS-A27 baffle physically displaced from the visual stimulus (the EMF displaying vocal-sac pulsations). The external speaker was oriented toward the test frog’s initial position and placed on the forest floor in one of three locations: (*i*) 12 cm from the center of the model frog (immediately adjacent to the log), (*ii*) 25 cm from the model frog, or (*iii*) ≈50 cm from the model frog. These distances are equivalent to angles of 3.6°, 7.2°, and 14.4°, respectively, subtended between the EMF and the EL as measured from the test male’s initial position, 2 m from the EMF. In addition, in some trials, the internal speaker that was used provided a spatial disparity of 2 cm between the auditory and visual cues. The acoustic stimulus used for all SI experiments was IS, and each trial lasted 402 s. Experimental results were compared and analyzed by using either ANOVA followed by Fisher’s protected least-squares difference post hoc test, χ^2 analysis, or one-tailed Fisher’s exact test. $P < 0.05$ was regarded statistically significant.

Results

Temporal Integration Experiments. Sixteen different males were tested in their natural habitat. Of these males, two never made a successful approach to the EMF, whereas among the remaining 14 males, there was a total of 38 successful approaches analyzed

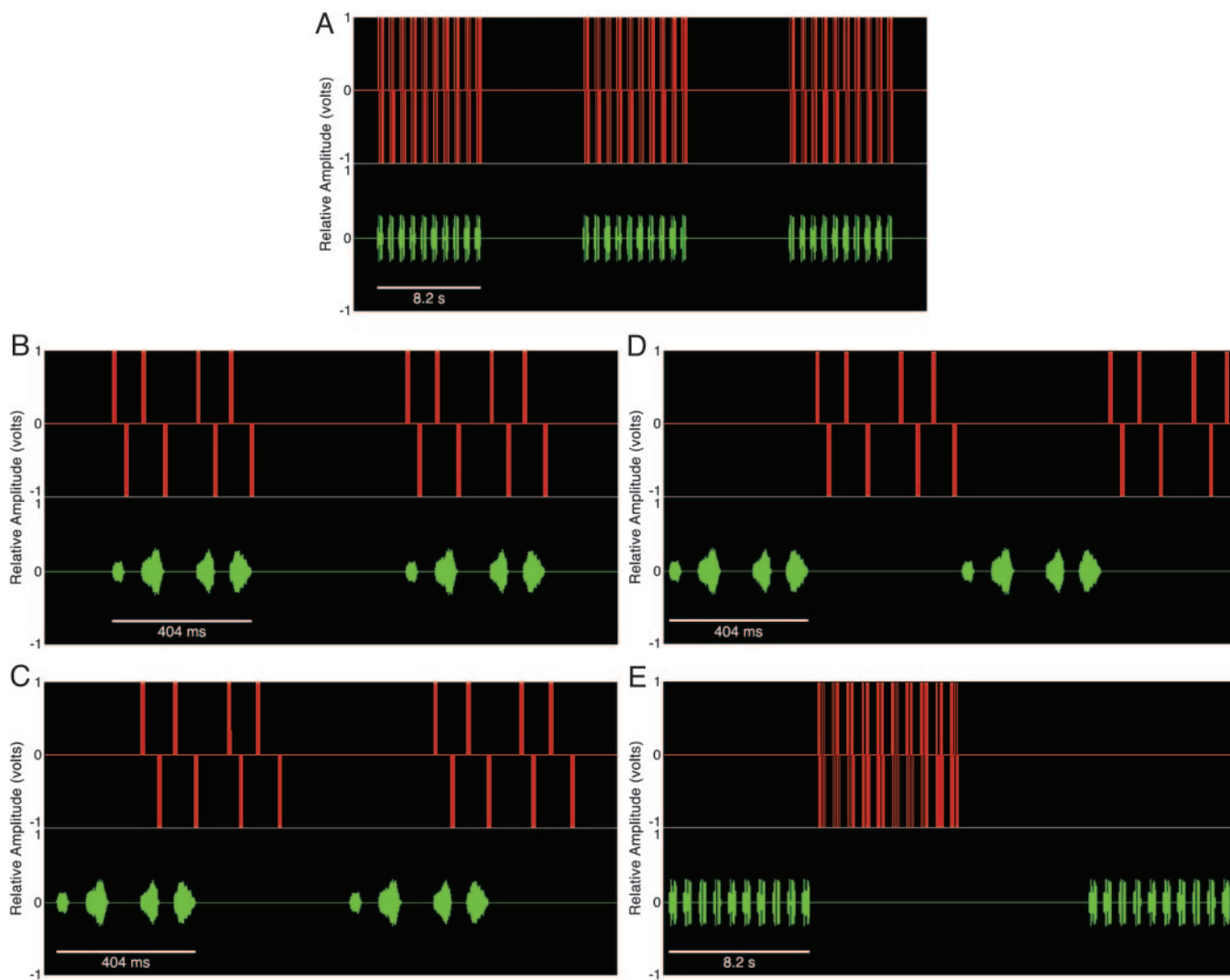


Fig. 1. Graphic representation of the temporal integration stimuli used in the experiment. The red traces show the vocal-sac control pulses for the EMF; the green traces show the natural call notes presented through the loudspeaker. (A) A portion of the IS stimulus used during a single trial showing three bouts of 10 natural calls, each accompanied by a perfectly synchronized vocal-sac pulsation. (B–E) Representative portions of the IS (B), OL (C), IL (D), and GA (E) stimuli illustrating the temporal relations between the call structure and the vocal-sac control pulses. All stimuli shown evoked contact and attack behavior from the males of *E. femoralis* tested except the GA stimulus.

from 47 completed trials (Table 1). Six partial trials (those terminated prematurely because of rain, darkness, etc.) were not included in the analysis. As expected, of the four stimuli presented to the frogs, the IS stimulus (natural call) evoked the highest ratio of attacks per successful approach (Table 1). The OL and IL stimuli evoked a marginally reduced attack rate, but the GA stimulus evoked significantly fewer attacks on the model

(none) than any of the other stimuli (one-tailed Fisher's exact test, $P < 0.005$, $n = 21$) (Table 1). The time required for a successful approach was stimulus-independent (Fig. 2A), consistent with the idea that the species-specific call alone is the long-distance attractant and the visual component of the call plays little or no role in this function (7). Mean approach velocity for all males making successful approaches to the model ($n = 41$ trials for 14 males) was $2.9 \text{ cm}\cdot\text{s}^{-1}$; three males reached a maximum approach velocity of $10.4 \text{ cm}\cdot\text{s}^{-1}$. Having made a successful approach, males spent significantly less time on the log during trials in which the GA stimulus was broadcast than during playback of either the IS stimulus ($P < 0.01$), the OL stimulus ($P < 0.05$), or the IL stimulus ($P < 0.005$) [ANOVA ($P < 0.01$), followed by Fisher's protected least-squares difference test ($P < 0.05$) in all cases] (Fig. 2B, green bars). Yet the percentage of males reaching the log was stimulus-type-independent (Fig. 2b, yellow bars). Thus, at close range, stimuli containing some degree of cross-modal temporal overlap (IS, OL, and IL) capture the attention of the male to a significantly greater extent than that containing none (GA). Mean calling times during a trial did

Table 1. Summary of *E. femoralis* aggressive behavior evoked during temporal integration experiments

Stimulus	Males tested, n	N_s	N_a
IS	15	12	8
OL	9	7	4
IL	12	9	5
GA	11	9	0
Totals	47	37	17

N_s and N_a are the number of successful approaches and attacks obtained for each stimulus type, respectively.

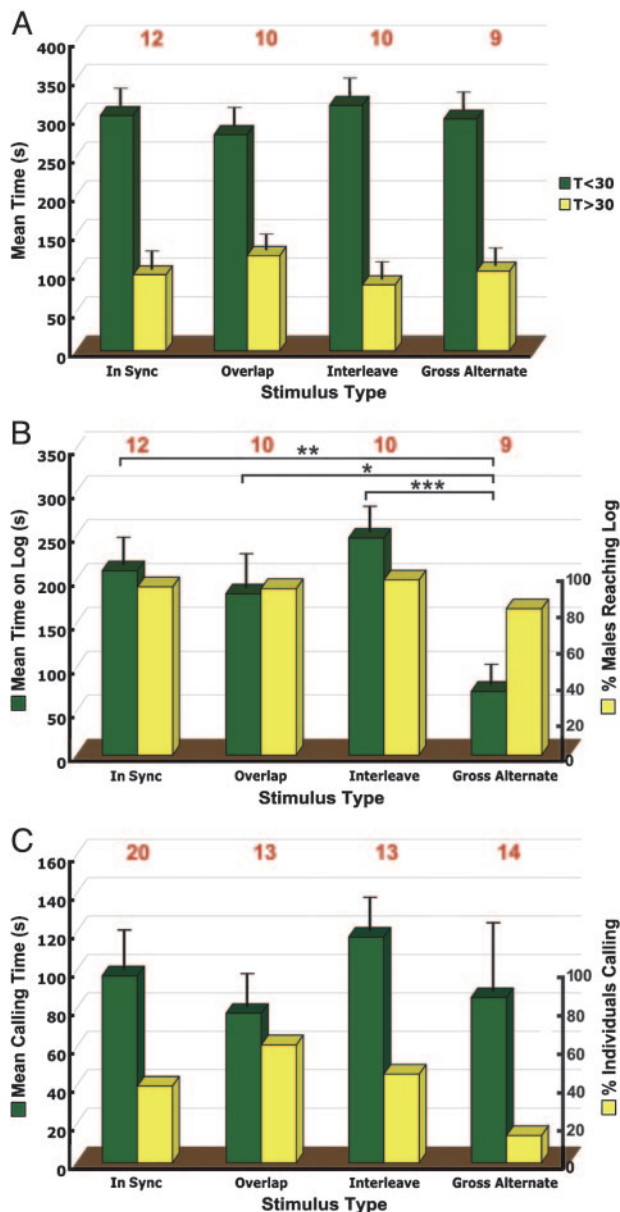


Fig. 2. Behavior of males of *E. femoralis* during temporal integration experiments. (A) Spatial location during trial. Mean times test males spent inside ($T_{<30}$) and outside ($T_{>30}$) a 30-cm perimeter around the EMF ($T_{<30} + T_{>30} =$ trial duration = 402 s). None of the $T_{<30}$ or $T_{>30}$ values is significantly different from the others (ANOVA, $P > 0.05$). (B) Attraction to stimuli. Mean times test males spent on the log. Males spent significantly less time on the log during GA stimulus trials compared with trials using other stimuli. ANOVA, followed by Fisher's protected least-squares difference test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.005$. The percentage of males reaching the log was stimulus-type-independent. (C) Calling behavior. Mean calling times during a trial did not differ among stimuli (ANOVA, $P > 0.05$), but a smaller percentage of males called during GA stimulus trials compared with trials using other stimuli. Numbers above histograms represent the number of frogs tested for each stimulus type. Error bars show SEM.

not differ among stimuli (ANOVA, $P > 0.05$) (Fig. 2C, green bars). Likewise the percentage of males calling during the experiment was stimulus-type-independent [ANOVA ($P > 0.05$) and one-tailed Fisher's exact test in all cases) (Fig. 2C, yellow bars).

SI Experiments. Twenty-five males were tested in their natural habitat. Of these, six never made a successful approach to the EMF,

Table 2. Summary of SI experiments for *E. femoralis*

D_s , cm	Males tested, n	Visual			Auditory	
		N_{SEM}	N_{aEMF}	Mean T_{EMF} , s	N_{SEL}	Mean T_{EL} , s
2	15	12	8	210.2	—	—
12	16	12	7	148.0	9	39.2
25	11	10	1	141.2	12	77.2
50	13	10	2	95.9	9	112.1
Totals	55	44	18	N/A	30	N/A

D_s is the spatial disparity between the EMF presenting the visual cue (vocal sac pulsations) and the EL presenting the auditory cue (advertisement call). N_{SEM} and N_{aEMF} are the numbers of successful approaches to and attacks on the model, respectively, obtained for all males tested. Mean T_{EMF} is the mean time spent on the EMF during a trial. N_{SEL} is the number of trials in which a male approached the EL, and Mean T_{EL} refers to the mean time spent by the test males on or within 25 cm of the EL. The data reported here correspond to one trial per male at each D_s . Males never attacked the EL. The advertisement call stimulus, IS, was used for all SI experiments. —, internal speaker used; N/A, not applicable.

whereas among the remaining 19 males, there were a total of 44 successful approaches analyzed from 55 completed trials (Table 2). In addition, of the 40 trials in which the EL was used, there were 30 trials in which the male made a successful approach to the EL. During 28 of these trials, males also made successful approaches to the EMF. Thus, males investigated the EL and EMF when presented with both. However, trials in which the loudspeaker was displaced 25 or 50 cm from the EMF (large displacement trials) (Fig. 3A) resulted in significantly fewer attacks on the model (data pooled; one-tailed Fisher's exact test, $P < 0.005$) (Table 2) than trials with loudspeaker displacements of 12 or 2 cm (small displacement trials) (Fig. 3B). Analyzed as absolute time during a trial spent contacting or attacking the model, large displacement trials resulted in 23.1% and 23.8% of the contact and attack time, respectively, on the model, whereas small displacement trials accounted for 76.9% and 76.2% of the contact and attack time, respectively, on the model. In addition, males spent progressively more time on or within 25 cm of the EL and progressively less time on the log as the spatial separation between these two devices increased (Table 2).

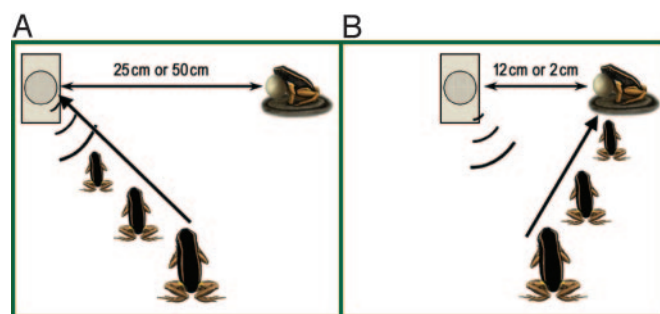


Fig. 3. Behavior of test males during SI trials. The field setup depicted male phototaxis behavior under two conditions: investigatory, with the EL placed at 50 or 25 cm from the model (data pooled) (A), and aggressive, with the EL placed at 12 or 2 cm from the model (data pooled) (B). The model and the loudspeaker were placed 2 m from the test male's calling position; his initial trajectory is indicated by the thick arrows. The time spent investigating the model did not differ significantly between playback configurations in A and B; nevertheless, conditions in A resulted in 23.1% and 23.8% of the contact and attack time for the EL placed at 50 and 25 cm, respectively, whereas conditions in B resulted in 76.9% and 76.2% of the contact and attack time for the EL placed at 12 and 2 cm, respectively. Thus, cross-modal SI in this species over a distance of ≤ 12 cm results in significantly higher incidence of aggressive behavior than over distances of 25 or 50 cm.

Discussion

Experiencing the environment requires constant integration of information from our different senses. A number of human studies have shown that visual cues can modulate the apparent location of auditory cues. This phenomenon is clearly seen in the “spatial ventriloquist effect,” where discrepancies in the spatial localization of synchronized auditory and visual events can lead to a bias of the perceived auditory localization toward the visual localization (13, 14). More recently, it has been demonstrated that a sound presented in close temporal proximity to a visual stimulus may be perceived as occurring simultaneously with the visual stimulus. Known as the “temporal ventriloquism effect,” this phenomenon can correct for asynchronous auditory and visual inputs by binding visual stimuli into temporal alignment with the appropriate auditory events (19–22). This phenomenon has not been investigated systematically in freely behaving animals in the field.

When delivered with an electromechanical model designed to mimic the frog’s morphology and behavior, simultaneous auditory (advertisement call) and visual (vocal-sac pulsations) cues are known to elicit aggressive behavior in the dart-poison frog *E. femoralis* (7). Here we show that, in *E. femoralis*, fighting behavior persists in response to a bimodal stimulus in which temporal asynchrony up to 434 ms (IL stimulus) has been introduced between its auditory and visual components. This result clearly supports our corollary that there is a range of non-zero temporal and spatial disparities between the bimodal cues for which aggressive behavior persists. Moreover, the long-distance attractiveness of these temporally asynchronous stimuli is not compromised (measured by comparing the times required to approach within 30 cm of the EMF when broadcasting these stimuli and the call mimic stimulus) (Fig. 2A). Nevertheless, bimodal temporal synchronous stimuli with some overlap (IS, OL, and IL) result in significantly longer attendance at the model and evoke calling in a higher percentage of individuals than the GA stimulus with no bimodal temporal overlap (Fig. 2B and C). Although exact temporal synchrony is thus not a prerequisite for evoking aggressive behavior, removal

of all bimodal temporal overlap certainly extinguishes it (Table 1). Thus, cross-modal temporal integration appears to be a prerequisite for evoking aggressive behavior in this species, lending support to our first hypothesis that the introduction of sufficient discrete temporal disparities between the bimodal cues substantially reduces or extinguishes integration and, hence, evoked aggression.

Results from our SI experiments with aggressive behavior as a metric suggest that spatial ventriloquism (cross-modal SI) also operates strongly in this amphibian for bimodal spatial disparities of 12 cm (equivalent to an initial angular separation of bimodal cues of 3.6°) and weakly for bimodal spatial disparities of 25 and 50 cm (equivalent to initial angular separations of bimodal cues of 7.2° and 14.4°, respectively). Although the large displacement trials (25 and 50 cm) resulted in significantly fewer contacts or attacks on the model, a bimodal spatial disparity of 50 cm was insufficient to extinguish aggressive behavior (Fig. 3). This result supports our second hypothesis that the introduction of sufficient spatial disparity between the bimodal cues substantially reduces or extinguishes integration and, hence, evoked aggression. Consistent with this hypothesis, as the spatial disparity between the bimodal cues was increased, the animal’s attention was increasingly “captured” by the auditory stimulus (Table 2 and Fig. 3). This system affords the possibility of potential future studies of cross-modal integration both behaviorally and neurophysiologically, topics that have received increased attention of late (5, 23–28).

We thank Muriel Nugent and the staff of the Association Arataï for consistent and generous assistance with this project; Carlos Martinez for help with the modifications and improvements to the control box; Mario Penna for help constructing the stimuli; Christian Proy for immeasurable assistance with the field studies; Yana Durmashkin and Sylvia Kurtovic for aid with the data analysis; Adolfo Amézquita, Nancy Kanwisher, and Greg Grether for providing helpful comments on an early version of the manuscript; and Margaret Kowalczyk for able assistance with the figure preparation. This work was supported by National Institutes of Health Grant DC00222 and Academic Senate Grant 3501 (to P.M.N.) and Austrian Science Foundation FWF P 15345 (to W.H.). K.K.S. is supported by the Canadian Institutes of Health Research and the Michael Smith Foundation for Health Research.

- Hauser, M. D. (1996) *The Evolution of Communication* (MIT Press, Cambridge, MA).
- Pandy, P. C., Kunov, H. & Abel, S. M. (1986) *J. Aud. Res.* **26**, 27–41.
- Grant, K. W. & Greenberg, S. (2001) in *Proc. Aud.-Vis. Speech Process.* (Univ. California, Santa Cruz), pp. 132–137.
- Alais, D. & Burr, D. (2004) *Curr. Biol.* **14**, 257–262.
- Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K. & Hallett, M. (2003) *Nat. Neurosci.* **6**, 190–195.
- Jack, C. E. & Thurlow, W. R. (1973) *Percept. Motor Skills* **37**, 967–979.
- Narins, P. M., Hödl, W. & Grabul, D. S. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 577–580.
- Narins, P. M., Lewis, E. R. & McClelland, B. E. (2000) *J. Zool.* **250**, 283–298.
- Rosenthal, G. G., Rand, A. S. & Ryan, M. J. (2004) *Anim. Behav.* **68**, 55–58.
- Hödl, W. & Amézquita, A. (2001) in *Anuran Communication*, ed. Ryan, M. J. (Smithsonian Inst. Press, Washington, DC), pp. 121–141.
- Amézquita, A. & Hödl, W. (2004) *Herpetologica* **60**, 20–29.
- Hödl, W. (1987) in *Proc. 4th Ordinary Gen. Meet. Soc. Eur. Herpetologica*, eds. Gelder, van, J. J., Strijbosch, H., Bergers, & Bergers, P. J. M. (Catholic Univ. of Nijmegen, Nijmegen, The Netherlands), pp. 201–204.
- Bertelson, P. & de Gelder, B. (2004) in *Crossmodal Space and Crossmodal Attention*, eds. Spence, C. & Driver, J. (Oxford Univ. Press, Oxford), pp. 141–177.
- Stekelenburg, J. J., Vroomen, J. & de Gelder, B. (2004) *Neurosci. Lett.* **357**, 163–166.
- Hofrichter, R. (2000) *The Encyclopedia of Amphibians* (Key Porter, Toronto).
- Duellman, W. E. (1999) in *Patterns of Distribution of Amphibians: A Global Perspective*, ed. Duellman, W. E. (Johns Hopkins Univ. Press, Baltimore), pp. 255–328.
- Hödl, W., Amézquita, A. & Narins, P. M. (2004) *J. Comp. Physiol.* **190**, 823–829.
- Roithmair, M. (1992) *Ethology* **92**, 331–343.
- Spence, C. & Squire, S. (2003) *Curr. Biol.* **13**, R519–R521.
- Morein-Zamir, S., Soto-Faraco, S. & Kingstone, A. (2003) *Cognit. Brain Res.* **17**, 154–163.
- Vroomen, J. & de Gelder, B. (2004) *J. Exp. Psychol. Hum. Percept. Perform.* **30**, 513–518.
- Fujisaki, W., Shimojo, S., Kashino, M. & Nishida, S. (2004) *Nat. Neurosci.* **7**, 773–778.
- Recanzone, G. H. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 869–875.
- Fuster, J. M., Bodner, M. & Kroger, J. K. (2000) *Nature* **405**, 347–351.
- Bushara, K. O., Grafman, J. & Hallett, M. (2001) *J. Neurosci.* **21**, 300–304.
- Stein, B. E., Huneycutt, W. S. & Meredith, M. A. (1988) *Brain Res.* **448**, 355–358.
- Spence, C. & Driver, J. (1997) *Percept. Psychophys.* **59**, 1–22.
- Calvert, G. A., Spence, C. & Stein, B. E. (2004) *The Handbook of Multisensory Processes* (MIT Press, Cambridge, MA).