Integrative and Comparative Biology

Integrative and Comparative Biology, pp. 1–8
doi:10.1093/icb/icx027

Society for Integrative and Comparative Biology

SYMPOSIUM

Perceived Synchrony of Frog Multimodal Signal Components Is Influenced by Content and Order

Ryan C. Taylor,[1,*,†] Rachel A. Page,[†] Barrett A. Klein,[‡] Michael J. Ryan[§,†] and Kimberly L. Hunter[*]

[*]Department of Biological Sciences, Salisbury University, 1101 Camden Avenue, Salisbury, MD 21801, USA; [†]Smithsonian Tropical Research Institute, Balboa Ancon, 56292 Panama, Republic of Panama; [‡]Department of Biology, University of Wisconsin—La Crosse, La Crosse, WI 54601, USA; [§]Department of Integrative Biology, University of Texas at Austin, Austin, TX 12330, USA

From the symposium "Integrating Cognitive, Motivational and Sensory Biases Underlying Acoustic and Multimodal Mate Choice" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 4–8, 2017 at New Orleans, Louisiana.

[1]E-mail: rctaylor@salisbury.edu

Synopsis Multimodal signaling is common in communication systems. Depending on the species, individual signal components may be produced synchronously as a result of physiological constraint (fixed) or each component may be produced independently (fluid) in time. For animals that rely on fixed signals, a basic prediction is that asynchrony between the components should degrade the perception of signal salience, reducing receiver response. Male túngara frogs, *Physalaemus pustulosus*, produce a fixed multisensory courtship signal by vocalizing with two call components (whines and chucks) and inflating a vocal sac (visual component). Using a robotic frog, we tested female responses to variation in the temporal arrangement between acoustic and visual components. When the visual component lagged a complex call (whine + chuck), females largely rejected this asynchronous multisensory signal in favor of the complex call absent the visual cue. When the chuck component was removed from one call, but the robofrog inflation lagged the complex call, females responded strongly to the asynchronous multimodal signal. When the chuck component was removed from both calls, females reversed preference and responded positively to the asynchronous multisensory signal. When the visual component preceded the call, females responded as often to the multimodal signal as to the call alone. These data show that asynchrony of a normally fixed signal does reduce receiver responsiveness. The magnitude and overall response, however, depend on specific temporal interactions between the acoustic and visual components. The sensitivity of túngara frogs to lagging visual cues, but not leading ones, and the influence of acoustic signal content on the perception of visual asynchrony is similar to those reported in human psychophysics literature. Virtually all acoustically communicating animals must conduct auditory scene analyses and identify the source of signals. Our data suggest that some basic audiovisual neural integration processes may be at work in the vertebrate brain.

Introduction

Animal signals are complex, often consisting of individual components transmitted and received through multiple sensory channels (Hebets and Papaj 2005; Higham and Hebets 2013; Hebets et al. 2016). Signal complexity has been an area of intense research for more than 15 years (Partan and Marler 1999), yet we understand little about how a signal component in one sensory channel influences the perception and corresponding behavioral response to a component in another channel. In animal courtship signals, for example, do individual components in the auditory and visual channels combine to increase female responses in an additive fashion? Alternatively, does the addition of a visual component induce an exponentially stronger response in receivers or even reduce their response relative to the acoustic signal alone? Some recent work in animal communication indicates that the perception and subsequent behavioral response to multisensory

Published by Oxford University Press on behalf of the Society for Integrative and Comparative Biology 2017.

signals is not additive or easily predicted (Taylor and Ryan 2013; Rubi and Stephens 2016; Stange et al. 2016). To date, the most comprehensive work on audiovisual integration and non-additive effects has been done in cats and primates, including work in human psychophysics (for review see Stein 2012).

The human psychophysical work has been critical for informing us about how the senses are integrated and how this integration allows individuals to make sense of a complex world around them. In particular the recruitment of additional senses, such as vision, is one mechanism that humans use to locate and recognize acoustic signals, increasing the efficacy of our auditory scene analyses (Sumby and Pollack 1954). Psychophysical techniques have been applied to a number of taxa, but frogs are especially amenable to these methods, allowing us to address questions about the perception of complex signals (Bee and Micheyl 2008; Bee 2015). Male frogs produce stereotyped advertisement (mating) calls and their neural systems are "tuned" to properties of these calls (Ryan 2001). In most species, females search out calling males and approach them to initiate mating. If the mating signals deviate too far from their species-specific properties, female receivers fail to perceive them as appropriate mating signals (Phelps et al. 2005). Because females readily respond to acoustic playbacks of male signals, and engage in mate searching behavior, they are easy to manipulate in behavioral tests of signal perception. These perceptions are directly relevant to understanding how communication signals evolve. In frogs, the sex ratio is typically highly skewed and male reproductive success is likewise skewed. Therefore, female mate choice generates strong selection on male signal evolution.

The túngara frog, *Physalaemus pustulosus*, is a small frog found from northern South America through southern Mexico. Like many frog species, they breed in ephemeral pools of water and males produce a conspicuous acoustic signal, the advertisement call. In túngara frogs, this advertisement call consists of two components. The first is the "whine" and the second is the "chuck." The whine is necessary and sufficient for mate attraction and males always produce this component. The chuck is neither necessary nor sufficient for mate attraction but males can facultatively append up to seven chucks onto the end of the whine (usually one to three). Chucks make the whine more attractive to females, and always follow the whine as a result of morphological constraint (Ryan and Guerra 2014). The advertisement call is also accompanied by the synchronous inflation of a conspicuous vocal sac that has been shown to make the call more attractive (Taylor et al. 2008). Thus, females assess both the call and the vocal sac inflation as part of a multimodal signal.

The visual cue of an inflating vocal sac increases the attractiveness of a call when it is added, but its effect can easily be overridden by an alternative call that contains more attractive properties. Thus, the acoustic signal component has primacy for female mate choice. The temporal arrangement of the call and vocal sac movement are critically important, however. If the vocal sac inflation is delayed, such that it lags the call in time, females strongly reject this asynchronous multisensory signal (Taylor et al. 2011). Alternatively, temporally sandwiching the vocal sac movement between the whine and chuck can restore the saliency of the overall mating signal (Taylor and Ryan 2013). For an individual male, temporal delays between the call and vocal sac movement are impossible due to morphological constraints. Our previous experimental data show that females strongly attend to temporal synchrony of the signal components, yet are flexible about how they perceive and respond to temporal variation.

Our current understanding of the túngara frog system suggests that two simple rules may govern female choice for multisensory signals. First, if the vocal sac inflates following a call, then reject the signal. Second, if the whine and chuck "book end" the vocal sac, then accept the signal. Despite these data, we still have a largely incomplete understanding of how all three components—whine, chuck, and vocal sac—interact to influence perception and female mate choice.

In this study we further probed how females respond to asynchronous signals. Specifically, we asked two questions. First, we asked if acoustic content matters. Do females find an asynchronous multimodal signal aversive, when one or more of the calls lack a chuck? This question is important because it helps to shed additional light on the cognitive/perceptual system that governs how the frog audiovisual system processes complex signals. Second, we asked if there is a syntactical order effect. That is, does a vocal sac that leads a call in time influence female choice as it would if it lags the call? This question is intriguing because for males, order of call components is fixed; vocal sac inflations always coincide with the call and chucks always follow whines. Females, however, show permissiveness for temporal arrangement of chuck placement in tests with no visual cue (Wilczynski et al. 1999).

Methods

We collected mated pairs of túngara frogs at choruses within 4 h after sunset. The frogs were collected at breeding sites near the Smithsonian Tropical Research Institute, Gamboa, Republic of Panama. We placed individual frog pairs into plastic bags and stored the frogs in a light-safe cooler (total darkness) for a minimum of 1 h prior to testing. This ensured that the frogs' eyes were dark-adapted after collection using flashlights. After testing, the frogs were toe-clipped, following guidelines of the American Association of Ichthyologists and Herpetologists, which allowed us to avoid using recaptures on subsequent nights. We released all collected frogs at their sites of capture at the end of the night, ensuring that they could breed in the wild. All procedures were approved by STRI IACUC (2011-0825-2014-02) and conducted with permits from Panama's ANAM permit No. SE/A-30-12. ANAM is now the Ministry of the Environment, MiAmbiente.

We conducted phonotaxis experiments in a hemi-anechoic chamber (Acoustic Systems, ETS-Lindgren, Austin, TX, USA) measuring 2.7 m × 1.8 m × 2 m. For the behavioral tests, we used a restraining funnel placed in the center of the chamber. The funnel kept the females equidistant (80 cm) from the two speakers (Mirage Nanosat, Klipsch Audio, Indianapolis, IN, USA) used to broadcast the male calls (Fig. 1a). Each speaker was separated by 80 cm and formed a triangle with ca. 60° separation relative to the female's release point. To generate a multisensory signal, we placed a robotic frog (robofrog) with an inflatable vocal sac in front of one speaker. We inflated the vocal sac of the robotic frog remotely using a pneumatic pump that was triggered by the computer producing the acoustic stimulus. By using a sound file to trigger the robofrog vocal sac inflation, we were able to precisely control the timing of the robofrog's inflation/deflation sequence relative to the calls produced at the speaker. Because the speaker broadcast the call from the same location as the robofrog, this closely matched the spatial location of the natural visual and acoustic signal components (Taylor et al. 2008; Klein et al. 2012).

We illuminated the test chamber with a single GE nightlight (ca. 2.27×10^{-8} W/cm^{-2}, model no. 55507; Fairfield, CT, USA). The spectrum and intensity of light at nocturnal breeding sites varies tremendously with location (forest cover vs. open), moon phase, and cloud cover. The light environment we provided was well within the range of what frogs naturally experience (Cummings et al. 2008; Taylor et al. 2008). For each trial, we placed a female under the funnel and broadcast digitally synthesized male vocalizations (see Ryan et al. [2003] for details on call synthesis). The robofrog vocal sac was also activated to inflate/deflate asynchronously with the call broadcast at the speaker (for more details see Taylor et al. 2011). These playbacks were broadcast for 2 min, which allowed the female to acclimatize to the playbacks while under the funnel. For all experiments, we used a synthetic, simple (whine), or complex (whine + one chuck) call broadcast at 82 dB SPL (re. 20 μ Pa; RMS, fast, C weighting) measured at the point of release for the females. We used Adobe Audition software (ver. 3.0) for playbacks and each call was played once every 3 s.

After the acclimation period, we lifted the funnel so the female was free to move around the test arena. We recorded a choice when a female approached to within 5 cm of a speaker or speaker/robofrog combination and remained there for 5 s. The 5 s rule avoided false positives or negatives caused by females simply walking by a stimulus. To control for side bias, we systematically alternated the sides on which the robofrog and calls were presented between trials. If a female did not move for 2 min after the funnel was raised or failed to enter a choice zone within 10 min, we discarded the trial from the data set due to a lack of motivation. Response rates by females were typically around 65% each night. We recorded female behavior using an infrared sensitive camera (Everfocus EHD500IR, Everfocus Electronics, Duarte, CA, USA) mounted on the ceiling of the chamber. A video feed allowed us to view the female's behavior in real time from outside the sound chamber, while simultaneously recording video (Ethovision® recording program).

Following these general procedures, we conducted three experiments. In Experiment 1, we presented females with a complex call (whine plus one chuck, hereafter "WC") versus a simple call (the whine alone, hereafter "W"). The WC had the visual component of a robofrog with inflating vocal sac added, but the vocal sac inflation lagged the call. The call and robofrog inflation were 100% out of phase such that the inflation began immediately following the terminus of the call (Fig. 1b). The temporal sequence of this stimulus was: whine, then chuck, then vocal sac inflation, hereafter abbreviated as (WC-robo). In Experiment 2, we presented females with the identical W call at each speaker. To one speaker we also added a robofrog with inflation following the whine, hereafter abbreviated as (W-robo). Here also, the inflation occurred 100% out of phase, immediately following the call (Fig. 1b). In the third experiment, we presented females with two identical WC calls, but one speaker again had a robofrog added. The robofrog vocal sac inflation preceded the call yielding a temporal sequence of: vocal sac inflation, then

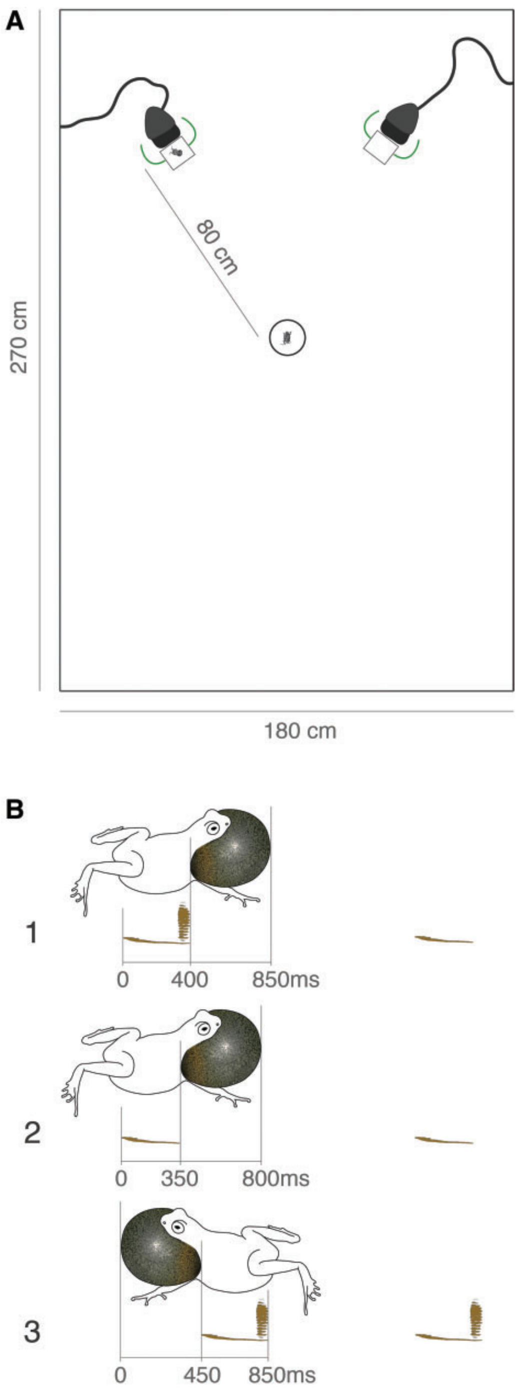

Fig. 1 (**A**) Diagram of two-choice test arena. Females could choose between two stimuli, a call only or a call with the asynchronously inflating robofrog placed in front of the speaker. (**B**) Detail of female choice tests. The asynchronous multimodal signals are depicted on the left side; the calls only are depicted by the sonograms on the right. In Experiment 1, the robofrog vocal sac inflation lagged the call (depicted in the timeframe above the whine–chuck sonogram). The alternative was a whine only. In Experiment 2, the robofrog vocal sac inflation lagged the call (depicted in the timeframe above the whine only sonogram). The alternative was also a whine only. In Experiment 3, the robofrog vocal sac inflation led the call (depicted in the timeframe above the whine–chuck sonogram). The alternative was also a whine chuck.

whine, then chuck, hereafter abbreviated as (robo-WC). Although the inflation preceded the call, the inflation still occurred 100% out of phase; the call began immediately following the deflation of the robofrog vocal sac (Fig. 1b).

Statistical analysis

All experiments consisted of a two-choice test, where females had the option of responding to a unimodal call (speaker only) or a multimodal signal (speaker plus the visual cue of a robofrog). The data were analyzed using a binomial exact test and the mid-*P* value (Agresti 2001). We previously showed that when the robofrog's vocal sac inflation temporally lagged the complex WC call by either 50% or 100%, females chose the multisensory signal only 25% of the time (Taylor et al. 2011). The timing of the lagging vocal sac in the current study matched the timing of the 100% from previous experiments. These experiments were later repeated (unpublished data), confirming the results. Given the repeatable and robust nature of the female preference function for a lagging visual component, we set our a priori expected binomial response to this asynchronous multisensory signal at 0.25 (Experiments 1 and 2). In Experiment 3, where the vocal sac inflation led the call, we had no prior data to suggest how females would respond to this particular temporal arrangement. Therefore, we set our a priori expected response rate at random choice = 0.5.

Results

In our first experiment, we presented females with a WC versus W call, but the robofrog was added to the speaker playing the WC and the vocal sac was inflated asynchronously following the call (WC-robo). Females chose the WC-robo in 75% of trials ($n = 24$; binomial test, expected = 0.25; $P < 0.0001$; Fig. 2). This reversed the general avoidance of the asynchronous multisensory signal when the calls broadcast from alternative speakers were held constant (both WC). This distribution is similar to female behavior in a standard WC versus W experiment when no robofrog is present (Gridi-Papp et al. 2006). Thus, the presence of the chuck at one call was enough to overcome the unattractiveness of the asynchronous signal when the alternative call was just the whine.

In the second experiment, we presented females with two identical calls consisting of the whine only. The speaker with the robofrog lagged the call (W-robo). Here, females also did not exhibit an overall aversion to the asynchronous multisensory signal. They chose it 60% of the time, significantly more often than expected ($n = 40$; binomial test, expected = 0.25; $P < 0.0001$; Fig. 2).

In the final experiment, we presented females again with two identical calls consisting of a WC. This time, the speaker with the robofrog inflated before the call (robo-WC). Females chose the asynchronous signal 40% of the time ($n = 40$; binomial test, expected = 0.5; $P = 0.21$; Fig. 2). Thus, females did not choose either the visually leading asynchronous multimodal signal or the unimodal call more often that expected from random chance.

Discussion

All else being equal, the presence of a synchronously inflating vocal sac makes a male's call more attractive to females (Taylor et al. 2008). Further, females tend to reject an asynchronous signal when the vocal sac inflation lags the call (Taylor et al. 2011). Male túngara frogs often call in dense choruses and due to physiological constraints cannot alter the timing of vocal sac inflation and call production. Taylor et al. (2011) suggested that the assessment of the vocal sac by females may provide a means of identifying individual callers with a chorus, much like a human reads lips at noisy parties (Sumby and Pollack 1954).

In this study, we show that the acoustic and visual signal components of the túngara frog's mating signal interact in complex ways to influence female choice. Since males cannot alter the timing of their audiovisual signals in nature, it seems intuitive that females would recognize any incongruency and adopt a simple rule that rejects any combination that does not match the natural template. Interestingly, there does not appear to be a set "rule" that governs a simple template recognition of signal synchrony by females (Taylor and Ryan 2013).

In our first experiment, where we played an asynchronous multimodal WC versus a unimodal W, females showed virtually no aversion to the asynchronous signal and responded to the WC almost as strongly as the same experiment, absent the visual component (85% preference Gridi-Papp et al. 2006; 75% this study). This indicates that although the asynchronous audio-visual signal is generally aversive, if one call contains a chuck, the asynchrony is still more attractive than an isolated whine.

In nature, chucks always follow whines. Wilczynski et al. (1999), however, showed that female túngara frogs are permissive to the temporal order of whines and chucks. In particular, they found that in stimuli where a chuck artificially preceded a whine, females found this as attractive as one that followed in the natural position. Given the difficult task that females have assigning calls to their source when many males are calling within a small area, one prediction might be that females use the chucks to determine when a call is finished. This should improve a female's ability to assign calls to their source. The data from Wilczynski et al. (1999) suggest that this is not true, at least when a female is presented with only two, spatially separated calling males. So for the acoustic component of the signal, syntax for female receivers is flexible. Farris and Ryan (2011, 2017) also demonstrated that female túngara frogs make relative comparisons when identifying callers acoustically. In a series of experiments, they showed that females perceptually group whines and chucks that are temporally and spatially separated, effectively responding as if the disparate components belong to the same source. Here again, the females show permissiveness for signal variation in time and space. They showed that females more readily group calls that have a smaller spatial separation and non-natural sequence relative to calls with a greater spatial separation but natural sequence (Farris and Ryan 2017). Although females perceptually weight spatial cues more, when multiple cues become available, females integrate these into their perceptual and decision making processes (Farris and Ryan 2011).

When the visual component is added to the signal, syntax becomes more important. In our second experiment where we removed the chucks altogether and just presented females with whines in the acoustic domain, the asynchronous multisensory signal (W-robo) was no longer aversive, and females chose this signal more often than expected. In the absence of the chuck, females are less likely to be influenced by the incongruency. This may indicate that when females are simultaneously evaluating acoustic and visual components, the chuck indicates that the call is finished, and any vocal sac inflation following this call does not belong. Thus like relative comparisons within the auditory domain (Farris and Ryan 2011, 2017), female túngara frogs also appear to make relative comparisons when integrating visual and acoustic cues (for other cross-modal comparisons, see also Halfwerk et al. 2014).

In our final experiment, we presented females with a pair of identical WCs, but at one speaker, the robofrog inflation preceded the call (robo-WC). Females responded to the asynchronous signal statistically as often as the unimodal call only. This suggests that females do not recognize the temporal asynchrony or that their perception of the leading

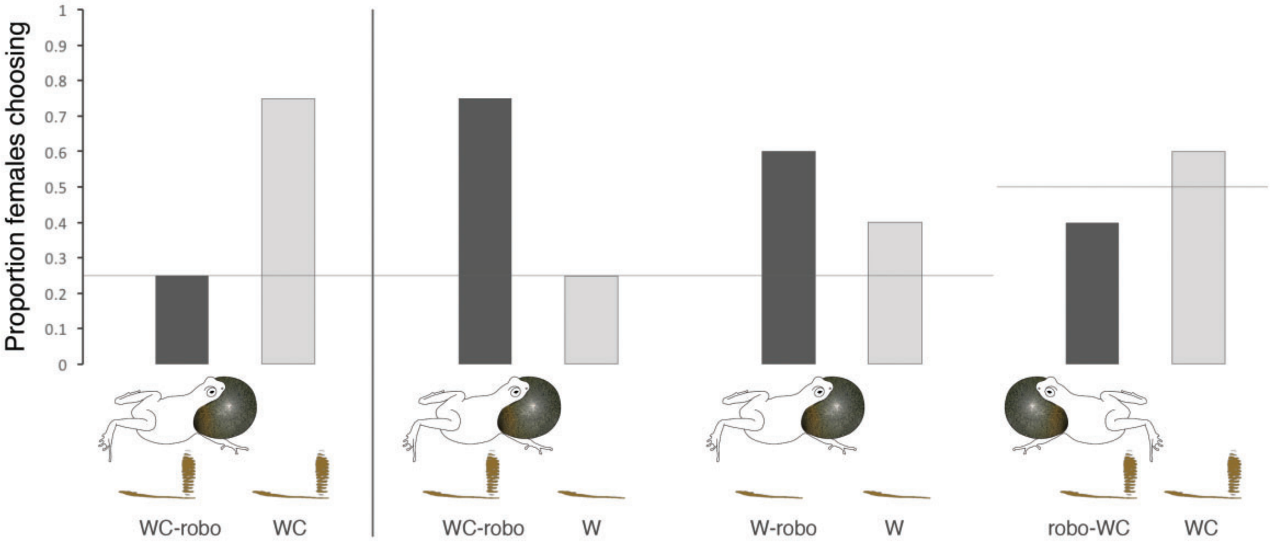

Fig. 2 Proportion of females choosing an asynchronous multimodal signal (audio + visual) versus an alternative unimodal signal (call only). The far left experiment separated by a vertical line is from Taylor et al. (2011) and was used to set prior expectation of asynchrony response at 0.25 (horizontal line). For Experiment 3 on the far right, the expected response was set at 0.5 (horizontal line). The *x*-axis legends refer to the temporal sequence of the stimuli. WC-robo = whine, then chuck, then robotic frog inflation. W-robo = whine, then robotic frog inflation. Robo-WC = robotic frog inflation, then whine, then chuck. The graphic of the timing of the robofrog inflation/sonogram follows from Fig. 1b.

visual signal is less aversive than when it lags the call. Interestingly, this behavior coincides with audiovisual discrepancy detection in human listeners. Human listeners, like many vertebrates, integrate auditory and visual signals and generate perceptions of synchrony as part of their overall auditory scene analysis (Stein 2012; Farris and Ryan 2017). Humans more easily detect asynchrony when a visual cue lags an acoustic signal versus one that leads (Dixon and Spitz 1980). Given that light travels dramatically faster than sound, audiovisual discrepancies occur in nature with increasing communication distances. Specifically, since sound naturally lags a visual stimulus, it might be expected that receivers, humans or otherwise, are somewhat permissive of lagging sound. For example, Navarra et al. (2009) showed that human listeners increased reaction times to audio signals that lagged the visual cue, but were unable to do this for lagging visual signals. They suggested this effect may result from auditory processing plasticity that can compensate for the normal temporal lag that occurs in nature, thereby improving the ability of the brain to bind relevant audiovisual cues into a coherent stimulus (also see Sugita and Suzuki 2003). Given stimulus transmission and neural transduction speeds, communication distances need to exceed about 10 m before audio signals begin to perceptually lag visual signals (Pöppel and Artin 1988) and human listeners remain unaware of asynchronies until the audio stimulus lags the visual by about 250 ms (Dixon and Spitz 1980).

Our results have important implications for our understanding of sensory ecology, perception, and multimodal signal evolution. First, for nocturnally communicating frogs that use multimodal signals, the evaluation distance is nearly always less than 10 m (personal observation). Thus, a female receiver is unlikely to experience a noticeable audiovisual asynchrony produced by a particular calling male. In light of this, there is no ecological reason why female frogs should be more sensitive to a lagging visual signal versus a lagging audio signal. Our data show that they are, however. One explanation may be that neural integration of auditory and visual signals, particularly the perception of synchrony, is a conserved process across many vertebrate taxa. In particular, if the vertebrate auditory processing is more plastic than the visual system (Navarra et al. 2009), then this may constrain receivers to be permissive of lagging audio signals, irrespective of whether they experience them in nature.

The second implication of our results is that contextual aspects of audiovisual integration may be as important as temporal structure per se. For túngara frogs, the chuck component of the call must be accompanied by the whine in order for females to even recognize it as a salient signal (Ryan 1985). Even so, once the context is set (e.g., the presence of the whine), the chuck strongly modulates female attraction, making the complex call five times more attractive as the whine only (Gridi-Papp et al. 2006).

The presence of the chuck also overrides the aversive nature of the lagging visual signal. Likewise, when the chucks are removed completely, females are no longer averse to the temporal asynchrony. In sum, females are permissive to variation in call syntax when presented with a call only (e.g., chuck precedes whine) and they are permissive of multisensory asynchrony when chucks are absent. The presence of the chuck, however, alerts females to the asynchrony of the multisensory signal (when the visual cue lags a standard complex call), and modulates their behavior.

We suggest that future studies of multimodal signaling should include experiments that are not only signal isolation tests (*sensu* Partan and Marler 2005), but also explore how different arrangements of both context and timing influence receiver behavior. Doing so is likely to reveal the full range of multisensory space over which receivers recognize and respond to conspecific signals (Smith and Evans 2013), including variations that don't naturally occur. This will provide insights into how neural integration and sensory perception can promote or constrain the evolution of complex signal design.

Acknowledgments

Joey Stein and Moey Inc. developed the robotic frog control system. We thank Nic Stange, Kyle Wilhite, and Kelsey Mitchell for help with data collection. We are grateful to the Smithsonian Tropical Research Institute for logistical support. Constructive criticism from one anonymous reviewer improved the quality of the manuscript. The work was conducted under STRI IACUC protocol No. 2011-0825-2014-02 and collecting permit from Panama's Autoridad Nacional del Ambiente (ANAM).

Funding

This work was supported by a National Science Foundation grant [IOS 1120031 to R.C.T., M.J.R., and R.A.P.].

References

Agresti A. 2001. Exact inference for categorical data: recent advances and continuing controversies. Stat Med 20:2709–22.

Bee MA. 2015. Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. Int J Psychophysiol 95:216–37.

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.

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 211:1203–10.

Dixon NF, Spitz L. 1980. The detection of auditory visual desynchrony. Perception 9:719–21.

Farris HE, Ryan MJ. 2011. Relative comparisons of call parameters enable auditory grouping in frogs. Nat Commun 2:410.

Farris HE, Ryan MJ. 2017. Schema vs. primitive perceptual grouping: the relative weighting of sequential vs. spatial cues during an auditory grouping task in frogs. J Comp Physiol A 203:175–82.

Gridi-Papp M, Rand AS, Ryan MJ. 2006. Animal communication: complex call production in the túngara frog. Nature 442:257.

Halfwerk W, Page RA, Taylor RC, Wilson PS, Ryan MJ. 2014. Crossmodal comparisons of signal components allow for relative-distance assessment. Curr Biol 24:1751–5.

Hebets EA, Barron AB, Balakrishnan CN, Hauber ME, Mason PH, Hoke KL. 2016. A systems approach to animal communication. Proc R Soc B Biol Sci 283:20152889.

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

Higham JP, Hebets EA. 2013. An introduction to multimodal communication. Behav Ecol Sociobiol 67:1381–8.

Klein BA, Stein J, Taylor RC. 2012. Robots in the service of animal behavior. Commun Integr Biol 5:466–72.

Navarra J, Hartcher-O'Brien J, Piazza E, Spence C. 2009. Adaptation to audiovisual asynchrony modulates the speeded detection of sound. Proc Natl Acad Sci U S A 106:9169–73.

Partan S, Marler P. 1999. Behavior–communication goes multimodal. Science 283:1272–3.

Partan SR, Marler P. 2005. Issues in the classification of multimodal communication signals. Am Nat 166:231–45.

Phelps SM, Rand AS, Ryan MJ. 2005. A cognitive framework for mate choice and species recognition. Am Nat 167:28–42.

Pöppel E, Artin TT. 1988. Mindworks: time and conscious experience. Harcourt Brace Jovanovich.

Rubi TL, Stephens DW. 2016. Why complex signals matter, sometimes. In: Bee MA, Miller C, editors. Psychological mechanisms in animal communication. New York (NY): Springer. p. 119–35.

Ryan MJ. 1985. The túngara frog: a study in sexual selection and communication. Chicago: University of Chicago Press.

Ryan MJ. 2001. Anuran communication. Washington, DC: Smithsonian Institution Press.

Ryan MJ, Guerra MA. 2014. The mechanism of sound production in tungara frogs and its role in sexual selection and speciation. Curr Opin Neurobiol 28:54–59.

Ryan MJ, Rand W, Hurd PL, Phelps SM, Rand AS. 2003. Generalization in response to mate recognition signals. Am Nat 161:380–94.

Smith CL, Evans CS. 2013. A new heuristic for capturing the complexity of multimodal signals. Behav Ecol Sociobiol 67:1389–98.

Stange N, Page RA, Ryan MJ, Taylor RC. 2016. Interactions between complex multisensory signal components result in unexpected mate choice responses. Anim Behav 116:83–7.

Stein BE. 2012. The new handbook of multisensory processing. Cambridge (MA): MIT Press.

Sugita Y, Suzuki Y. 2003. Audiovisual perception: implicit estimation of sound-arrival time. Nature 421:911.

Sumby WH, Pollack I. 1954. Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26:212–5.

Taylor RC, Klein BA, Stein J, Ryan MJ. 2008. Faux frogs: multimodal signalling and the value of robotics in animal behaviour. Anim Behav 76:1089–97.

Taylor RC, Klein BA, Stein J, Ryan MJ. 2011. Multimodal signal variation in space and time: how important is matching a signal with its signaler? J Exp Biol 214:815–20.

Taylor RC, Ryan MJ. 2013. Interactions of multisensory components perceptually rescue túngara frog mating signals. Science 341:273–4.

Wilczynski W, Rand AS, Ryan MJ. 1999. Female preferences for temporal order of call components in the tungara frog: a Bayesian analysis. Anim Behav 58:841–51.