

Joshua J. Schwartz · Bryant W. Buchanan
H. Carl Gerhardt

Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments

© Springer-Verlag 2001

Abstract We studied female mate choice by *Hyla versicolor* in three venues to examine how acoustic and spatial complexity, background noise, and the calling behavior of males might influence preferences manifest in previous laboratory two-stimulus choice tests. Our laboratory-based two-stimulus choice tests with and without broadcasts of chorus noise demonstrated that females prefer long calls relative to short calls when calling efforts of alternatives are equivalent. Background noise impaired the ability of females to discriminate in favor of longer over shorter calls, but the magnitude of the effect was small. Captures of females at eight speakers broadcasting 6- to 27-pulse calls at the edge of a pond revealed strong discrimination against only the shortest call variant. In natural choruses, females may only rarely encounter males using such unattractive vocalizations. Female phonotaxis at an artificial pond with caged and electronically monitored calling males also indicated that consequences of female preferences for temporal aspects of calling observed in two-stimulus choice tests are much attenuated in choruses and explain only small portions (<10%) of the variation in male mating success. Nevertheless, relatively high call duration and calling effort increased male attractiveness. Acoustic interference emerged as another significant factor influencing male mating success and possibly the differences in female choice observed in laboratory and chorus settings. We suggest that the bias of females against both overlapped

and very short calls may help explain why males lengthen their calls but lower their rate of delivery in response to increases in chorus size.

Keywords Frog · Female · Mate choice · Chorus · Pattern recognition

Introduction

Attributes of the study organism, the particular topic of interest, and the resources of the investigators will usually determine whether questions about animal behavior are addressed in the field, laboratory, or both settings. The typically reduced complexity of the environment in the laboratory compared to a natural setting has both advantages and disadvantages. For example, in studies of mate choice in anuran amphibians, the behavior of females has often been tested in an arena or chamber isolated from the natural chorus (Ryan 1985; Gerhardt 1995; Murphy and Gerhardt 1996). The main advantage of this approach is the ability to work in a stable environment with a relatively quiet acoustic background where the experimenter can present alternative stimuli whose differences are precisely controlled at the female's decision point. A disadvantage of this approach is that patterns of discrimination observed in the laboratory are likely to exaggerate the selectivity of the behavior in the field, where extraneous variables can influence choices or mask them. Indeed data, unfortunately available for only a small number of species, argue for caution when extrapolating from laboratory results on female preferences to patterns of mate choice in the field (Gerhardt 1982; Gerhardt et al. 1987; Sullivan and Hinshaw 1992; Bertram et al. 1996). For example, the extent to which patterns of female pairing may conform with expectations based on female preferences for particular call features in two-stimulus choice tests may depend on chorus size (e.g., Gerhardt 1982; Telford et al. 1989; Márquez and Bosch 1997). If loud enough, noise, which is typically considerable in choruses, can so mask communica-

J.J. Schwartz (✉) · H.C. Gerhardt
Division of Biological Sciences, University of Missouri, Columbia,
MO 65211, USA

B.W. Buchanan
Department of Natural Sciences, Stephens College, Columbia,
MO 65215, USA

Present addresses:

J.J. Schwartz, Department of Biological Sciences, Pace University,
861 Bedford Road, Pleasantville, NY 10570, USA
e-mail: jschwartz2@pace.edu
Tel.: +1-914-7733507, Fax: +1-914-7733634

B.W. Buchanan, Department of Biology,
Utica College of Syracuse University, Utica, NY 13502, USA

tion signals that a receiver may not detect conspecific signals (Gerhardt and Klump 1988; Narins and Zelick 1988; Narins 1992; Simmons et al. 1992) or discriminate among them (Wollerman 1995). The spectral characteristics of the noise relative to the spectral content of the signals determine its impact on both of these tasks (Schwartz and Wells 1983). Background noise can also induce threshold shifts as well as shifts in the linear operating range of auditory fibers (Young et al. 1983; Narins and Zelick 1988; Schwartz and Gerhardt 1998). Even if background noise levels are low, the fine-scale timing of signals given by males in a chorus can, by degrading the species-specific pulse rate of calls of some species, have a profound effect on mate choice decisions by females (Schwartz 1987).

One way to augment the typical laboratory-based female choice experiments with frogs is to gradually add to the testing environment elements of a natural chorus such as background noise (Schwartz and Gerhardt 1998), structural complexity (Gerhardt and Rheinlaender 1982; Passmore et al. 1984), and additional sources of call stimuli (Gerhardt 1982; Schwartz 1994). Another approach is to make observations of natural mate choice in the field. This should be done in a way that is both non-intrusive and allows for subsequent analysis of male characteristics that influence female decisions (Passmore et al. 1992; Schwartz et al. 1995; Sullivan et al. 1995; Grafe 1997a). Two additional methodologies incorporate elements of this field approach and tests in a laboratory chamber. First, speakers can be used to broadcast calls in the field, giving the researcher control over the alternatives available to females (Forester et al. 1989; Schwartz 2001). Second, gravid females may be released within enclosed groups of males (e.g., Telford et al. 1989) whose calls and behaviors are monitored prior to female choice.

In this study, we examine the choice behavior of female gray treefrogs, *Hyla versicolor*, in three different experimental venues: a laboratory, an enclosed artificial pond where we monitored the calling activity of males, and at a natural pond, where we monitored female attraction to an array of speakers. We could thereby exploit some of the advantages and compensate for some of the disadvantages of each approach. We also gained some understanding of factors in the natural environment that affect a female's ability to select a particular caller in the chorus.

Males of the gray treefrog advertise for mates in aggregations in which the acoustic environment can be quite complex. Not only are numerous calling males present, but the signals of males overlap and affect the call duration and call timing of neighboring callers (Wells and Taigen 1986; Schwartz 1987; Klump and Gerhardt 1992). The spatial arrangement of males and the habitat adds additional complexity to the acoustic environment. In our field study site, males advertise from vegetation on the surface of a pond, from the shoreline, or from bushes and trees along its shore from mid April to mid July. Chorusing begins near dusk and typically

wanes near midnight. Gravid females approach the pond from the surrounding forest, perch in vegetation near the shore and, after an indeterminate time, approach and pair with a male.

Laboratory-based two-choice experiments with *H. versicolor* have demonstrated that call duration and call rate, which multiplicatively constitute calling or pulse effort, help to mediate discrimination by females of different conspecific males (Gerhardt et al. 1996 and references therein). The number of pulses in the amplitude-modulated advertisement calls (which determines call duration at a particular temperature) is especially important in this regard, and females may gain indirect benefits in fitness from their preference for males giving longer calls (Welch et al. 1998). In fact, females preferred long calls (24 pulses) to short calls (12 pulses) even when calling effort of the choice stimuli were equivalent (Klump and Gerhardt 1987; Gerhardt et al. 1996). Males maintain a relatively stable calling effort by adjusting call rate when they add or drop pulses from their calls in response to increases or decreases in acoustic stimulation from real or simulated competitors (Wells and Taigen 1986; J.J. Schwartz, B.W. Buchanan, H.C. Gerhardt, unpublished data).

The work described in this paper focuses on those temporal aspects of calling that determine calling effort. Call amplitude was also considered in our evaluation of female choice behavior in the artificial pond. For brevity, "call duration" is often used in lieu of "the number of pulses in a call." In a strict sense, the terminology is not interchangeable for conditions of *different* temperature, because pulse duration and rate are temperature dependent (Gayou 1980), and so calls with the same number of pulses can differ in duration.

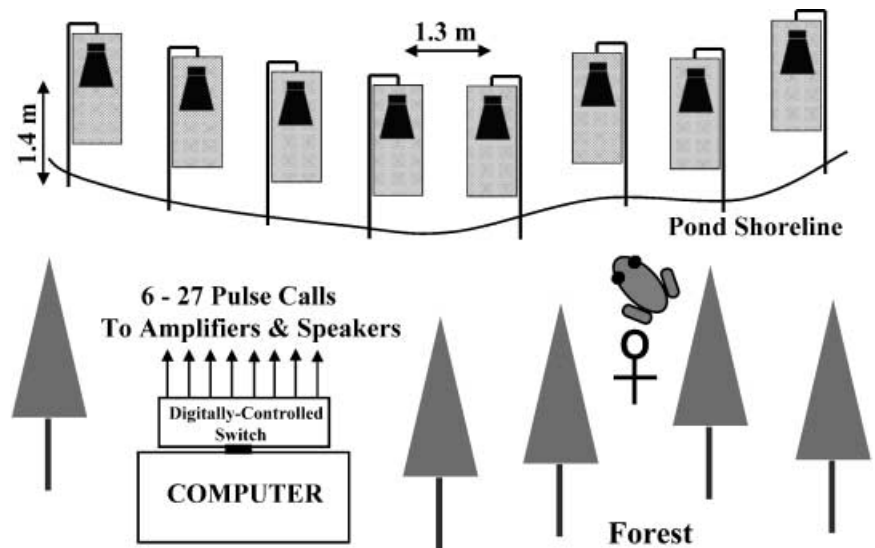
Methods

Discrimination of pulse number in an artificial chorus in the field

We performed two experiments at our field site (the Baskett Wildlife Area near Ashland, Missouri) using an artificial chorus assembled at the edge of the pond (see Fig. 1). In these tests, calls were broadcast from eight 360-degree speakers (Radio Shack catalog no. 40-1312 patio speakers) separated from one another by about 1.3 m. The speakers were individually housed within cylindrical screen cages constructed from plastic needlepoint canvas (3.5×3.5 mm mesh) that were capped at the top and bottom with fiberglass screen. Each speaker was suspended at an elevation of approximately 140 cm from a metal shepherd's hook, and the surrounding trap was supported by three metal stakes. Females, attracted from the adjacent forest to the broadcast calls, gained access to a speaker via one of nine windows with inward-facing flaps cut into the plastic mesh.

We used a Commodore Amiga 500 computer and custom software to present the synthetic calls, which were routed from the two audio output lines of the computer to individual speakers with a custom digitally controlled analog switch attached to the parallel port of the computer. The computer was powered with a 17-Ah 12-V sealed rechargeable battery and a 140-W DC-to-AC inverter (Radio-Shack 22-132). We used two four-channel Clarion APA 4202 75-W power amplifiers (powered by a 12-V 24-Ah sealed rechargeable battery) to amplify the calls, and adjusted the stimulus intensity to 88±2 dB sound pressure level (SPL; Fast RMS,

Fig. 1 Schematic diagram of the equipment used to simulate a chorus of eight males calling at the edge of the pond at our study site in Ashland, Missouri. Gravid females typically approached the pond and speaker array from the bordering forest. The two four-channel amplifiers, DC-to-AC inverter, and two batteries are not shown



C-weighting) at 1 m with a Radio-Shack 33-2050 sound level meter and individual gain controls that had been installed on each speaker. We ensured that the peak SPL from all speakers was equivalent by making the gain adjustments while broadcasting a 27-pulse call from each speaker.

The stimulus set consisted of eight non-overlapping synthetic calls ranging from 6 to 27 pulses in length (in steps of 3 pulses) randomly assigned to each of the speakers prior to testing every night. The durations of the stimulus calls encompassed over 97% of the natural range of variation in average pulses per calls of males recorded at our field site (mean=17.9, range=7.4–32, $n=349$ males). We created the stimuli (J.J. Schwartz, unpublished software) using design parameters of calls synthesized for our laboratory-based tests of female phonotaxis (e.g., Gerhardt et al. 1996). Because the pulse rate of calls is positively correlated with body temperature in *H. versicolor* (Gerhardt 1978), we measured the cloacal temperature of two calling males near our speaker array before beginning the playbacks. The mean temperature determined which one of a set of previously synthesized 27-pulse calls we loaded into the memory of the computer.

We started experimental broadcasts shortly after males began to call at the pond (8:00–8:30 p.m.) and continued until midnight. Just prior to and during the playbacks, we removed males calling in the vicinity of our array of speakers and checked the traps for females approximately every 20 min. We removed captured females from each cage, placed them in a marked plastic box, and recorded the stimulus identity and times of capture.

We ran test 1 on 27 nights from 11 May through 12 June 1998 and on 14 nights from 1 May through 22 June 1999. In this test, calls were presented at equal average rate (9.1/min) from each speaker; therefore, calling effort was unequal. The randomized broadcast sequence from the speakers was rerandomized after all speakers had broadcast their call (a chorus bout); this resulted in some variation in call period between bouts. We ran test 2 on 29 nights from 14 June through 17 July 1998 and on 36 nights from 3 May through 11 July 1999. In this test (unequal call rate, equal calling effort), the mean pulse effort per speaker was set to the average used in test 1 (150 pulses/min) and the call sequence was randomized among bouts of our simulated chorus.

Arena-based two-choice tests

General methods

In all tests, as described above, the synthetic stimulus calls resembled the natural calls of *H. versicolor*. Additional details on synthe-

sis and a description of the choice arena and general playback procedures are described in Gerhardt (1994) and Gerhardt et al. (1996).

Best discrimination of pulse number

We tested females with synthetic advertisement calls in a quiet background and one with natural chorus noise. A stereo tape of the noise was created by digitizing (16 bits/sample, 22.05 kHz) and digitally looping a 55-s segment of a stereo recording of the chorus at the Ashland field site. We broadcast it from two Analog-Digital-Systems 200 loudspeakers, separated by 92 cm, and suspended from brackets 165 cm above the release point on the floor of the testing chamber at 77 dB SPL (Fast RMS, C-weighting; Larsen-Davis 800B). This amplitude was close to the average value measured adjacent to our field pond during chorusing (mean=77.6 dB, median=75 dB, $n=74$ nights; 9:30–10:30 p.m.). The call stimuli, broadcast with a call period of 4.5 s at an SPL of 85 dB (measured at the central release point of females), ranged from 14 to 20 pulses in length. Each female was initially given a choice between calls of 18 and 20 pulses. If the female approached the shorter of the two calls, the next choice was made easier by reducing the pulse number of the shorter call by two pulses. If the female approached the longer of the two calls, the choice was made more difficult by increasing the length of the shorter call by one pulse. Using this up-down staircase procedure, we estimated the discrimination ability of females for pulse number differences with possible stimulus pulse number pairings of 14 versus 20, 15 versus 20, 16 versus 20, 17 versus 20, 18 versus 20, and 19 versus 20. If a female stopped responding before the test protocol was completed, results for that individual were not used. We alternated whether we started testing with or without the background noise among testing periods and frequently switched the side of the choice arena from which we broadcast the shorter calls.

Pulse discrimination with equal pulse effort stimuli

We gave females a choice between calls of 12 and 18 pulses timed such that their mean pulse effort was identical (240 pulses/min). Therefore, the broadcast rate of the shorter call was 1.5 times that of the longer call (20 calls/min versus 13.33 calls/min). The computer software (SiliconSoft Inc., San Jose, Calif.) we used for our arena playbacks loops a fixed sequence of calls from two audio channels. To obtain alternative stimuli with equal call rates, we typically adjust call timing such that the intervals between the midpoints of alternative calls on the two playback channels are

Fig. 2 The two arrangements of the timing of calls used in the two-stimulus female choice tests (EPE-1, EPE-2) in which the 12- and 18-pulse alternatives were broadcast with equal pulse efforts

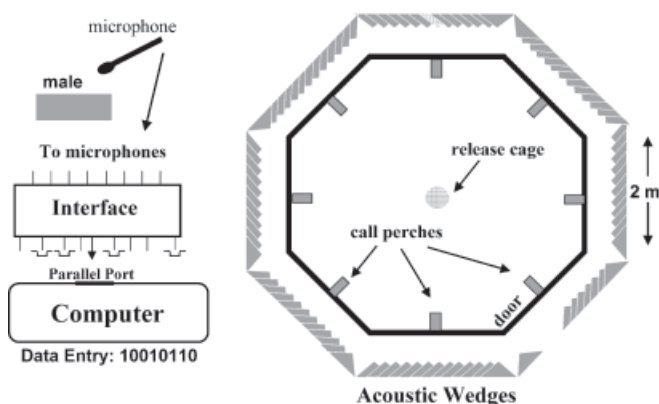
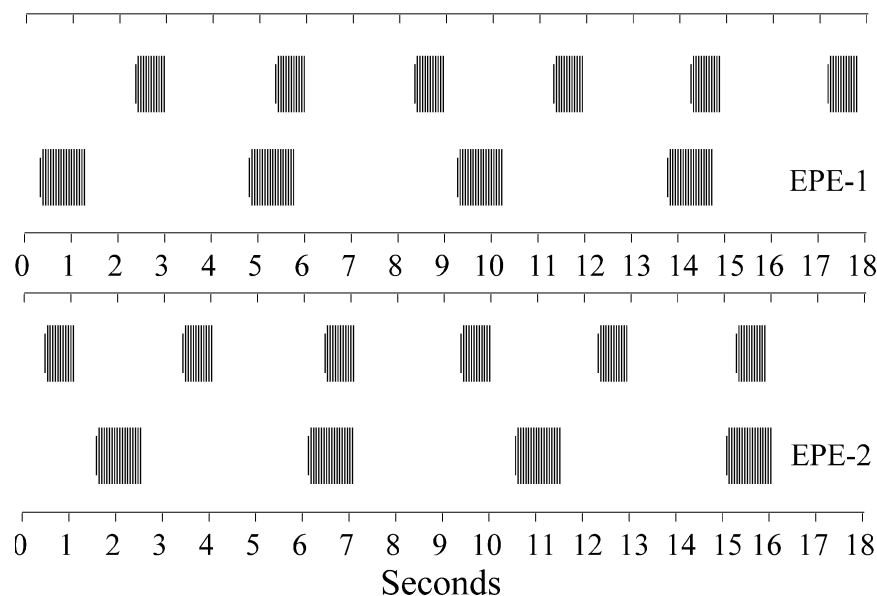


Fig. 3 Schematic diagram of the artificial pond in the greenhouse as viewed from above (*right*) and operation of the data acquisition system. We monitored the calls of up to eight males with a computer and an eight-channel custom-built interface board. Above each of the cinderblock perches, we suspended a unidirectional microphone whose cable was run to the computer interface outside the enclosure. When frogs called, voltage transitions were produced by the interface, as shown, coincident with each pulse of each male's call. The voltage levels on each of the eight lines were detected via the corresponding data lines of the parallel port of the computer and encoded as either a 0 or 1 in each byte of data sampled by the computer. This process is illustrated for males calling at four positions in the pond

equal. However, with stimuli of unequal call rates such as those used here, this is not possible. Accordingly, one must select a particular static pattern of relative call timing to store in the memory of the computer. Because an arbitrarily chosen pattern of call delivery could conceivably bias our results, we conducted two tests (EPE-1, EPE-2) with different relative timing arrangements of the calls for the equal pulse effort tests (Fig. 2). As described previously for the best discrimination tests, these tests were conducted with and without background chorus noise.

Tests in the artificial pond

In 1995 and 1998, we conducted female choice experiments ($n=21$ and $n=24$ tests, $n=15$, $n=21$ females, respectively) in the artificial

pond (see Fig. 3) with calling males ($n=47$, $n=51$, respectively) at individual calling sites. The octagonal pond (2 m on each side) was constructed within a greenhouse at the University of Missouri and enclosed in screen. Acoustic foam wedges (Soundcoat Inc.) surrounded the enclosure and opaque plastic sheeting served to shield the testing area from light. To monitor the calling behavior of the males, we used the system described in detail in Schwartz (1993). A directional microphone suspended above each call site was connected to a custom battery-powered interface board that sent its output to a computer (Commodore Amiga 600) via the parallel port. Although the system does not encode spectral information, the specific frogs that are calling, or not calling, at a particular time are saved. The rate at which the computer sampled data from the parallel port (100 Hz) and the temporal resolution settings of the hardware interface were sufficient for counting the individual pulses within the calls of each male.

We released up to 30 males within the enclosure up to a day prior to testing. These males had been captured at our field site and marked on the back with a numerical freeze brand. Shortly after sunset, we broadcast (approximately 75-dB SPL, Fast RMS, C-weighting at center of pond) a recording of a natural chorus of gray treefrogs from two speakers (Realistic 40-1242) suspended outside the pond. Each evening, we attempted to conduct our choice tests with a chorus of eight males. However, because it was often not possible to stimulate this number of male frogs to call reliably, choruses ranged in size from four to eight males (four males, $n=3$ choruses; five males, $n=10$ choruses; six males, $n=7$ choruses; seven males, $n=10$ choruses; eight males, $n=15$ choruses; mean=6.53; total $n=45$ tests with the 98 males providing 294 potential mate options for the females). After calling in the chorus was sustained and vigorous, we placed a male on each of up to eight cinderblock perches equally spaced around the periphery of the pond. A small screen cage placed on the perch prevented each male from moving away from his call site. We next removed the extra males from the enclosure, allowed the calling of our subjects to recover and stabilize, and then turned off the background chorus stimulus. Typically, we monitored the calls of males for at least an additional 30 min before initiating our tests with females.

At our study pond in Ashland, we captured the experimental females in amplexus the night prior to pairing, brought them to the laboratory, and refrigerated (2–5°C) them until shortly before testing. In 1998, we also measured SPL (Fast RMS, C-weighting; $n=56$ males, 158 potential mate options, 24 tests) from a centrally placed microphone mounted vertically near the ceiling of the chamber (Radio-Shack 33-2050 with a custom-installed extension cable). We obtained SPL measurements ($n=3$ per male) by moni-

toring the meter and the LEDs on the case of our interface board and noting the SPL when only one male called. This procedure usually required 15–25 min. Values were subsequently corrected for the directionality of the microphone of the meter, which we had previously determined.

For each test, we placed a female in a centrally located release cage constructed of hardware screen. We gave her approximately 2 min to adjust to her new surroundings before raising the lid of the cage with a string leading to the outside of the enclosure. At this time, we started acquisition of calling data on the computer while one of us monitored the female with an AMT-M02 infrared night vision scope and dictated information on her behavior to the other person, who took notes. We defined a choice as contact with the cage of a male. If a female failed to make a choice within 10 min, we usually terminated the test, unless the female was already moving toward a male. If available, we tested additional females after the calling behavior of males recovered following the brief disturbance caused by movements of an experimenter in the pond to retrieve the previous test subject.

We used custom software to calculate the time of occurrence and number of pulses in the calls of each male during each choice test and to calculate the call overlap between chorus members. We analyzed the resulting secondary data files using SAS (SAS Institute, Cary, N.C.) by quantifying aspects of the calling behavior of chosen and not chosen males from the time a female left her release cage until she made a choice. These aspects of calling were mean pulses per call, calls given, the total pulses given, and the proportion of calling time that was free from overlap by any other frogs. For the sample of males for which we had data on call SPL, we compared “successful” and “unsuccessful” individuals for this attribute as well as measures of the first three temporal characteristics weighted by the intensity level of each male’s calls. We obtained the weighting factor by converting SPL in dB to μbars , calculating a mean in μbars for all males, and dividing the SPL in μbars of each male by this mean. Results were also analyzed using logistic regression to determine how much of the variation in male success might be explained by one or more of the call behavior variables.

Results

Discrimination of pulse number in an artificial chorus in the field

The distribution of captures for the eight stimuli did not differ significantly among the two tests ($\chi^2=3.793$, $P=0.998$, $df=15$). In test 1, we captured 81 females. There was a non-random pattern of capture at the speakers broadcasting the different stimuli ($\chi^2=21.82$, $P=0.0027$, $df=7$; Fig. 4a). Most of the difference in our results from the expectation of no discrimination was because calls with 6 pulses attracted only one female ($\chi^2=9.4$, $P=0.0022$, $df=1$). We did, however, detect a significant deviation from the null expectation for the remaining seven stimuli ($\chi^2=14.0$, $P=0.0296$, $df=6$). In test 2, we captured 87 females. Again, we obtained a non-random pattern of captures for the different call stimuli ($\chi^2=17.18$, $P=0.0162$, $df=7$; Fig. 4b), with only one capture for the 6-pulse call. The result for the shortest call also accounted for the bulk of the difference in our results from the null expectation ($\chi^2=10.25$, $P=0.00137$, $df=1$); however, in this test, there was no significant deviation from expectation for the remaining seven stimuli ($\chi^2=6.14$, $P=0.408$, $df=6$).

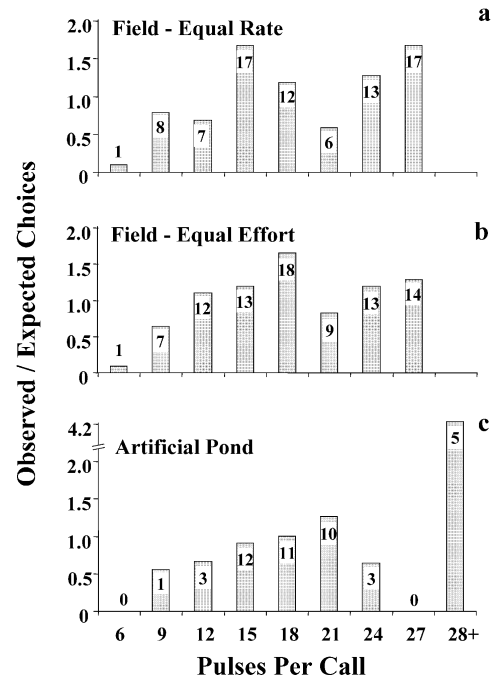


Fig. 4 **a** Captures ($n=81$) of females at the speaker array simulating a chorus of eight males for stimuli of 6–27 pulses per call broadcast at equal call rate. Each histogram bin presents the data as the ratio of observed captures relative to expected captures (10.125 captures). The actual number of females is given inside each bar. **b** Captures ($n=87$) when calls were broadcast at equal calling effort. **c** Results obtained during female choice tests in the artificial greenhouse pond. Mean numbers of pulses per call have been grouped to fall into the eight stimulus categories (bin values are midpoints) used in the two field experiments along with an additional bin for the five males who averaged more than 28 pulses per call. The 6-pulse category includes all values equal to or below 7 pulses per call. The expected number of captures for the 6- and 27-pulse category was 0.325 and 4.63 females, respectively. Expectation was calculated from the individual odds of mating in each chorus (e.g., 20% for a test with five males) summed for each ordinal category over all tests. We obtained ordinal categories using the transformation: pulses per call = integer (pulses per call + 0.5). There was no significant difference between observed and expected matings for this calling feature (all choices: $\chi^2=24.91$, $df=26$; subset of choices: $\chi^2=24.69$, $df=26$).

Arena-based two-choice tests

Best discrimination of pulse number

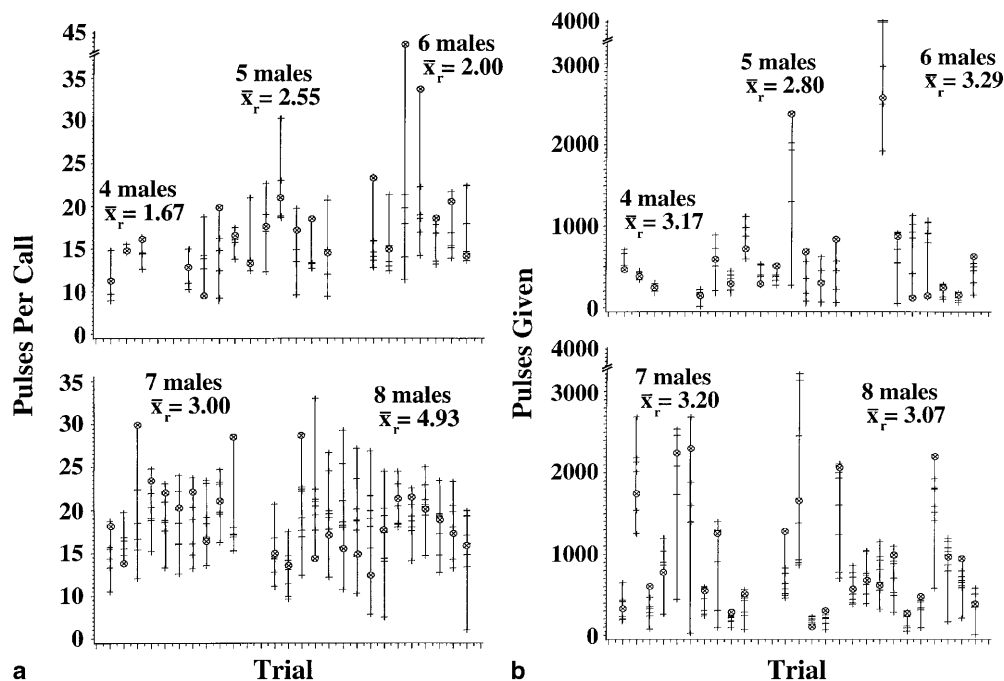
We tested 47 females. Of these, 16 either initially failed to respond or failed to complete the required series of pairwise tests. Of the responding 31 females, 6 exhibited finer discrimination in noise than in quiet, 17 worse discrimination, and 8 showed no change in behavior ($\chi^2=6.65$, $P=0.036$). The best discrimination in quiet averaged 1.52 pulses (median=1), while with the chorus noise it averaged 2.29 pulses (median=2). The average change in discrimination ability in noise relative to quiet was 1.83 pulses (median=1.5) for the 6 females that had improved discrimination in noise and -2.06 pulses (median= -2) for those that had reduced discrimination in noise (mean change for all females= -0.774 pulses).

Table 1 Results of the female choice tests in the artificial pond when data from all tests were pooled before analysis. Values for call features are means. Values in parentheses give the number of males used in the tests that were either successful or unsuccessful in attracting a female. The results labeled *All choices* include data

from females tested more than once (see text). *Z*- and *P*-values (two-sided) are for a Wilcoxon test, and an *asterisk* adjacent to a call feature indicates that there was a significant difference between males that were or were not successful

Call feature	All choices				Subset of choices			
	Successful		<i>Z</i>	<i>P</i>	Successful		<i>Z</i>	<i>P</i>
	No	Yes			No	Yes		
Unweighted by SPL								
Pulses/call	17.3 (249)	18.8 (45)	1.21	0.225	17.2 (199)	19.5 (36)	1.73	0.083
Calls given	45.6 (249)	46.8 (45)	0.63	0.526	43.6 (199)	43.7 (36)	0.36	0.721
Pulses given	744.3 (249)	818.7 (45)	0.98	0.327	716.2 (199)	787.6 (36)	0.82	0.414
SPL	85.8 (134)	86.2 (24)	0.57	0.566	86.2 (118)	86.1 (21)	0.13	0.900
Weighted by SPL								
Pulses/call*	18.5 (134)	22.8 (24)	2.13	0.033	18.4 (118)	23.0 (21)	2.01	0.044
Calls given	38.8 (134)	41.2 (24)	0.58	0.561	39.0 (118)	39.9 (21)	0.36	0.720
Pulses given	685.4 (134)	818.7 (24)	1.29	0.197	682.3 (118)	803.4 (21)	1.13	0.260

Fig. 5 a The mean number of pulses per call given by males in each of the 45 female choice tests in the artificial pond for chorus sizes of four to eight males. A vertical line connects the values for each test and $\otimes$ indicates the male selected by the female. The mean rank ($\bar{x}_r$; 1=highest pulses per call) of chosen males is given for each chorus size. Note that in three tests, the female selected the lowest-ranked caller. These males ranked as follows for other aspects of their calling. Five-male chorus, trial 2: calls=1.5, pulses given=3; seven-male chorus, trial 2: calls=4, pulses given=5; eight-male chorus, trial 4: calls=1, pulses given=1. **b** The total number of pulses produced by each male during these 45 tests



Pulse discrimination with equal pulse effort stimuli

We tested a total of 67 females. Of these, 62 responded at least once. In both EPE-1 and EPE-2, females discriminated in favor of 18-pulse relative to 12-pulse calls with and without chorus noise (EPE-1 without noise, 42 females chose 18-pulse calls, 5 females chose 12-pulse calls, $P < 0.001$; EPE-1 with noise, 32:3, $P < 0.001$; EPE-2 without noise, 27:8, $P = 0.019$; EPE-2 with noise, 23:10, $P = 0.0351$; two-tailed binomial test). Results obtained with the two timing arrangements differed significantly only when the background noise was broadcast ($G = 3.99$, $P = 0.0458$); females were more selective in their responses in tests with the EPE-1 stimulus.

Tests in the artificial pond

We performed analyses using data from all tests (i.e., including the nine tests when a particular female was tested more than once) as well as just the subset of our data that included only the first test with each female. Subsequently, we refer to these two situations as “all choices” ($n = 45$) and “subset of choices” ($n = 36$), respectively. We present the mean pulses per call and pulses produced for the successful and unsuccessful males in each choice test in Fig. 5a, b.

We first performed a coarse-grained analysis by pooling the results obtained from all the separate choice tests (Table 1). We found no significant difference in the

Table 2 Results of the female choice tests in the artificial pond using an analysis based on the z-scores of successful and unsuccessful males. Values in parentheses give the number of males used in the tests that were either successful or unsuccessful in attracting a female. The results labeled *All choices* include data from

females tested more than once (see text). Z- and P-values (two-sided) are for a Wilcoxon test, and an asterisk adjacent to a call feature indicates that there was a significant difference between males that were or were not successful

Call feature	All choices				Subset of choices			
	Successful		Z	P	Successful		Z	P
	No	Yes			No	Yes		
Unweighted by SPL								
Pulses/call*	−0.07 (249)	0.36 (45)	2.69	0.007	−0.09 (199)	0.47 (36)	3.12	0.002
Calls given	−0.03 (249)	0.17 (45)	1.26	0.207	−0.01 (199)	0.08 (36)	0.56	0.575
Pulses given*	−0.56 (249)	0.31 (45)	2.40	0.016	−0.04 (199)	0.24 (36)	1.75	0.080
SPL	−0.003 (134)	0.02 (24)	0.29	0.773	0.12 (118)	−0.69 (21)	−0.250	0.802
Weighted by SPL								
Pulses/call*	−0.08 (134)	0.42 (24)	2.48	0.013	−0.72 (118)	0.41 (21)	2.25	0.025
Calls given	−0.04 (134)	0.24 (24)	1.54	0.124	−0.03 (118)	0.18 (21)	1.22	0.222
Pulses given*	−0.09 (134)	0.51 (24)	3.38	0.001	−0.08 (118)	0.47 (21)	3.07	0.002

pooled distributions of mean pulses per call per frog, calls given, pulses given, and SPL for unsuccessful versus successful males. Weighting by SPL yielded a significant difference for mean pulses per call but did not change the results for the remaining temporal characteristics. Figure 4c shows the data for pulses per call grouped into bins so as to match the stimulus categories used in the field-based choice experiment.

As in a study of mating success in choruses of *H. marmoratus* (Passmore et al. 1992), we obtained a measure of the performance of each male relative to the average performance of chorus members during each test by calculating the z-score for each individual [(male performance–chorus mean)/chorus standard deviation]. This form of rescaling the data is important because males alter their calling behavior in response to changes in the acoustic conditions (e.g., number of other callers: the same males will give shorter calls in a small chorus than in a large chorus; Schwartz 2001). Accordingly, pooling untransformed data before analysis might obscure or exaggerate patterns of female preference in *H. versicolor* obtained on different nights. When we compared the z-scores of successful and unsuccessful males, we found significant differences for all measures of calling behavior (weighted or unweighted by SPL) except calls given and SPL (Table 2, Fig. 6).

We performed two analyses to test whether females might be better able to discriminate at times (1) when there were greater differences among chorus members and (2) when choruses were smaller. Neither of these possibilities was realized. There was no correlation (Spearman rank) between the difference in the calling behavior of a successful male and the mean behavior of unsuccessful males in a chorus and the range of calling behavior present in a chorus (pulses per call: $r_s=0.0286$, $P=0.852$; calls given: $r_s=0.214$, $P=0.159$; pulses given: $r_s=0.194$, $P=0.202$; SPL: $r_s=-0.0414$, $P=0.848$). Nor was there a significant negative correlation between this

difference in calling behavior and chorus size (pulses per call: $r_s=-0.0175$, $P=0.252$; calls given: $r_s=0.278$, $P=0.0643$; pulses given: $r_s=0.200$, $P=0.188$; SPL: $r_s=-0.133$, $P=0.567$). Results were similar when the difference in calling behavior was calculated from the z-scores of calling performance computed for each male during each choice test (pulses per call: $r_s=-0.0193$, $P=0.203$; calls given: $r_s=0.283$, $P=0.0594$; pulses given: $r_s=0.217$, $P=0.153$; SPL: $r_s=-0.121$, $P=0.573$). This analysis controlled for differences in the variability among test choruses in the measures of calling behavior.

When we examined call overlap in our choruses, we found a significant difference between successful and unsuccessful males. Successful males had a greater proportion of their total calling time free from overlap with the calls of other chorus members than unsuccessful males. This was the case when the data were pooled from all choice tests (all choices: unsuccessful males, $\bar{x}=0.41$, successful males, $\bar{x}=0.51$, $Z=3.39$, $P=0.0007$; subset of choices: $\bar{x}=0.42$, $\bar{x}=0.51$, $Z=2.97$, $P=0.003$; Wilcoxon two-sample test). The results of an analysis based on z-scores were similar (all choices: 34 of 45 chosen males had positive z-scores, $P=0.00082$; subset of choices: 26 of 36 chosen males had positive z-scores, $P=0.0113$; two-tailed binomial test). Successful males also had a higher average number of pulses per call that were free from overlap (i.e., unmasked by the call of another frog) than did unsuccessful males (all choices: $\bar{x}$ of means=6.11, 8.09; $Z=3.37$, $P=0.0007$; $n=249$, 45; subset of choices: $\bar{x}$ of means=6.28, 8.37; $Z=3.108$, $P=0.0019$; $n=199$, 36; Wilcoxon two-sample test, data pooled from all tests). The z-score analysis was consistent, indicating that significantly more successful males had above- rather than below-average pulses per call that were not overlapped than did unsuccessful males (all choices: 31 of 45 chosen males had positive z-scores, $P=0.0169$; subset of choices: 26 of 36 chosen males had positive z-scores, $P=0.0113$). Successful males also had significantly

Table 3 Results of logistic regression analyses of choice tests in the artificial pond using just a single independent variable at a time. The variables were z-scores (see text) of call parameters of males ranked within each test. Values not in parentheses are for all choices (not weighted by SPL: $n=45$ tests, 45 successful males, 249 unsuccessful males; SPL or weighted by SPL: $n=34$ tests, 134

unsuccessful males). Values in parentheses are for the subset of choices (not weighted by SPL: $n=24$ tests, 134 unsuccessful males; SPL or weighted by SPL: $n=21$ tests, 118 unsuccessful males). Each P -value is for a test (using $-2 \log$ likelihood, distributed as a χ^2) of the null hypothesis that the regression coefficient of the model is zero

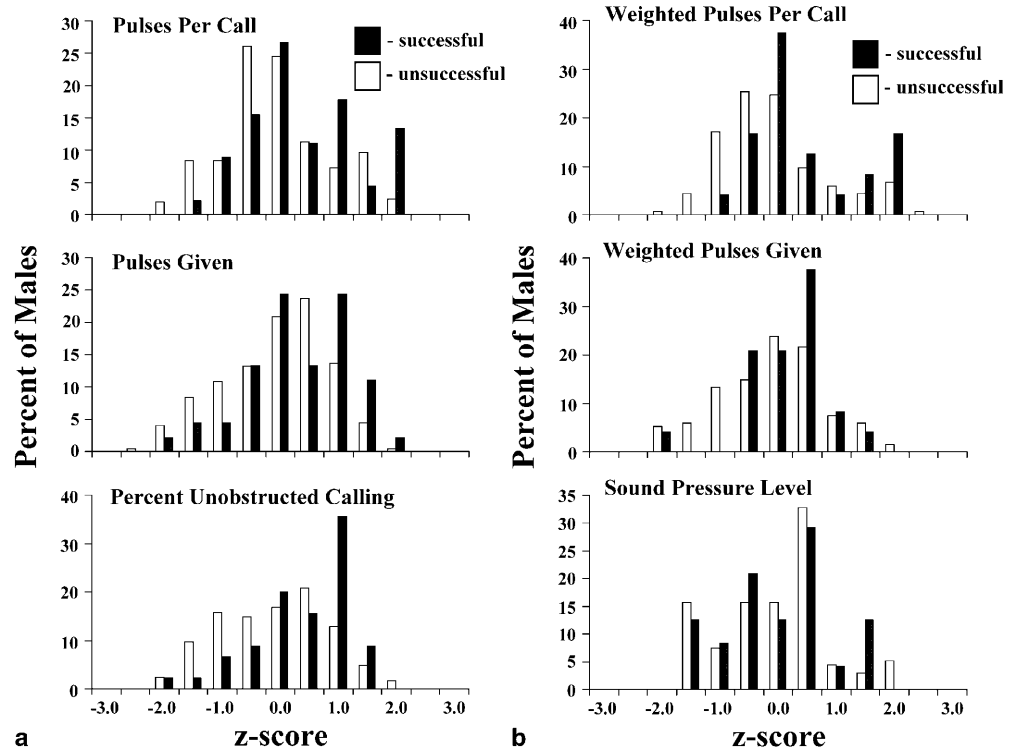
Call variable ^a	RSQ ^b	P	Odds ratio ^c
Percent unobstructed	0.053 (0.048)	0.025 (0.010)	1.74 (1.69)
All pulses	0.037 (0.022)	0.012 (0.085)	1.60 (1.43)
All pulses (weighted)	0.099 (0.833)	0.002 (0.009)	2.33 (2.05)
Pulses/call	0.048 (0.081)	0.004 (0.001)	1.66 (1.94)
Pulses/call (weighted)	0.061 (0.013)	0.018 (0.034)	1.72 (1.68)
Calls	0.011 (0.002)	0.177 (0.595)	1.27 (1.11)
Calls (weighted)	0.013 (0.007)	0.157 (0.315)	1.42 (1.30)
SPL	0.000 (0.002)	0.918 (0.710)	1.03 (0.91)

^a *Percent unobstructed* percentage of calling time unobstructed by the calling of any other males, *All pulses* total pulses produced, *All pulses (weighted)* total pulses produced weighted by relative call intensity, *Pulses/call* mean pulses per call, *Pulses/call (weighted)* mean pulses per call weighted by relative call intensity, *Calls* total calls produced, *Calls (weighted)* total calls produced weighted by relative call intensity

^b RSQ gives the maximum rescaled R^2 . As calculated by SAS, R^2 is the coefficient of determination generalized by Cox and Snell (1989) for general linear models. Since this R^2 cannot attain a value of 1 for a discrete model, SAS also provides the coefficient of determination rescaled [as $R^2/\max(R^2)$] such that a maximum of 1 is possible (see Nagelkerke 1991)

^c *Odds ratio* gives the change in the odds of attracting a female with a one-unit increase in the value of the call variable

Fig. 6 **a** z-Scores (see text) of successful and unsuccessful males in the female choice tests ($n=45$) in the artificial pond calculated for three features of calling behavior. **b** z-Scores for SPL and two features of calling behavior when weighted by SPL ($n=24$ tests)



more clear pulses per call than unsuccessful males (all choices: $Z=3.21$, $P=0.007$; subset of choices: $Z=2.79$, $P=0.005$; Fig. 6).

In the analyses using logistic regression, we used the z-scores obtained from the earlier analysis because this circumvented the problems created by changes in chorus size and other factors that might influence calling behavior of males among tests. Using the z-scores of the calling parameters individually (Table 3), less than 10% of the total variance in choice results was explained by any

independent variable using either the data that were or were not weighted by call SPL. The results indicated that call overlap, pulses per call, and calling effort had a significant effect on a male's success. In a stepwise logistic regression, only the z-scores of both pulses per call and percentage of calling time free of overlap entered the model at an acceptable level of significance for the call variables unweighted by SPL. This model explained close to 10% of the variation in male success in all tests ($r^2=0.083$; subset of tests: $r^2=0.111$; Table 4). When the

Table 4 Results of stepwise logistic regression analysis of choice tests in the artificial pond using more than one independent variable. The variables were z-scores (see text) of call parameters of males ranked within each test. Model 1: success=percent unobstructed pulses/call. Model 2: success=percent unobstructed all pulses (weighted) pulses/call (weighted) (all choices). Model 3: success=percent unobstructed pulses/call (weighted) (subset of choices). Values not in parentheses are for all choices [not weighted by SPL (model 1): $n=45$ tests, 45 successful males, 249 un-suc-

cessful males; weighted by SPL (model 2): $n=34$ tests, 134 unsuccessful males]. Values in parentheses are for the subset of choices [not weighted by spl (model 1): $n=24$ tests, 134 unsuccessful males; weighted by SPL (model 3): $n=21$ tests, 118 unsuccessful males]. The P -value of each model is for a test (using $-2 \log$ likelihood, distributed as a χ^2) of the null hypothesis that the regression coefficients of the model are zero. The P -values for the independent variables are for the Wald χ^2 -statistic with 1 df . The variables are defined in Table 3

Call variable	RSQ	P	Odds ratio
Model 1	0.083 (0.111)	0.001 (<0.001)	
Percent unobstructed		0.016 (0.043)	1.61 (1.57)
Pulses/call		0.024 (0.004)	1.49 (1.79)
Model 2	0.182	0.001	
Percent unobstructed		0.032	1.87
All pulses (weighted)		0.050	1.68
Pulses/call (weighted)		0.050	1.66
Model 3	(0.136)	(0.004)	
Percent unobstructed		(0.013)	(2.02)
Pulses/call (weighted)		(0.037)	(1.70)

data for all choices were weighted by SPL, call overlap, pulses per call, and calling effort together explained 18.2% of the variation in male success. Using just the subset of choices, only pulses per call and call overlap entered the model at an acceptable level of significance and together explained 13.6% of the variation in success.

Discussion

Females of *H. versicolor* clearly have the ability to discriminate subtle differences in call duration. This is known from previous tests in our laboratory (Gerhardt et al. 1996, 2000) and was evident in our tests of best discrimination without noise in which we observed a mean difference of less than 2 pulses for calls close to the mean duration in our population (18 pulses). With background noise that simulated a typical chorus, this ability of females deteriorated only slightly (by less than 1 pulse/call) in our arena-based two-stimulus choice tests. In a natural chorus, noise levels are spatially and temporally heterogeneous, and some males give calls considerably longer and other males considerably shorter than those used in this experiment. Although the calling performance of males in a chorus relative to one another tends to be quite stable (Runkle et al. 1994; Schwartz 2001; J.J. Schwartz, B.W. Buchanan, H.C. Gerhardt, unpublished data), from moment to moment the ranking of individuals in the chorus can shift. Depending on the acoustic conditions or situation, females may or not respond to these short-term shifts in relative performance (see Dyson et al. 1994; Gerhardt and Watson 1995; Gerhardt et al. 1996). In a chorus, moreover, females will often choose their mate when more than two individuals are calling near to one another. Therefore, there may be situations where the females' ability to exercise the fine degree of discrimination exhibited in our laboratory will be obscured from observers or actually impaired. Evidence supporting this statement is

available from other experiments we performed in this study.

In both natural choruses and playback tests, *H. versicolor* males add pulses to their calls in response to increasing levels of acoustic stimulation (Wells and Taigen 1986). However, *H. versicolor* males compensate for the increase in call duration with a decrease in rate (Wells and Taigen 1986). If a male is to benefit from a shift to the longer calls he gives in response to the increased vocal activity of neighbors, females must give more weight to the former than the latter aspect of calling. This prediction of Wells and Taigen (1986) was first confirmed by Klump and Gerhardt (1987). Using alternative stimuli with duty cycles (i.e., pulse efforts) equal to within 1–2%, they found that females discriminated in favor of longer calls when these were twice as long as the shorter calls (24 versus 12 pulses). However, when the difference in duration was only 50% (18 versus 12 pulses), females failed to discriminate. In these tests, the relative timing of the alternative calls was not fixed but varied. The results of our tests using equal pulse effort stimuli with two timing arrangements of calls augment the earlier findings of Klump and Gerhardt (1987). The data indicate that the preference for longer calls is quite robust and can extend to less extreme differences in call duration (e.g., 18 versus 12 pulses), provided the relative timing of the alternative calls is static. Furthermore, the preference for longer calls is maintained even in the presence of background noise. The stronger preference for the 18-pulse call with the EPE-1 stimulus calls than EPE-2 stimulus calls, which was significant with noise, may be related to the near alignment of pulses in the first pair of overlapped calls within the set of EPE-1 but not the EPE-2 stimulus calls. This timing arrangement would tend to preserve the inherent pulse structure of the overlapped calls.

The results from the choice experiment in the field are consistent with previous data from two-stimulus choice tests in that they show duration can be an impor-

tant choice criterion. This interpretation is strengthened by the similarity in the distribution of captures for the tests using stimuli with equal and unequal calling efforts. With equal calling efforts, long-term (i.e., over minutes) differences in the sound energy broadcast among the stimulus sources were absent and so one might have expected a more even distribution of captures than with unequal calling efforts. An important difference between the field and two-stimulus choice test results was that in the field venue, the extent of the preference for longer calls was quite circumscribed. Our data indicated that females discriminate against calls of 6 pulses when confronted with a range of longer alternatives. Under the test condition of equal calling efforts, there were no significant differences in the relative attractiveness of the remaining call durations (9–27 pulses), although the call of 9 pulses attracted half as many females as the 27-pulse call. We randomized the broadcast positions of the different stimuli among nights and the sequence with which the stimuli were broadcast among bouts. Therefore, our results are not the consequence of some particular spatial order or temporal pattern in our array of simulated males. In the field experiment, we did not vary the number of speakers or their separation; such an approach could help elucidate the spatial scale of female choice (Morris 1989) and might reveal situations in which there is stronger evidence of discrimination based on call duration than described here.

Despite the profound bias against very short calls, in natural choruses of more than two males, females may rarely, if ever, exercise this preference because males in choruses rarely, if ever, give calls of 6 pulses or less (see Fig. 7). Indeed, the data from the choice tests in the artificial pond using real males revealed only weak discrimination based on average pulses per call. This was evident in the analysis based on the z-scores of successful and unsuccessful males, while the logistic regression analysis indicated that pulses per call explained just a small portion of the total variation in male success. In fact, females selected males that gave fewer than the average number of pulses per call in 17 of 45 tests and occasionally even selected males that were ranked at the bottom of their chorus (Fig. 5a). In addition to a contribution of pulses per call, the data from the artificial pond suggest that females in choruses are also responsive to differences in calling effort by males (total pulses given), especially when any differences are augmented by differences in call intensity. Thus, males that produce more sound energy are at a slight advantage. SPL did not, on its own, play a significant role in mate choice decisions by females. This is not too surprising, as the modal range in SPL in our choice tests was only 2.5 dB (median=4.5). Interestingly, when females came near or climbed upon the cages confining males, the inverse relationship between call duration and call rate could briefly be suspended and males typically increased their calling effort by lengthening their calls, as first reported by Fellers (1979a), and increasing call rate. Although it seems unlikely that males could differ consistently in their ability

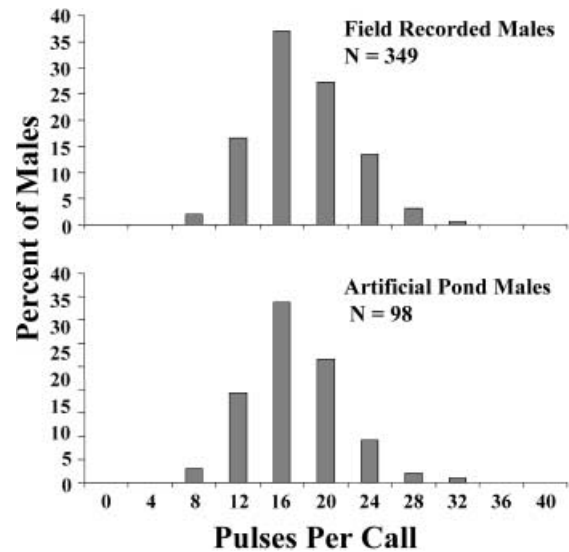


Fig. 7 Distributions of mean pulses per call for males recorded at the field site and in the artificial pond. The distributions did not differ significantly ($P=0.06$, Wilcoxon two-sample test)

to reduce acoustic interference in a chorus (Schwartz 2001; see below), our analyses also demonstrate that less acoustic interference with one's neighbors can act in a male's favor. In fact, the robust female preference for longer calls observed in laboratory two-stimulus tests was evident when we calculated mean pulses per call free from overlap. Nevertheless, the bulk of the variation in male mating success in the artificial pond was not explained by any of the components of male vocal behavior that we examined. Thus, while our data confirm that call properties can affect male mating success through female choice in the directions predicted by two-stimulus choice tests, they demonstrate the great extent to which the expression of female preferences may be attenuated in the chorus environment. Our results with *H. versicolor* make this point more forcefully than did field data on *H. microcephala* (Schwartz 1994) and *Hyperolius marmoratus* (Passmore et al. 1992). The studies on these other species revealed a higher expectation of mating success for chorus males that ranked high rather than low in note or call rate, respectively. Our data from *H. versicolor* may explain the observations of Sullivan and Hinshaw (1992) who, with small sample sizes in 2 of 3 years of their study, found little evidence that call rate or call duration influenced male mating success in natural choruses on a seasonal or nightly time scale. As stressed by Bertram et al. (1996), the number of nights on which a male advertises in the chorus is probably the most important factor influencing his chance of pairing with a female (but see Fellers 1979b), largely superseding other aspects of calling behavior as an explanation for variation in mating success. Occupation of particular calling sites may also, perhaps by virtue of an effect on sound transmission, impact on a male's chances of attracting a mate (Fellers 1979b).

The explanation for the difference between results obtained on female discrimination of pulse number in two-stimulus choice tests and real or simulated choruses with more than two call sources is not entirely clear. The difference is probably not due to noise per se, because addition of chorus noise close to typical background levels at the Ashland pond resulted in only a small drop in discrimination ability by females in our arena-based tests. Rather, what may be more important is interference with calls of relatively close neighbors whose vocalizations are sufficiently loud to disrupt the temporal structure of a call as perceived by a listening female (Schwartz 1987; Schwartz and Gerhardt 1995). Another potentially important factor may be the increase in acoustic or spatial complexity of a natural chorus relative to that of a two-stimulus choice test.

A recent result from two-stimulus choice tests is particularly relevant to our data from the field-based choice experiment. Namely, the strength of female preferences based on duration is a non-linear function of stimulus duration (Gerhardt et al. 2000; also see Gerhardt et al. 1996). When this strength was measured as the difference in SPL required to abolish the preference for the longer call, females exhibited a stronger preference for longer calls when the stimulus alternatives were short (e.g., 8 vs 12 pulses) than when the alternatives were long (18 vs 27 pulses). Therefore, data from our field experiment and Gerhardt et al. (2000) indicate that there is particularly strong discrimination by females against those calls with pulse numbers near the left tail of the natural distribution of males for this call feature (see Fig. 7).

If call duration, per se, is only weakly reflected in the choices of females at a chorus, why do males respond to the calls of others by adding pulses to their own calls while simultaneously reducing call rate (see Wells and Taigen 1986 and Grafe 1997b for a discussion of energetic considerations)? Late at night or on evenings when relatively few males are calling, females may sometimes face a choice between only two males. In these situations, based on extensive data from two-stimulus choice tests with and without background noise, all else being equal, the male giving longer calls will probably have a clear advantage. This improvement in mating success along with the very weak one present in larger choruses may translate into a significant selective advantage for males able to produce the longest calls (but see Sullivan and Hinshaw 1992). Thus, in a more competitive environment, males alter their calling behavior in a way that renders them relatively more attractive than would otherwise be the case (Wells 1988). In addition, when increased numbers of males are calling, adding pulses to calls may be necessary because the presence of more callers increases the likelihood that portions of a male's calls will be overlapped. For example, in tests in which we manipulated chorus size, we found that within groups of just two males, on average 92% of a male's calling time (percentage of time producing sound) was free of overlap in the artificial pond. In choruses of four and

eight males, this figure dropped to 57% and 32%, respectively (Schwartz 2001; J.J. Schwartz, B.W. Buchanan, H.C. Gerhardt, unpublished data). Moreover, the ability of males to actively avoid acoustic interference (determined by generating null expectations of overlap) appeared to be absent in these larger groups. Broadcasting unobstructed trains of pulses may be critical if a male is to attract a mate. Females discriminated against overlapped calls (Schwartz 1987; Schwartz and Gerhardt 1995) in arena-based choice tests and showed a significant preference for those males that had more of their calling time free from overlap than other males in our artificial pond. An additional important factor is that, based on our field data, females may find calls with 6 or fewer unobstructed pulses especially unattractive relative to longer calls. Therefore, if a male failed to lengthen his calls and, for example, simply increased his call rate with increases in chorus size or activity, many of his calls might fall below this approximate threshold of acceptability (Fig. 8). In fact, without the sharply increasing or step-like pattern of attractiveness for call pulse number exhibited in choruses, there might be little advantage to lengthening calls (while maintaining a stable calling effort). This is because, assuming random timing of calls among males, the expected proportion of all pulses produced that would be overlapped for a male that partitioned his calling effort into a lower number of long calls would be the same as that of a male that produced a higher number of short calls.

Levels of acoustic interference among a subset of the males used in our choice experiments in the artificial pond are consistent with this hypothesis as to why males shift the two components of calling effort in the direction of increased call duration and lower call rate in dense choruses. The subset we examined were those males in the same seven-male or eight-male choruses who had equivalent (within about 10%; mean=4.6%) calling efforts but very different call durations (>25%; mean=57.1%). We found in 26 pairwise comparisons that the male producing longer calls had a greater total number of non-overlapped pulses in his calls beyond the 6-pulse putative threshold than the male producing shorter calls significantly more often (19:7, $P=0.029$, two-tailed binomial test). However, when we ignored the 6-pulse criterion and counted all clear pulses, there was no significant relationship between call duration and overlap (11:15, $P=0.56$).

The choice behavior of females in our simulated and male-populated choruses formed by more than two sound sources helps explain another curious aspect of male vocal behavior observed during our manipulations of chorus size in the artificial pond. Males tended to maintain their ranking relative to other chorus members with respect to call duration while adjusting call duration between treatments (Schwartz 2001; J.J. Schwartz, B.W. Buchanan, H.C. Gerhardt, unpublished data). In other words, a male giving the longest calls in a high-density chorus typically continued to produce the longest calls after chorus size was reduced, although these calls were

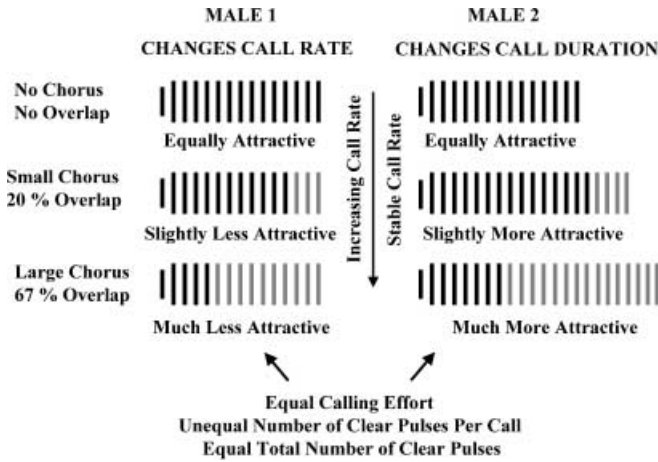


Fig. 8 Graphical representation of the increase in acoustic interference and its consequences for attracting mates as discussed in the text. Imagine two hypothetical males, with identical calling efforts, who respond to increases in chorus size by increasing calling effort; however, male 1 accomplishes this by elevating call rate while male 2 increases call duration. When each male is isolated from other males, or perhaps calling near just one other male, there is no call overlap and females can clearly discern species-specific temporal information transmitted in the pulses and their delivery pattern within the 15-pulse calls. In this situation, the two males are equally attractive. Now imagine in a small chorus that on average 20% of the pulse pattern in the calls of each male is obscured (*stippled pulses*) from females through call overlap with neighbors (for simplicity this is illustrated at the end of each call). In this case, male 1 broadcasts on average 12 clear pulses per call in his 15-pulse calls. Male 2, who has lengthened his calls, broadcasts 16 clear pulses, and females find him slightly more attractive than male 1. In the large chorus, there is considerably more overlap of calls and, on average, only 5 pulses within the calls of male 1 are not obscured. If there is a step-like or steep increase in call attractiveness as call duration increases beyond 6 pulses, females will find male 2 much more attractive than male 1. Note that the specific values chosen for this example are arbitrary and that, in reality, males increase call duration while maintaining about the same calling effort in response to increases in chorus size. However, the scenario described does not change if calling efforts are static for both males. Also note that the general scenario is valid if there is a gradual increase in advantage to males who produce more pulses per call that are not overlapped relative to other males. A step-like preference function only strengthens the advantage to such males. The precise location of the "step" is also not critical (e.g., Fig. 4a shows another abrupt increase in odds of success between 12 and 15 pulses)

shorter than those he had given initially. Males that were low in rank in a high-density chorus were also low in rank in choruses of lower density even though they were clearly capable of producing calls of sufficient duration to achieve a much higher rank in a low-density chorus! Our data on choice behavior suggest that, unless a male gives unusually short calls, his chances of attracting a female will not be seriously compromised by such a downward shift in duration accompanying moderate drops in chorus size, provided all remaining males behave in a similar fashion.

Alder and Rose's (1998) recent investigation of the temporal integration properties of neurons in the auditory midbrain of frogs bear upon our present and previous

findings on female selectivity based on pulse number. In electrophysiological recordings from the torus semicircularis of *Rana pipiens* and *Hyla regilla*, they found a class of pulse rate-sensitive neurons that required a minimum number of stimulus pulses (median=8.5, $n=18$) to fire. Conceivably, similar neurons might evaluate fine temporal features like pulse shape. Sound processing by these units could serve to improve the odds that certain neural and behavioral responses (e.g., phonotaxis) follow only those acoustic stimuli with amplitude modulation rates (as emphasized by Alder and Rose) or pulse shapes of conspecific calls. Thus, through increased sampling of fine temporal information in a sound, the probability of erroneous responses by females (e.g., approaching a heterospecific male or some other inappropriate sound source) should be reduced. While the fundamental role of this kind of neural processing is pattern recognition, a behavioral byproduct of the mechanism could be a female bias for long calls. Such a bias could set the stage for the evolution of more extreme preferences for signal duration if females received additional benefits from such behavior (*sensu* Welch et al. 1998).

If units like those described by Alder and Rose (1998) are present in the central nervous system of female *H. versicolor* and help mediate choice behavior, one would predict a non-linear functional relationship between female preference strength and the number of pulses in calls. The function would be characterized by both an initial abrupt rise in relative attractiveness as short calls increase in pulse number followed by a decline in relative preference strength differences as call alternatives increase in length. Such a neural mechanism would also help explain the preference for long over short calls when such stimuli are presented at equal calling effort. Additional experiments are planned or in progress to evaluate the relevance of Alder and Rose's findings to the behavior of *H. versicolor* females.

Acknowledgements We thank Benny Abraham, Dallas Nivens, Kathleen Sadlo, Michelle Grigsby, Christian Dornhoefer, Kerry Stevison, and Candy Corrigan for help with experiments. Leslie Schwartz helped construct cages and Johannes Schul provided technical assistance. Financial support was provided by grants from the National Science Foundation to J.J. Schwartz and H.C. Gerhardt. Use of experimental animals was consistent with US laws and also conformed to protocols approved by the Animal Care and Use Committee of the University of Missouri.

References

- Alder TB, Rose GJ (1998) Long-term temporal integration in the anuran auditory system. *Nat Neurosci* 1:519–523
- Bertram S, Berrill M, Nol E (1996) Male mating success and variation in chorus attendance within and among breeding seasons in the gray treefrog. *Copeia* 1996:729–794
- Cox DR, Snell EJ (1989) The analysis of binary data, 2nd edn. Chapman & Hall, London
- Dyson ML, Henzi SP, Passmore NI (1994) The effect of changes in the relative timing of signals during female phonotaxis in the reed frog, *Hyperolius marmoratus*. *Anim Behav* 48:679–685
- Fellers GM (1979a) Aggression, territoriality, and mating behavior in North American treefrogs. *Anim Behav* 27:107–119

- Fellers GM (1979b) Mate selection in the gray treefrog, *Hyla versicolor*. *Copeia* 1979:286–290
- Forester DC, Lykens DV, Harrison WK (1989) The significance of persistent vocalization by the spring peeper, *Pseudacris crucifer* (Anura: Hylidae). *Behaviour* 108:197–208
- Gayou DC (1980) The effect of temperature on the mating call of *Hyla versicolor*. MSc thesis, The University of Missouri, Columbia
- Gerhardt HC (1978) Temperature coupling in the vocal communication system of the gray treefrog *Hyla versicolor*. *Science* 199:992–994
- Gerhardt HC (1982) Sound pattern recognition in some North American treefrogs (Anura: Hylidae): implications for mate choice. *Am Zool* 22:581–595
- Gerhardt HC (1994) Reproductive character displacement of female mate choice in the gray treefrog *Hyla chrysoscelis*. *Anim Behav* 47:959–969
- Gerhardt HC (1995) Phonotaxis in female frogs and toads: execution and design of experiments. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) *Methods in comparative psychoacoustics*. Birkhäuser, Basel, pp 209–220
- Gerhardt HC, Klump GM (1988) Masking of acoustic signals by the chorus background noise in the green treefrog: a limitation on mate choice. *Anim Behav* 36:1247–1249
- Gerhardt HC, Rheinlaender J (1982) Localization of an elevated sound source by the green treefrog. *Science* 217:663–664
- Gerhardt HC, Watson GF (1995) Within-male variability in call properties and female preference in the grey treefrog. *Anim Behav* 50:1187–1191
- Gerhardt HC, Daniel RE, Perrill SA, Schramm S (1987) Mating behaviour and male mating success in the green treefrog. *Anim Behav* 35:1490–1503
- Gerhardt HC, Dyson ML, Tanner SD (1996) Dynamic properties of the advertisement calls of gray tree frogs: patterns of variability and female choice. *Behav Ecol* 7:7–18
- Gerhardt HC, Tanner SD, Corrigan CM, Walton HC (2000) Female preferences based on call duration in the gray treefrog (*Hyla versicolor*): influence of relative and absolute duration on preference strength. *Behav Ecol* 11:663–669
- Grafe TU (1997a) Costs and benefits of mate choice in the lek-breeding reed frog, *Hyperolius marmoratus*. *Anim Behav* 53:1103–1117
- Grafe TU (1997b) Use of metabolic substrates in the gray treefrog *Hyla versicolor*: implications for calling behavior. *Copeia* 1997:356–362
- Klump GM, Gerhardt HC (1987) Use of non-arbitrary acoustic criteria in mate choice by female gray treefrogs. *Nature* 326:286–288
- Klump GM, Gerhardt HC (1992) Mechanisms and function of call-timing in male-male interactions in frogs. In: McGregor PK (ed) *Playback and studies of animal communication*. Plenum, New York, pp 153–174
- Marquez R, Bosch J (1997) Female preference in complex acoustical environments in the midwife toads *Alytes obstetricans* and *Alytes cisternasii*. *Behav Ecol* 8:588–594
- Morris MR (1989) Female choice of large males in the treefrog *Hyla chrysoscelis*: the importance of identifying the scale of choice. *Behav Ecol Sociobiol* 25:275–281
- Murphy CG, Gerhardt HC (1996) Evaluating the design of mate-choice experiments: the effect of amplexus on mate choice by female barking treefrogs, *Hyla gratiosa*. *Anim Behav* 51:881–890
- Nagelkerke NJD (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78:691–692
- Narins PM (1992) Biological constraints on anuran acoustic communication: auditory capabilities of naturally behaving animals. In: Webster DB, Fay RR, Popper AN (eds) *The evolutionary biology of hearing*. Springer, Berlin Heidelberg New York, pp 439–454
- Narins PM, Zelick R (1988) The effects of noise on auditory processing and behavior in amphibians. In: Fritsch B, Wilczynski W, Ryan MJ, Hetherington T, Walkowiak W (eds) *The evolution of the amphibian auditory system*. Wiley, New York, pp 511–536
- Passmore NI, Capranica RR, Telford SR, Bishop PJ (1984) Phonotaxis in the painted reed frog (*Hyperolius marmoratus*). *J Comp Physiol A* 154:189–187
- Passmore NI, Bishop PJ, Caithness N (1992) Calling behavior influences mating success in male painted reed frogs, *Hyperolius marmoratus*. *Ethology* 92:227–241
- Runkle LS, Wells KD, Robb CC, Lance SL (1994) Individual, nightly, and seasonal variation in calling behavior of the gray treefrog, *Hyla versicolor*: implications for energy expenditure. *Behav Ecol* 5:318–325
- Ryan MJ (1985) *The tungara frog: a study in sexual selection and communication*. University of Chicago Press, Chicago
- Schwartz JJ (1987) The function of call alternation in anuran amphibians: a test of three hypotheses. *Evolution* 41:461–471
- Schwartz JJ (1993) Male calling behavior, female discrimination and acoustic interference in the Neotropical treefrog *Hyla microcephala* under realistic acoustic conditions. *Behav Ecol Sociobiol* 32:401–414
- Schwartz JJ (1994) Male advertisement and female choice in frogs: new findings and recent approaches to the study of communication in a dynamic acoustic environment. *Am Zool* 34:616–624
- Schwartz JJ (2001) Call monitoring and interactive playback systems in the study of acoustic interactions among male anurans. In: Ryan MJ (ed) *Recent advances in anuran communication*. Smithsonian Institution Press, Washington, DC, pp 183–204
- Schwartz JJ, Gerhardt HC (1995) Interspecific interactions and species recognition. In: Heatwole H, Sullivan BK (eds) *Amphibian biology*, vol 2. Social behaviour. Surrey Beatty, Chipping Norton, Australia, pp 603–632
- Schwartz JJ, Gerhardt HC (1998) The neuroethology of frequency preferences in the spring peeper. *Anim Behav* 56:55–69
- Schwartz JJ, Wells KD (1983) The influence of background noise on the behavior of a Neotropical treefrog, *Hyla ebraccata*. *Herpetologica* 39:121–129
- Schwartz JJ, Ressel S, Bevier CR (1995) Carbohydrate and calling: depletion of muscle glycogen and the chorusing dynamics of the Neotropical treefrog *Hyla microcephala*. *Behav Ecol Sociobiol* 37:125–135
- Simmons AM, Schwartz JJ, Ferragamo M (1992) Auditory-nerve representation of a complex communication sound in background noise. *J Acoust Soc Am* 91:2831–2844
- Sullivan BK, Hinshaw SH (1992) Female choice and selection on male calling behaviour in the grey treefrog *Hyla versicolor*. *Anim Behav* 44:733–744
- Sullivan BK, Ryan MJ, Verrell PA (1995) Female choice and mating system structure. In: Heatwole H, Sullivan BK (eds) *Amphibian biology*, vol 2. Social behaviour. Surrey Beatty, Chipping Norton, Australia, pp 469–517
- Telford SR, Dyson ML, Passmore NI (1989) Mate choice occurs only in small choruses of painted reed frogs *Hyperolius marmoratus*. *Bioacoustics* 2:47–53
- Welch AM, Semlitsch RL, Gerhardt H (1998) Call duration as a reliable indicator of “good genes” in the gray treefrog. *Science* 280:1928–1930
- Wells KD (1988) The effects of social interactions on anuran vocal behavior. In: Fritsch B, Wilczynski W, Ryan MJ, Hetherington T, Walkowiak W (eds) *The evolution of the amphibian auditory system*. Wiley, New York, pp 433–454
- Wells KD, Taigen TL (1986) The effect of social interactions on calling energetics in the gray treefrog (*Hyla versicolor*) *Behav Ecol Sociobiol* 19:9–18
- Wollerman L (1995) *Acoustic communication and acoustic interference in a neotropical treefrog*. PhD thesis, University of North Carolina, Chapel Hill
- Young ED, Costalupes JA, Gibson DJ (1983) Representation of acoustic stimuli in the presence of background sounds: adaptation in the auditory nerve and cochlear nucleus. In: Klinke R, Hartmann R (eds) *Hearing: physiological bases and psychophysics*. Springer, Berlin Heidelberg New York, pp 119–124