J Comp Physiol A (2005) 191: 1173–1183
DOI 10.1007/s00359-005-0041-1

ORIGINAL PAPER

H. E. Farris · A. Stanley Rand · Michael J. Ryan

The effects of time, space and spectrum on auditory grouping in túngara frogs

Received: 16 February 2005 / Revised: 27 June 2005 / Accepted: 5 July 2005 / Published online: 9 August 2005

Abstract Male túngara frogs (*Physalaemus pustulosus*) produce complex calls consisting of two components, a ~350 ms FM sweep called the "whine" followed by up to seven ~40 ms harmonic bursts called "chucks". In order to choose and locate a calling male, females attending to choruses must group call components into auditory streams to correctly assign calls to their sources. Previously we showed that spatial cues play a limited role in grouping: calls with normal spectra and temporal structure are grouped over wide angular separations (≤ 135°). In this study we again use phonotaxis to first test whether an alternative cue, the sequence of call components, plays a role in auditory grouping and second, whether grouping is mediated by peripheral or central mechanisms. We found that while grouping is not limited to the natural call sequence, it does vary with the relative onset times of the two calls. To test whether overlapping stimulation in the periphery is required for grouping, the whine and chuck were filtered to restrict their spectra to the sensitivity ranges of the amphibian and basilar papillae, respectively. For these dichotic-like stimuli, grouping still occurred (albeit only to 45° separation), suggesting that stream formation is mediated by central mechanisms.

Keywords Cocktail party effect · Auditory stream · Auditory scene analysis · Complex call · Mate choice · Chorus · Phonotaxis · Amphibian · Frog · Túngara · *Physalaemus pustulosus*

H. E. Farris (✉) · M. J. Ryan
Section of Integrative Biology C0930,
University of Texas, 1 University Station,
Austin, TX, USA
E-mail: Hfarri@lsuhsc.edu

H. E. Farris
Center for Neuroscience, LSUHSC,
2020 Gravier St., New Orleans, LA 70112, USA

A. S. Rand · M. J. Ryan
Smithsonian Tropical Research Institute,
Balboa, Panama

Introduction

A poorly studied aspect of sexual communication is a receiver's ability to assign signals to their correct source. Indeed, whereas most studies of sexual communication have focused on the limits of a receiver's ability to discriminate signals along some physical dimension (e.g., amplitude, size, concentration; Andersson 1994, Ryan and Keddy-Hector 1992), few studies have addressed which cues are used to assign them to their correct sender. Complicating the problem of source assignment is the fact that signals often do not vary along single dimensions and may consist of multiple components from multiple modalities (e.g., see Gibson and Bradbury 1985). Thus, in order to assess the quality of potential mates using such complex signals, receivers must first group the signal components into the appropriate complex (i.e., produced by a single male) and then correctly assign them to their source (i.e., which male; cf. humans: Mellinger and Mont-Reynaud 1996). This study examines which cues are used to solve these problems in female frogs faced with choruses of males producing complex calls.

Auditory scene analysis or the phenomena of source determination and group formation (Bregman 1990) were initially studied by Cherry (1953). His 'cocktail party problem', in which humans attend to one sound source amidst overlapping sounds from other sources and ambient noise, is similar to that faced by receivers across disparate taxa and ecologies. For example, male túngara frogs (*Physalaemus pustulosus*) produce complex calls containing two distinct components, a ~350 ms FM sweep called the whine followed by up to seven ~40–80 ms harmonic bursts called chucks. Commonly produced by males in multi-male choruses (Ryan 1980), all calls contain a whine, which is necessary and sufficient to elicit phonotaxis by females. The chuck alone rarely elicits phonotaxis, but when broadcast with the whine results in a call more attractive than the whine alone (Ryan and Rand 1990). This preference for the

complex call does not result from an increase in call energy (Ryan and Rand 1990) or duration (Wilczynski et al. 1999), suggesting the components are grouped as a stream. In a multi-male chorus consisting of up to 80 males with overlapping complex calls (Ryan 1985), what cues do females use for group formation (i.e., which whine goes with which chuck) and source determination (i.e., who produced the call)?

In humans, for whom there are the most data, auditory grouping and source determination result from several different cues. The first class are the so-called primitive cues: location (common interaural time and intensity differences), temporal cues (use of synchrony in different frequency bands such as harmonics to assign a common source), spectral cues (common changes in FM or harmonic relationship), and amplitude modulation (correlated envelope fluctuations). Called primitive because these cues may be associated with all sounds (i.e., even arbitrary), they require no a priori knowledge on the part of the receiver to form groups and segregate sources (Yost and Sheft 1993; Darwin and Carlyon 1995). A second class of cues includes those which are context dependent such as syntax. Like humans, female túngara frogs may use cues from either of these classes for group formation. With regard to primitive cues, calls from different males will vary in location, duration, onset and offset times, frequency, and amplitude. In addition, because males only produce chucks at the end or after a whine (Ryan and Drewes 1990), a syntax-like rule exists: chucks that precede or overlap earlier parts of whines should not be from the same male.

Using female phonotaxis, we recently tested the effect of spatially separated whines and chucks on group formation. Interestingly, although females rarely exhibit phonotaxis to a chuck alone, phonotaxis to a spatially separated chuck can be elicited when presented with a whine over a wide range of separation angles ($\leq 135°$; Farris et al. 2002). Because this conditional response to the chuck is not elicited over all separations ($> 135°$), the whine is not simply playing an alerting role (Richards 1981; Hebets 2005), suggesting the two components are assigned to a single source.

In this study we have continued to use the conditional attractiveness of spatially separated chucks (i.e., responses gated by a whine) as a bioassay of the cues and mechanisms used in the grouping of complex call components. With regard to cues, we examined whether the natural call sequence could explain grouping by varying the order of component presentation. Túngara frogs have two auditory end organs, the amphibian papilla (AP, low frequency) and the basilar papilla (BP, high frequency) which are predominantly responsible for the initial processing of the whine and chuck, respectively (Ryan et al. 1990). Thus, with regard to mechanisms for grouping, we exploited this auditory anatomy to test whether or not overlap in the periphery contributes to auditory grouping: by frequency filtering the two call components to restrict their spectral range to the sensitivity range of either the AP or BP we are able to create single ear (monotic) or double ear (dichotic) stimuli in the free-field, addressing the relative roles of central and peripheral processing in group formation.

Materials and methods

Subjects

Females were collected in amplexus in Gamboa, Panama approximately 3 h after sunset in August, 2002 and July, 2003. Because all animals were returned to the field following the trials, animals were marked (toe clipped) to prevent retesting on subsequent nights. All procedures with animal subjects were performed in accordance with the guidelines established by the Smithsonian Tropical Research Institute and the University of Texas at Austin Animal Care and Use Committee.

Experimental procedure

In order to assess grouping (or conditional responses to the chuck), phonotaxis to either the whine or chuck must not be ambiguous: the call components must be spatially separated so that differences in phonotaxis can be measured. Thus, the experimental procedure is the same as that used by Farris et al. (2002). Phonotactic responses were measured to whines and chucks presented alone or in combination at various spatial separations (0°, 45°, 90°, 135° and 180°). For each trial, gravid female *P. pustulosus* were placed beneath a plastic cone (10 cm diameter) at the center of a circular array of five speakers on the floor of the sound chamber. To ensure that movement by females was phonotaxis and not aversion from being handled, the females remained under the cone during the first 3 min of stimulus broadcast. Subsequently, the cone was removed allowing a subject to move freely. Using an infrared camera, we (minimum of three observers) recorded the position (in 5° intervals) at which females exited the full perimeter. All trials were videotaped to confirm measurements. A positive response was recorded only if the female crossed the perimeter of the array within 15 min. ''No-choice'' trials were scored when females (1) failed to leave the 10 cm center circle in 5 min, (2) remained stationary for 2 min, or (3) remained within the perimeter for 15 min. To ensure that ''no-choice'' scores were due to the stimuli and not a lack of female motivation, females exhibiting consecutive ''no-choice'' responses to stimuli that included a whine were not tested further and excluded from the analysis. Females were tested only once per stimulus. Subjects controlled their orientation with respect to the speakers at all times during the experiments (i.e., in or out of the cone). All trials were performed within 12 h of sunset under infrared illumination only. Ambient temperature for all experiments was ~27°C. The peak amplitude of all chucks was 6 dB re. whine amplitude (90 dB SPL; amplitude of call at 50 cm, Ryan

1985). These amplitudes are within the natural range produced by males. For each experiment, the order of the experimental stimuli was presented randomly. All females were presented with the normal whine and chuck alone as control stimuli.

Stimuli

The standard call (i.e., prior to variation in temporal sequence or spectrum; see below) was an average call from one of the sites where we collected females (Fig. 1). The average was calculated from 14 acoustic variables from 250 complex calls of 50 males (5 calls per male). Using multiple dimensional scaling to reduce the multivariate data to a two-dimensional map, we chose the call at the center of the distribution as representing a 'typical' or 'average' call for the population (see Ryan and Rand 2003). For consistency, this is the same call used by Farris et al. (2002). Call period for all stimuli was 2 s (Rand and Ryan 1981).

Stimuli were generated using Signal 16 bit, digital-to-analog converters and software (50 μs sample period). Stimuli were amplified using a Pioneer A-105 amplifier and broadcast from either Radioshack 4 in. broadband speakers (Cat. No. 40-1040) or ADS (L200C) speakers positioned along the perimeter of a 75 cm radius arc inside an Acoustic Systems (Austin, TX, USA) sound chamber (2.75×1.83 m) lined with additional anechoic foam along the bottom 0.6 m of each wall (Sonex, 1.5 in.; NRC 0.8). Prior to each night's trials, the peak amplitude of the whine and chuck stimuli were calibrated using a 500 Hz continuous tone, a GenRad 1982 sound level meter (Fast, linear weighting) and 0.5 in. microphone placed 3 cm above the floor of the arena at the center of the arc (grid on, 90° angle of incidence). All sound pressure levels (dB SPL) are referenced to 20 μPa. To control for potential position effects introduced by the chamber and to reduce any phonotactic bias due to potential speaker variability, speakers were randomly switched between positions along the array nightly (i.e., after 2–3 frogs) and stimulus orientation was varied for each broadcast condition in every trial.

Experiment 1: effects of whine-chuck sequence

Does call timing affect the spatial limitations of grouping?

When produced by a single male, the temporal sequence of call production consists of a whine followed by 0–7 chucks. In light of the fact that whine-chuck spatial separation does not appear to play a role in group formation (Farris et al. 2002), an alternative rule for grouping the components from a single male could be, "assign all chucks following a whine to a single male". By varying the onset times of the two call components to create calls that are not produced by single males, we tested whether the natural sequence was required for grouping spatially separated calls. Three sequences were

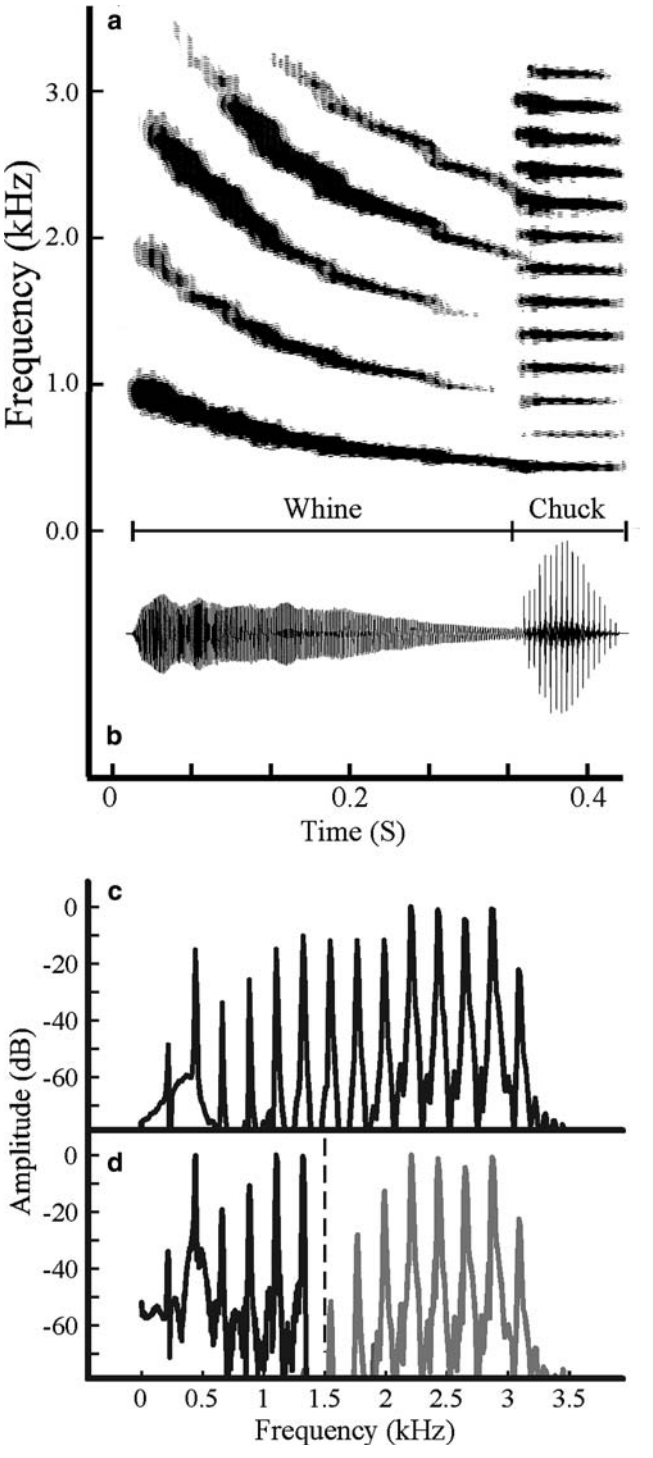

Fig. 1 **a**/**b** spectrogram and oscillogram of the complex call with natural sequence and spectrum. **c**/**d** Amplitude spectra of the normal and filtered chucks. The *black* and *grey plots* are the spectra of the low-pass and high-pass chucks, respectively. Filtering cutoff frequency was 1.5 kHz

tested covering different positions of temporal overlap with the whine (Figs. 2, 3, 4, 5). Chuck onset times (re. whine onset) were −80 ms (chuck precedes whine without overlap), 0 ms (simultaneous onset, overlapping beginning of whine) and 50 ms (overlap of a portion of

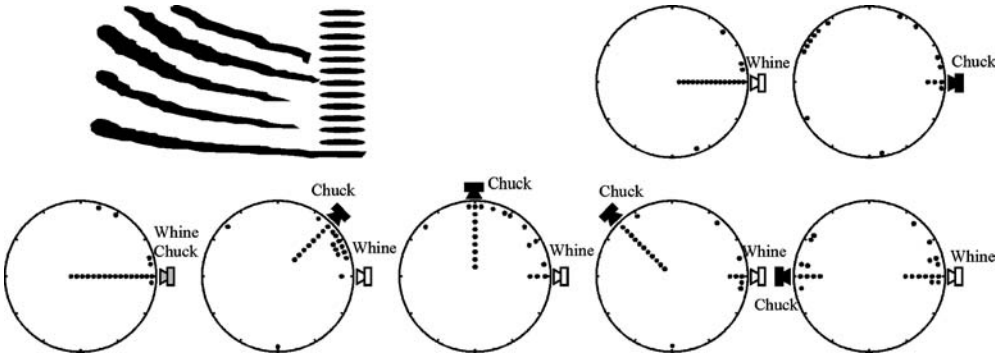

Fig. 2 Data previously reported by Farris et al. (2002). Each point represents the exit angle of a female presented with the complex call components (natural spectra) in their natural sequence, at various separation angles. Control distributions are shown for the components presented alone. The conditional response to the chuck is significant to 135°. *Inset* shows a spectrogram (without intensity information) of the stimulus. Circular arena has 1.5 m diameter

the whine critical for eliciting phonotaxis, Wilczynski et al. 1995). The relative chuck onset time for the natural sequence (i.e., tested by Farris et al. 2002) is 329 ms, the duration of the whine. Whereas females collected in 2002 were tested with the 0 and 50 ms sequences, females in 2003 where tested with the −80 ms sequence. Each experiment was conducted until a minimum of 20 females completed the entire stimulus set (e.g., all spatial separations for each temporal sequence) including broadcasts of the individual components: chuck alone and whine alone.

Experiment 2: monotic vs. dichotic stimuli

Does grouping occur for signals processed by different end organs?

Theories for the underlying mechanisms of stream segregation in humans using monaural cues have primarily focused on mechanisms at the periphery, where grouping is accomplished using the low level processing in single channels in the cochlea (van Noorden 1977; Beauvois and Meddis 1996; Rose and Moore 2000). In contrast, peripheral processing of the whine and chuck by túngara frogs is accomplished largely by separate end organs (Ryan et al. 1990). Whereas the AP is primarily tuned to lower frequencies (< 1.2 kHz), the BP exhibits best sensitivity near 2.1 kHz. With respect to phonotaxis, the whine is primarily processed by the AP, as its low frequency fundamental is necessary and sufficient to elicit attraction (and the higher frequency harmonics do not have an influence; Wilczynski et al. 1995). The chuck, however, appears to be processed by the BP. Previous studies have shown that the enhanced attractiveness of the whine produced by the addition of a chuck can be accomplished using the energy remaining in high-pass filtered chucks which stimulates the BP only, but not with two call components low-pass filtered for the AP (Wilczynski et al. 1995). This means that the peripheral mechanisms implicated in the grouping of auditory streams in humans may not be responsible for the grouping of the components of the complex call in *P. pustulosus*. To test this, the standard call was digitally filtered to produce either a low-pass call that is monotic with both the whine and chuck stimulating the AP or a dichotic call with the whine stimulating the AP and a high-pass chuck stimulating the BP (Fig. 1). Filtering removed all of the harmonics from the whine (leaving the fundamental). The chuck was low and high-pass filtered at 1.5 kHz (> 98 dB/octave) to produce the AP and BP stimuli, respectively. Because the grouping response varies with the relative amplitude of the two components (Farris et al. 2002), the relative amplitude of the filtered components was adjusted to the pre-filtered level (i.e., chuck 6 dB re. whine). This control was particularly important given that the low and high-pass filtered chucks had different amplitudes. Thus, these data address whether auditory grouping can be accom-

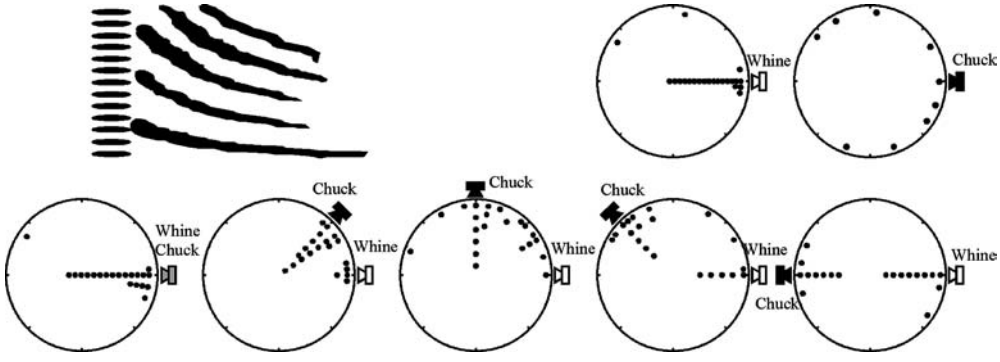

Fig. 3 Exit angles for females presented with calls (natural spectra) in reversed sequence (− 80 ms). Control distributions are shown for the components presented alone. The conditional response to the chuck is significant to 135° (Table 5). *Inset* shows a spectrogram (without intensity information) of the stimulus

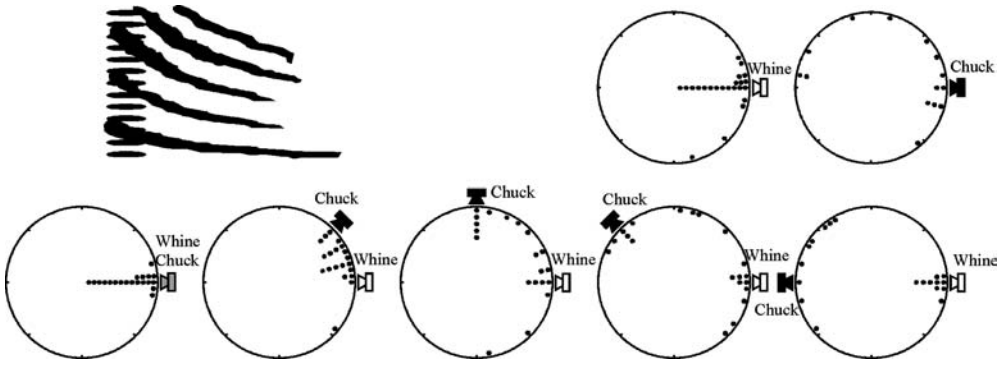

Fig. 4 Exit angles for females presented with calls (natural spectra) with simultaneous onset (0 ms). Control distributions are shown for the components presented alone. The conditional response to the chuck is significant to 90° (Table 5). *Inset* shows a spectrogram (without intensity information) of the stimulus

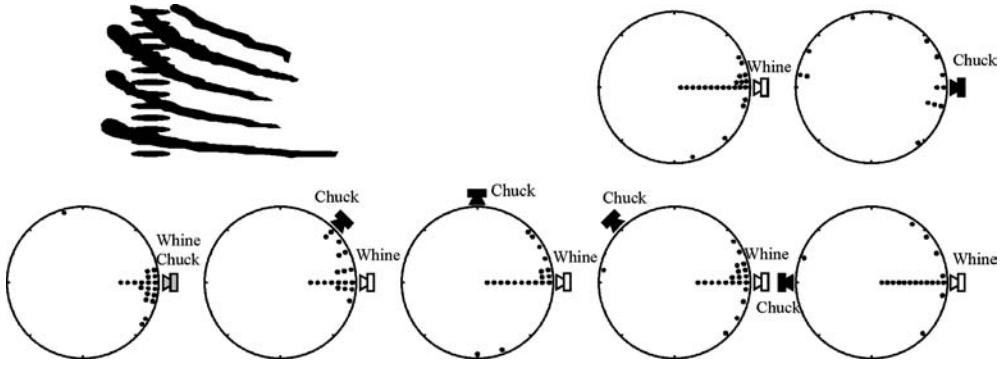

Fig. 5 Exit angles for females presented with calls (natural spectra) in which the onset of the chuck is 50 ms that of the whine. Control distributions are shown for the components presented alone. The conditional response was not measured for this sequence at any separation (Table 5). *Inset* shows a spectrogram (without intensity information) of the stimulus

plished in the absence of peripheral sensory overlap, a condition more often associated with stimuli of different modalities (i.e., visual and auditory; Narins et al. 2003; McGurk and McDonald 1976). Females collected in 2002 and 2003 were tested with the low-whine hi-chuck and the low-whine low-chuck stimuli, respectively. Each experiment was conducted until a minimum of 20 females completed the entire stimulus set (e.g., all spatial separations for each type of filtered call) including broadcasts of the individual components by themselves: the whine, chuck, low whine, low chuck and hi chuck.

Analysis

Using the female exit angles from the circular arena, we employed both categorical and circular analyses to assess which stimuli elicited a conditional response to the chuck, an indication of perceptual group formation. For the categorical analysis, exit angles were grouped into one of two categories. The first category included those responses in which females exited the perimeter and made contact with the speaker broadcasting the chuck (a 20° arc; 13 cm). The second category consisted of all other exit angles as well as those trials exhibiting "no choice" responses. No choice responses, especially to the chuck alone are biologically relevant but cannot otherwise be included in the circular analyses below. A Fisher exact test was used to compare the probability of phonotaxis to chucks broadcast with the whine to that for the simple broadcast condition of the chuck alone, which served as the null hypothesis.

For the circular analyses, the effect of chuck position on phonotactic direction was analyzed using three different tests. First, a Rayleigh test for circular uniformity tested whether or not the exit angles for the chuck-alone stimulus were randomly distributed. Second, a *V*-test was used to determine whether responses were localized at a particular exit angle (Zar 1999). The *V*-test, however, is most powerful when the distribution of exit angles is unimodal. Because phonotaxis is also directed at the whine in some trials, the *V*-test has poor resolution for smaller call separations. For example, when the whine and chuck are separated by 45°, phonotaxis to the nearby whine (broadcast from 0°) can erroneously lead to the conclusion that all exit angles are indeed located at 45° (the chuck position): due to the experimental design, the criteria for significance is not uniform for all separations. Thus, to complement the *V*-test, differences in mean exit angles between various broadcast conditions were analyzed using a third method, the Watson-Williams test (Zar 1999) with alpha correction (i.e., Bonferroni) for multiple comparisons. This test reveals whether the distribution of exit angles was affected by the position of the chuck even though they may not be localized there. Whereas the *V*-test determines if phonotaxis is localized at a certain separation angle, the Watson-Williams test assesses whether the distribution is different from that for call components with no spatial separation (0°). Unlike the categorical analysis, these

analyses were restricted to trials in which the females exited the perimeter, excluding "no choice" responses.

Results

Experiment 1

Few females responded during the chuck-only stimulus and their exit angles were randomly distributed (2002: $n = 14$, $P > 0.05$; 2003: $n = 9$, $P > 0.5$). When the chuck was presented with the whine, however, conditional phonotaxis was elicited to spatially separated chucks in certain of the altered sequences (Tables. 1, 4, 5). For example, when the temporal order of the two components is reversed so that the chuck precedes the whine (−80 ms relative chuck onset), phonotaxis to the chuck is exhibited up to 135° whine-chuck separation, identical to the natural sequence (Figs. 2, 3; Farris et al. 2002). In contrast, however, when chuck onset occurs during the whine, the effect of the chuck on phonotaxis is comparatively limited and depends on the time of overlap. Similar to the natural and reversed sequences, when the two components have simultaneous onsets (0 ms relative chuck onset), phonotaxis is directed toward the chuck up to 90° (Tables 4 and 5; Fig. 4). Most striking was the relative failure of conditional phonotaxis to the chuck for the 50 ms relative onset time, showing only an effect at 45° when analyzed using the most liberal test (*V*-test, Table 4; Fig. 5).

Table 1 Categorical analysis of chuck attractiveness in simple vs. complex broadcasts

Angle of Chuck (re. whine)	# to chuck	# Non-chuck responses	Total	Fisher Exact *P*
Chuck only Null *Ho*: −80 ms	1	21	22	–
0°	19	4	23	< 0.0005
45°	13	8	21	< 0.0005
90°	10	12	22	0.0019
135°	9	13	22	0.0047
180°	7	12	19	0.0124
Chuck only Null *Ho*: 0 ms	3	17	20	–
0°	19	1	20	< 0.0005
45°	4	16	20	0.296
90°	6	14	20	0.161
135°	5	17	22	0.254
180°	1	20	21	0.236
50 ms				
0°	15	5	20	< 0.0005
45°	3	17	20	0.338
90°	0	22	22	0.099
135°	0	22	22	0.099
180°	0	20	20	0.115

Table shows an increased response to the chuck stimulus when presented with the whine. Columns are the angle of the chuck stimulus (re. the whine position) for three different relative onset times, number of individuals attracted to the chuck, number of non-chuck responses, total responses and *P* value for the Fisher exact comparison using the responses to the chuck-only stimulus as the null hypothesis

Table 2 Categorical analysis of chuck attractiveness in simple vs. complex broadcasts

Angle of Chuck (re. whine)	# to Chuck	# Non-Chuck responses	Total	Fisher Exact *P*
Lo Chuck only Null *Ho*: Lo Wh Lo Ch	1	20	21	–
0°	12	11	23	< 0.0006
45°	3	19	22	0.321
90°	1	21	22	0.727
135°	3	21	24	0.356
180°	0	21	21	0.5
Hi Chuck only Null *Ho*: Lo Wh Hi Ch	0	20	20	–
0°	18	4	22	< 0.0005
45°	6	17	23	0.016
90°	1	23	24	> 0.5
135°	2	21	23	0.280
180°	1	22	23	> 0.5

Table shows an increased response to the chuck stimulus when presented with the whine. Columns are the angle of the chuck stimulus (re. the whine position) for AP only and AP-BP stimuli, number of individuals attracted to the chuck, number of non-chuck responses, total responses and *P* value for the Fisher exact comparison using the responses to the chuck-only stimulus as the null hypothesis

Table 3 Relationship between call spectral content, position and the direction of phonotaxis

Angle of Chuck (re. whine)	*N*	Mean angle (°)	Vector length (*r*)	*P*
Lo Wh Lo C				
0°	21	4.11	0.751	< 0.0005
45°	19	16.78	0.944	< 0.025
90°	21	24.79	0.728	> 0.25
135°	22	21.18	0.666	> 0.25
180°	20	355.9	0.781	> 0.25
Lo Wh Hi C				
0°	21	1.09	0.984	< 0.0005
45°	22	21.31	0.941	< 0.0005
90°	24	9.49	0.824	> 0.1
135°	22	11.44	0.729	> 0.25
180°	23	5.29	0.858	> 0.25

Columns are the angle of the chuck stimulus (re. the whine position) for AP only and AP-BP stimuli, number of individuals exhibiting positive responses, mean exit angle, length of the mean vector (varies from 0 to 1 and is inversely correlated to the variance in exit angles), and P value for a V-test for circular uniformity at the position of the chuck speaker

Experiment 2

Low-pass filtering the whine did not affect phonotaxis, as exit angles are still localized at the speaker (Fig. 6; $r = 0.87$, $P < 0.0005$). When presented in the absence of a whine, females responded to the filtered chucks as they did in the previous experiment to the unfiltered chucks, randomly (low: $r = 0.276$, $P > 0.5$; hi: $r = 0.243$, $P > 0.2$).

Table 4 Relationship between chuck onset time, position and the direction of phonotaxis

Angle of Chuck (re. whine)	*N*	Mean angle (°)	Vector length (*r*)	*P*
−80 ms onset				
0°	21	358.9	0.908	< 0.0005
45°	20	29.73	0.937	< 0.0005
90°	22	68.85	0.817	< 0.0005
135°	20	91.84	0.518	< 0.01
180°	19	13.5	0.054	> 0.25
0 ms onset				
0°	20	0.99	0.997	< 0.0005
45°	20	19.91	0.948	< 0.0005
90°	20	33.9	0.694	< 0.0005
135°	21	41.36	0.470	> 0.25
180°	20	35.16	0.235	> 0.25
50 ms onset				
0°	20	2.23	0.915	< 0.0005
45°	19	8.97	0.957	< 0.0005
90°	20	2.52	0.880	> 0.25
135°	22	2.22	0.872	> 0.25
180°	19	5.12	0.845	> 0.25

Columns are the angle of the chuck stimulus (re. the whine position) for three different relative onset times, number of individuals exhibiting positive responses, mean exit angle, length of the mean vector (varies from 0 to 1 and is inversely correlated to the variance in exit angles), and *P* value for a *V*-test for circular uniformity at the position of the chuck speaker

Table 5 Comparison of mean exit angles for spatially separated calls to that with 0° separation using a Watson-Williams test (Zar 1999) with alpha correction for multiple comparisons ($\alpha = 0.0125$)

Angle of Chuck (re. whine)	*N*	Mean angle (°)	*P*
−80 ms			
0°	21	358.9	–
45°	20	29.73	0.0001*
90°	22	68.85	< 0.0001*
135°	20	91.84	< 0.0001*
180°	19	13.5	> 0.5
0 ms			
0°	20	0.99	–
45°	20	19.91	< 0.005*
90°	20	33.9	0.009*
135°	21	41.36	0.028
180°	20	35.16	0.323
50 ms			
0°	20	2.23	–
45°	19	8.97	0.222
90°	20	2.52	> 0.5
135°	22	2.22	> 0.5
180°	19	5.12	> 0.5

Asterisks mark significant differences from the 0° condition

There was little evidence for auditory grouping using the low-pass filtered calls (AP only), as the conditional response to the chuck was only found using the most liberal test (*V*-test) for the 45° separation (Tables 2, 3, 6). For the dichotic stimulus (AP whine, BP chuck), however, all three analysis techniques revealed a significant response to the high-passed chuck at the 45° separation (Fig. 7). Albeit reduced relative to calls with

Table 6 Comparison of mean exit angles for spatially separated calls to that with 0° separation using a Watson-Williams test (Zar 1999) with alpha correction for multiple comparisons ($\alpha = 0.0125$)

Angle of Chuck (re. whine)	*N*	Mean angle (°)	*P*
Lo Wh Lo Ch			
0°	21	4.11	–
45°	19	16.78	0.488
90°	21	24.79	0.293
135°	22	21.18	> 0.5
180°	20	355.9	> 0.5
Lo Wh Hi Ch			
0°	21	1.09	–
45°	22	21.31	< 0.0005*
90°	24	9.49	> 0.5
135°	22	11.44	> 0.5
180°	23	5.29	0.767

Asterisks mark significant differences from the 0° condition

normal spectra, responses to the low whine-hi chuck stimulus suggest that grouping can be processed centrally and overlapping stimulation of the same peripheral frequency channels (i.e., critical bands) is not required. Indeed, when limited to such stimulation, as with the low whine-low chuck, the conditional response to the chuck was not measured.

Discussion

Because females exhibit both a conditional response to chucks (i.e., evidence for group formation) and perform phonotaxis to a particular source (i.e., evidence for source determination), the bioassay used here appears to satisfy both criteria for testing auditory scene analysis (see Bregman 1990; Yost and Sheft 1993). The demonstration of this phenomenon notwithstanding, we have used this experimental paradigm to examine which cues enable female frogs to perform this task. Initial work on this phenomenon in humans revealed that the spatial separation between sources plays a significant role in scene analysis (Cherry 1953; Kubovy et al. 1974). For túngara frogs, however, we found that for a single whine and chuck, spatial separation plays a surprisingly minor role in grouping the two call components (Farris et al. 2002). Thus, in this study we have followed those earlier results by exploring two other likely cues (temporal sequence and spectral overlap) leading to the grouping of call components, a requirement for scene analysis.

Call sequence

In various taxa, the sequence of complex call components has been shown to affect receiver behavior. For example, in a sequence of call components preceding components may alert receivers to subsequent ones. Although the alerting component may contain no information and elicit no response on its own, its inclusion in a particular sequence can affect the proba-

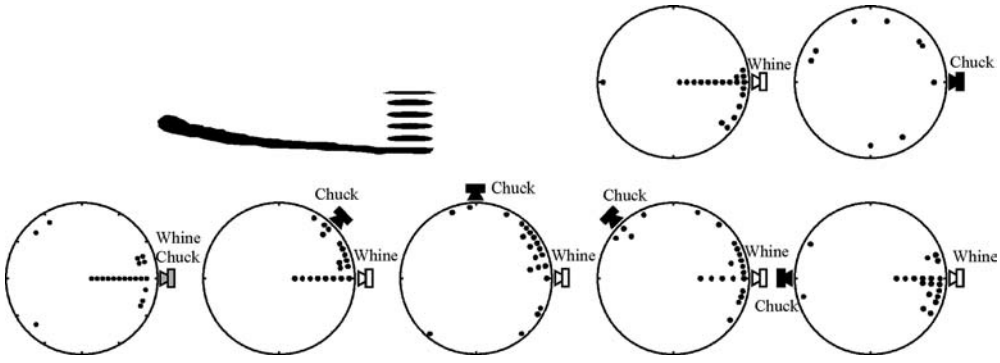

Fig. 6 Exit angles for females presented with monotic calls. The call components were filtered to the spectral range of the amphibian papilla: the whine consists of only the fundamental and the chuck has been low-pass filtered at 1.5 kHz (98 dB/octave). *Inset* shows a spectrogram (without intensity information) of the stimulus

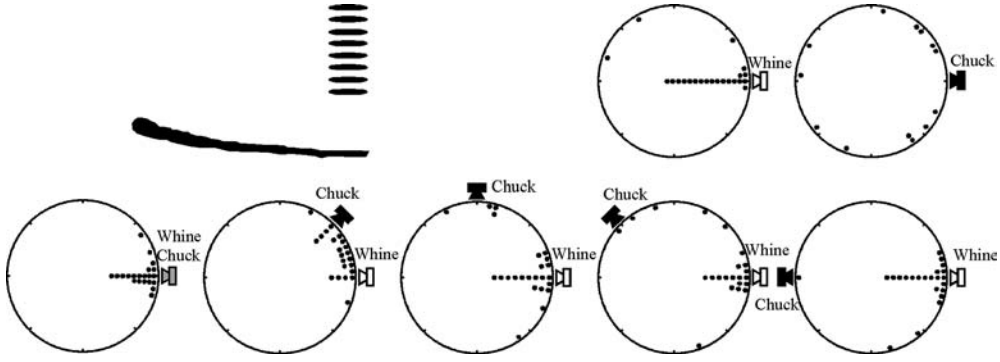

Fig. 7 Exit angles for females presented with dichotic calls. The call components were filtered to the spectral ranges of the amphibian and basilar papillae: the whine consists of only the fundamental and the chuck has been hi-pass filtered at 1.5 kHz (98 dB/octave). *Inset* shows a spectrogram (without intensity information) of the stimulus

bility of detecting following stimuli (Richards 1981; Hebets 2005). The alerting role notwithstanding, responses may vary with call sequence even when the probability of detection is not different for the various call components (cf. Narins and Capranica 1978). For chickadees (*Parus atricapillus*) and veeries (*Catharus fuscenscens*) for example, simple calls or complex calls in the natural sequence more quickly elicit responses than complex calls with the components reversed (Ratcliffe and Weisman 1986; MacNally et al. 1986). In túngara frogs, the effect of the whine-chuck sequence on female phonotaxis has primarily been tested in choice paradigms. Wilczynski et al. (1999) compared the relative attractiveness of a complex call broadcast in a variety of temporal sequences to either a whine alone or a whine-chuck in the natural sequence. Thus, those experiments asked whether the increased attractiveness bestowed on the whine by a chuck is robust to its relative temporal position. With few exceptions (see below), the relative temporal position of the chuck, including those sequences that are never produced by males in nature (Ryan and Drewes 1990), had little effect on the attractiveness of the complex (Wilczynski et al. 1999).

Although our study measured the probability of grouping and source determination (rather than discrimination and detection) our hypothesis is similar to Wilczynski et al. (1999) in that we examine whether responses deteriorate for certain call sequences which single males do not produce. Like those data from preference tests our data show that call sequence per se does not play a role in grouping the components: for the natural and reversed sequences, we found grouping to occur over the same spatial separation (135°; Tables 4, 5 and Farris et al. 2002). This response is only slightly diminished for overlapping calls with a simultaneous onset. Here, grouping is significant up to 90° and nearly so at 135° (Tables 4, 5).

Of particular interest is the sequence in which the chuck begins 50 ms after the start of the whine. In choice tests, responses to this sequence suggest that females do not attend to the chuck, as this portion of the whine is critical to the whine's recognition (Wilczynski et al. 1995, 1999). Interestingly, we also found a reduction in the conditional response to chucks overlapping this particular whine segment. When compared to the exit angles for 0° spatial separation, calls with this sequence do not elicit conditional responses to the chuck at any spatial separation and thus do not show evidence for grouping (Table 5). The similarity between these results and those from choice tests (Wilczynski et al. 1999) suggests a similar mechanism in both experimental paradigms (grouping and choice) in which attention to the chuck is gated by the presence of a particular segment of the whine. Due to the frequency modulation of the whine, however, it is not clear from these results whether gating (and thus grouping) depends on the temporal and/or spectral properties of this segment.

From a comparative point of view, temporal and spectral overlap are important for grouping in humans. Sounds with overlapping frequency bands and similar temporal modulation are more likely to be processed as originating from the same source (see Yost and Sheft

1993, for review). A cost to such processing is the potential for masking or suppression; the underlying mechanisms for such phenomena are well known in behavioral and physiological tests (Kiang 1965; Ehret et al. 1983; Farris and Hoy 2002). For example, masking is defined as an increase in detection threshold of one stimulus by the presence of another (see Delgutte 1996 for review). Known mechanisms for masking include excitatory effects, in which the masker causes a reduction in the signal-to-noise ratio in the neural code; neural adaptation, in which adaptation caused by the excitatory masker reduces the excitatory response to the signal; or biomechanical suppression or neural inhibition. The lack of response to the chuck for a particular whine-chuck sequence (overlap) raises the possibility that the whine is masking (or suppressing) the chuck and that grouping is influenced by such mechanisms in the periphery. Unfortunately these mechanisms do not appear consistent with the responses to the varying call sequences. For example, >90% of the chuck's power is above 1.5 kHz and processed by the BP. In contrast, only the harmonics of the whine, which do not influence phonotaxis, are processed there. Meaning that although the overall power ratio (chuck/overlapping whine segment) is −3.5 dB, the power ratio in the BP's frequency band is 5.3 dB. Thus, channels processing these upper frequencies are unlikely to receive excitatory or adaptive masking input by the frequencies of the whine. Furthermore, excitatory masking is also unlikely in the light of the positive grouping response to components with simultaneous onsets, as masking of the chuck would be expected here too. With respect to 2-tone suppression, high frequencies are commonly more effective in suppressing lower ones (Ehret et al. 1983). Given that detection of the higher frequency chuck rather than the whine appears suppressed at the 50 ms overlap segment, 2-tone suppression also seems unlikely. The results of these experiments with varying call sequences indirectly suggest that typical mechanisms in the periphery are not playing a role in grouping the two call components. Thus, experiment 2 was carried out to directly test whether peripheral overlap is required.

Dichotic vs. monotic grouping

Monotic and dichotic stimuli are useful in probing the location of the underlying mechanisms of certain auditory processes (Hicks and Bacon 1995). Indeed, if responses to dichotic stimuli (sound is different at each ear) are different from those in the monotic condition it suggests that some stimulus overlap in the periphery is required to maintain the same processing. Here we exploited the frequency sensitivity of the two auditory end organs in túngara frogs in order to present these two stimulus conditions in the free field. Thus, what is commonly accomplished in humans using headphones to separately stimulate the mechanically uncoupled left and right ears was done by either high-pass or low-pass filtering the two call components at 1.5 kHz so that they either stimulated the amphibian or the basilar papillae. In choice tests, the effect of filtering the calls depends on call amplitude: whereas the energy remaining in a high-pass filtered chuck is sufficient to increase the attractiveness of a whine (Wilczynski et al. 1995), a low-pass filtered chuck is effective only when its amplitude is corrected to its pre-filtered level (Ryan and Rand 1990). In the grouping paradigm used here, call amplitude did not vary and filtered calls were amplified to their pre-filtered levels. Yet, we measured the conditional response only to the high-pass chuck, suggesting grouping varies with spectral content and thus peripheral overlap. For monotic stimuli in which only the amphibian papillae were stimulated, no grouping was observed when analyzed using the most conservative test (Table 6). We feel that this inconsistency with data previously collected in the choice paradigm in which AP only stimuli were found to be as attractive as normal calls (Ryan and Rand 1990), is likely due to the limited size of separation angles tested; low-frequency, AP-only stimuli could still be grouped but not at 45°. In contrast, in the dichotic condition (no peripheral overlap) grouping was observed up to a 45° whine-chuck separation. Although this reduction in grouping from 135° (normal chuck) to 45° (high-pass chuck) suggests peripheral overlap is playing a role in grouping, it is not necessary and suggests that it may not be sufficient (failure of low-pass chuck). Thus, these latter results show that central mechanisms are capable of mediating auditory grouping of the whine and chuck.

In the chorus

At least two simple decision rules for grouping whines and chucks are based on their spatial and temporal separations. Smaller spatial separations between components should increase the likelihood of coming from the same male. In the temporal domain, because males exhibit a strict syntax rule in producing the two call components (chucks follow whines) grouping decisions based on component temporal overlap should also improve the rate of correct grouping. In a chorus, however, our data like those for treefrogs (Schwartz and Gerhardt 1995) would suggest extremely poor spatial and temporal acuity for grouping and source determination and predict numerous errors in phonotaxis: 135° resolution in a chorus should make almost every male a potential source for each call. We addressed simple decision rules, however, and our results point to the use of more complicated rules not tested in this experimental paradigm. For example, with respect to spatial cues, a possible decision rule could direct phonotaxis to the chuck closest to a whine, a possibility we could not test since we used only a single whine and a single chuck. Under such a rule, increased spatial and temporal acuity would only be expressed when multiple chucks are presented. In the temporal domain, aspects of grouping could be

affected by variance in the chuck–whine phases (cf. Bosch et al. 2002) as opposed to a particular phase (tested here): females would pay attention to temporal consistency between a whine and chuck. Similarly, in the amplitude domain, mechanisms of selective attention in túngara frog choruses (Greenfield and Rand 2000) could increase the probability of correct grouping by exploiting component amplitude to reduce the number of alternative calls detected by the female. Beyond these auditory cues, grouping may yet be improved using visual cues. Like numerous other anurans, the throat sac of male túngara frogs inflates during the production of each call. Rosenthal et al. (2004) showed that females prefer calls associated with video of male throat sac inflation raising the possibility that temporal coherence between this visual cue and the call increases the likelihood of grouping. Thus, the exclusion of spatial and temporal cues from decisions for grouping single whines and chucks does not yet allow us to understand grouping dynamics in a natural chorus.

In conclusion, this study shows that although the two call components may play different roles in directing phonotaxis when presented together, the whine for identity and the chuck for location (Farris et al. 2002), túngara frogs do not use the available cues in call sequence to group them into a stream. Furthermore, by using the monotic—dichotic stimulus presentations traditionally employed in psychoacoustics, we show that central mechanisms play a role in auditory grouping and source determination. Consequently, túngara frogs now provide the opportunity to examine central mechanisms for streaming independently from those in the periphery in freely behaving animals.

Acknowledgements We thank Rex Cocroft, Dennis McFadden, Carl Gerhardt, Walt Wilczynski and two anonymous reviewers for comments on the project and/or manuscript. Thanks to X. Bernal, F. V. Candioti, K. Lynch, K. Boul, J. Scarl, and S. Barrionuevo for help with data collection. This research was supported by NSF IBN 98 16564, and a Smithsonian Scholarly Studies Grant to MJR and ASR. HEF was supported by funds from the College of Natural Science, University of Texas at Austin. We thank the Smithsonian Tropical Research Institute for the their hospitality and logistical support. The experiments comply with the NIH guidelines for animal care (pub. #86-23).

References

Andersson MB (1994) Sexual selection. Princeton University Press, Princeton

Beauvois MW, Meddis R (1996) Computer simulation of auditory stream segregation in alternating-tone sequences. J Acoust Soc Am 99:2270–2280

Bosch J, Rand AS, Ryan MJ (2002) Response to variation in chuck frequency by male and female túngara frogs. Herpetologica 58:95–103

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge

Cherry EC (1953) Some experiments on the recognition of speech, with one and two ears. J Acoust Soc Am 25:975–979

Darwin CJ, Carlyon RP (1995) Auditory grouping. In: Moore BCJ (ed) Hearing. Academic Press, San Diego, pp 387–424

Delgutte B (1996) Physiological models for basic auditory percepts. In: Hawkins HL, McMullen TA, Popper AN, Fay RR (eds) Auditory computation. Springer, Berlin Heidelberg New York, pp 157–220

Ehret G, Moffat AJM, Capranica RR (1983) Two-tone suppression in auditory nerve fibers of the green treefrog (*Hyla cinerea*). J Acoust Soc Am 73:2093–2095

Farris HE, Hoy RR (2002) Two-tone suppression in the cricket *Eunemobius carolinus* (Gryllidae, Nemobiinae). J Acoust Soc Am 111:1475–1485

Farris HE, Rand AS, Ryan MJ (2002) The effects of spatially separated call components on phonotaxis in túngara frogs: Evidence for auditory grouping. Brain Behav Evol 60:181–188

Gibson RM, Bradbury JW (1985) Sexual selection in lekking sage grouse: phenotypic correlates of male mating success. Behav Ecol Sociobiol 18:117–123

Greenfield MD, Rand AS (2000) Frogs have rules: selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). Ethology 106:331–347

Hebets EA (2005) Attention-altering signal interactions in the multimodal courtship display of the wolf spider *Schizocosa uetzi*. Behav Ecol 16:75–82

Hicks ML, Bacon SP (1995) Some factors influencing comodulation masking release and across-channel masking. J Acoust Soc Am 98:2504–2514

Kiang NY-S (1965) Discharge patterns of single fibers in the cat's auditory nerve. Research monograph no. 35, The M. I. T. Press, Cambridge

Kubovy M, Cutting JE, McGuire RM (1974) Hearing with the third ear: dichotic perception of melody with monaural familiarity cues. Science 186:272–274

MacNally RC, Weary DM, Lemon RE, Lefebvre L (1986) Species recognition by song in the veery (*Catharus fuscescens*: Aves). Ethology 71:125–139

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

Mellinger DK, Mont-Reynaud BM (1996) Scene analysis. In: Hawkins HL, McMullen TA, Popper AN, Fay RR (eds) Auditory computation. Springer, Berlin Heidelberg New York, pp 271–331

Narins PM, Capranica RR (1978) Communicative significance of the two-note call of the treefrog *Eleutherodactylus coqui*. J Comp Physiol 127:1–9

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

van Noorden LPAS (1977) Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. J Acoust Soc Am 61:1041–1045

Rand AS, Ryan MJ (1981) The adaptive significance of a complex vocal repertoire in a neotropical frog. Ethology 57:209–214

Ratcliffe L, Weisman RG (1986) Song sequence discrimination in the black-capped chickadee (*Parus atricapillus*). J Comp Psychol 100:361–367

Richards DG (1981) Alerting and message components in songs of rufous-sided towhees. Behaviour 76:223–249

Rose MM, Moore BCJ (2000) Effects of frequency and level on auditory stream segregation. J Acoust Soc Am 108:1209–1214

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

Ryan MJ (1980) Female mate choice in a neotropical frog. Science 209:523–525

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

Ryan MJ, Drewes RC (1990) Vocal morphology of the *Physalaemus pustulosus* species group (Leptodactylidae): morphological response to sexual selection for complex calls. Biol J Linn Soc 40:37–52

Ryan MJ, Fox JH, Wilczynski W, Rand AS (1990) Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. Nature 343:66–67

Ryan MJ, Keddy-Hector A (1992) Directional patterns of female mate choice and the role of sensory biases. Am Nat 139:S4–S35

Ryan MJ, Rand AS (1990) The sensory basis of sexual selection of complex calls in the túngara frog, *Physalaemus pustulosus* (sexual selection for sensory exploitation). Evolution 44:305–314

Ryan MJ, Rand AS (2003) Sexual selection and female preference space: how female túngara frogs perceive and respond to complex population variation in acoustic mating signals. Evolution 57:2608–2618

Ryan MJ, Tuttle MD, Taft LK (1981) The costs and benefits of frog chorusing behavior. Behav Ecol Sociobiol 8:273–278

Schwartz JJ, Gerhardt HC (1995) Directionality of the auditory system and call pattern recognition during acoustic interference in the gray tree frog *Hyla versicolor*. Audit Neurosci 1:195–206

Wilczynski W, Rand AS, Ryan MJ (1999) Female preferences for temporal order of call components in the túngara frog: a bayesian analysis. Anim Behav 58:841–851

Wilczynski W, Rand AS, Ryan MJ (1995) The processing of spectral cues by the call analysis system of the túngara frog, *Physalaemus pustulosus*. Anim Behav 49:911–929

Yost WA, Sheft S (1993) Auditory perception. In: Yost WA, Popper AN, Fay RR (eds) Human Psychophysics. Springer, Berlin Heidelberg New York, pp 193–236

Zakon HH, Wilczynski W (1988) The physiology of the anuran VIIIth nerve. In: Fritzsch B, Ryan MJ, Wilczynski W, Hetherington TE, Walkowiak W (eds) The evolution of the amphibian auditory system. Wiley, New York, pp 125–155

Zar JH (1999) Biostatistical Analysis. Simon and Schuster, New Jersey