

ARTICLE

Received 25 Feb 2011 | Accepted 30 Jun 2011 | Published 2 Aug 2011

DOI: 10.1038/ncomms1417

Relative comparisons of call parameters enable auditory grouping in frogs

Hamilton E. Farris¹ & Michael J. Ryan^{2,3}

Whereas many studies on mate choice have measured the relative attractiveness of acoustic sexual signals, there is little understanding of another critical process: grouping and assigning the signals to their sources. For female túngara frogs, assigning the distinct components of male calls to the correct source is a challenge because males sing in aggregations, producing overlapping calls that lead to perceptual errors analogous to those of the ‘cocktail party problem’. Here we show that for presentation of >2 call components, however, subjects are more likely to group the two components with the smallest relative differences in call parameters, including relative spatial separation (a primitive acoustic cue) and relative similarity to the species-specific call sequence (a schema-based cue). Thus, like humans, the cognitive rules for the perception of auditory groups amidst multiple sound sources include the use of relative comparisons, a flexible strategy for dynamic acoustic environments.

¹ Neuroscience Center and Kresge Hearing Labs, Louisiana State University Health Sciences Center, New Orleans, Louisiana 70112, USA. ² Section of Integrative Biology, 1 University Station C0930, University of Texas, Austin, Texas 78712, USA. ³ Smithsonian Tropical Research Institute, Balboa, Panama. Correspondence and requests for materials should be addressed to H.E.F. (email: hfarri@lsuhsc.edu).

Auditory object formation involves sorting sounds into groups and assigning them to their correct sources^{1–3}. Object formation is complicated by the fact that sounds from different sources combine linearly and arrive at a receiver as a single complex wave which is not source specific, meaning groups and source assignment must be ‘deduced’^{4–7}. Thus, during acoustic communication⁸, perception of information and identification of its sender are inherently linked to how receivers make such deductions. The auditory system often accomplishes this by deconstructing the complex wave into its individual components, a result of processing so called ‘primitive’ or ‘general purpose’ acoustic cues (for example, temporal, spectral, spatial, and amplitude parameters)^{3,4,7,9}. Correct grouping and source assignment result from the ability to compare within and between these cues, a process made more difficult when there are multiple sources producing similar sounds. Using an approach most often applied with humans, this study determines how mate-searching female frogs deduce the source of complex male signals in breeding aggregations on the rain forest floor.

Tests of object formation often measure the cue or acoustic parameter values that elicit grouping within, but not beyond a certain boundary^{10–14}. Under conditions in which multiple cues overlap, however, absolute boundaries may not explain all perceptual grouping, as listeners may employ a relative strategy in which objects are formed by grouping stimuli with the most similar acoustic parameters and excluding those more distinct. Indeed, in humans variations of the streaming paradigm have shown that groups are not based on absolute tone differences, but result from relative comparisons among the tones available for grouping. For example, O’Conner and Sutter¹⁵ presented an ascending series of tones that, without background sounds, elicits perception of a single auditory object (that is, the target). If a sequence of background tones is added before and after the target sequence two groups are formed, one of which includes portions of the target that are perceptually captured into the background depending on target-background frequency separation¹⁶. Thus, the addition of the background ‘breaks’ the perceptual group that would otherwise be formed in isolation. Taken together with the fact that a complex grouping response-function emerges when the frequency separation between the target tones is also varied, it leads to the conclusion that relative (not absolute) frequency separation can determine the auditory stream^{3,11}.

In addition to primitive cues, object formation is also affected by experience or internal state^{6,11,17,18}. These are components of schema-driven object formation, in which the probability of grouping varies with a stimulus’ similarity to stored information, such as an innate or learned (for example, instructed) template^{7,11}. With respect to communication sounds, such a strategy is known from research on human speech^{5,6,9,19}. Thus, perceptual groups based on relative comparisons may not be limited to the relative differences in primitive cues, as similarity to schematic templates such as species-specific signal patterns could also be used to deduce auditory objects.

The wealth of data on humans notwithstanding, the challenge of object formation is not restricted to our own species. Various species across disparate taxa must also identify and locate complex acoustic signals in multi-source environments^{20–22}. Indeed, several groups of animals produce acoustically complex mating calls in multi-male choruses. As with humans during the ‘cocktail party problem’²³, females attending to male choruses must correctly identify and group these overlapping sounds to choose an appropriate mate^{24–28}. Understanding how females perform this task is critical for evaluating the underlying mechanisms of mate choice and how those mechanisms (and the signals) evolve. Furthermore, this understanding also informs us as to the degree of convergence/divergence in humans and other animals in responding to this challenge.

Using female túngara frogs (*Physalaemus pustulosus*), this study tested the hypothesis that auditory grouping of call components results from the comparison of relative spatial (primitive) and call

sequence (schema) cues. In multi-male choruses males produce complex calls consisting of two components: the frequency modulated ‘whine’, which is necessary and sufficient to elicit phonotaxis; followed by 0–7 broadband ‘chucks’. Chucks are never produced alone and can only follow a whine due to male vocal anatomy²⁹, creating a schema-based cue for whine-chuck grouping. Farris *et al.*^{30,31} showed that these two call components are perceptually distinct, as the whine is used in both ‘what and where’ decisions, and the chuck is only used in ‘where’ decisions. Important to this study is that phonotaxis to a chuck is contingent on the presentation of a whine^{30,31}. This combination sensitivity, in part, reveals auditory grouping of the whine and chuck. Previous tests of whine-chuck grouping have shown that absolute spatial and temporal differences have a limited effect (for example, grouping up to 135° whine-chuck separation and for non-natural whine-chuck sequences) and predict that in a dense male chorus whines and chucks from different sources could be incorrectly grouped^{30,31}. The stimuli in previous tests, however, used only two call components, a single whine and a single chuck. Thus, in this study we determine if such poor grouping acuity would be eliminated in a multi-source (> 2 sources) environment by testing whether object formation results from comparisons between available call components: we test the hypothesis that components with the smallest relative differences are more likely to be grouped¹⁵.

Results

Control stimuli. As with our previous work^{30,31}, responses to control stimuli confirm that the whine is necessary and sufficient to elicit and direct phonotaxis (Fig. 1). Data show that exit angles are non-random and localized to the position of the whine (Fig. 2; Table 1). New to this study is the control showing that the presentation of two chucks without a whine elicits the same response as that for

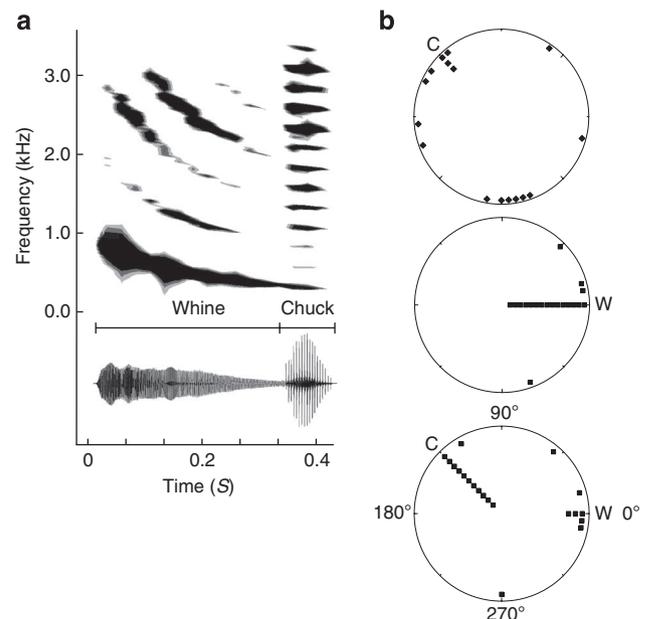


Figure 1 | Túngara frog call components and conditional response

to the chuck. (a) Spectrogram and time waveform of male call.

(b) Demonstration of auditory grouping of a single whine and chuck.

Squares are the exit angles from the circular arena for individual females.

(Top) Responses to the chuck alone are not oriented. (Middle) Whines alone are necessary and sufficient to elicit directed phonotaxis. (Bottom)

When a whine and chuck (with normal timing) are spatially separated up to 135°, females move to the chuck, indicating these two stimuli are perceived as a single object at the location of the chuck. Data are from a previous study³⁰. Polar coordinates are the same in all figures and tables. C, location of a chuck; W, location of a whine.

the presentation one chuck^{30,31}. First, chucks often do not elicit a response (Table 1). Second, circular analysis shows that if responses do occur, their direction is randomly distributed. Third, when compared to the categorical responses to a single chuck in our two previous studies^{30,31} (that is, 11.6% of females are expected to exit the arena within 10° of any chuck), the percentage of responses to either chuck of the 2-chucks stimulus was not significantly different (Fig. 2; Table 1). Thus, the failure of the two chucks to elicit a directed response maintains the critical control for concluding auditory grouping of the complex call components: phonotaxis to any chuck occurs only in the presence of the whine.

Spatial separation on same side of the whine. Responses to broadcasts of a whine and a pair of chucks with different relative whine-chuck separations show that grouping is more likely for the chuck closest to the whine (Fig. 3; Table 2). When the closest chuck has a 0° spatial separation relative to the whine, exit angles were non-randomly distributed and the mean exit angles did not differ from 0°. Furthermore, the presence of the second, more spatially separated chuck did not alter responses when compared to those for a single whine and chuck at 0° spatial separation (Fig. 3a–d; Table 2). Confounded in this stimulus set, however, is the fact that the chuck closest to the whine is broadcast from the same position of the whine, meaning the design can not distinguish conditional responses to the closest chuck from those to the whine. Using broadcasts of two spatially separated chucks with the closest at 45° (re. whine) revealed the same response, however. First, control data with two chucks of equal whine-chuck separation show that 16/24 trials exhibited grouping. There was no bias to either chuck, as equal separation elicits equal probability of being grouped with the whine (7/24 vs 9/24; $P = 0.38$;

Fig. 3e). When one chuck is separated farther, however, data show strong grouping of the closest chuck: the response is not random and the mean exit angle is not different from 45°. Furthermore, the mean exit angle is nearly identical to that for a single whine-chuck stimulus at 45° spatial separation (Fig. 3f–h; Table 2). This means that grouping of the closest chuck is categorical, as there is no evidence that females averaged the position of the two chucks.

Spatial separation on opposite sides of the whine. The previous experiments created whine-chuck separations in one hemisphere of the circular arena. This means that when facing the whine the two chucks would be lateralized to the same side of the frog, potentially engaging different mechanisms for assessing spatial coherence than if the chucks were lateralized to different sides. Thus, we tested whether stimulus orientation affects spatial grouping such that the two chucks are now on opposite sides of the whine. Control data for equal separation (Fig. 3i; 45° vs 45°) again showed no bias in grouping responses to either of the chucks (10/25 vs 11/25; $P = 0.5$). For different separations, only the 45° vs 135° chuck separation matched the response of the same-side stimuli above, as the mean exit angle was significantly localized at the closest chuck and the distribution was not different from that for a whine and single chuck at 45° (Fig. 3k; Table 2). Grouping of the chuck closest to the whine, however, did not hold when the relative spatial separations of the two chucks were 45° and 90° on opposite sides of the whine (Fig. 3j; Table 2), as responses were more similar to the 45° vs 45° control.

Masking verses grouping spatial acuity. The results above suggest that while grouping whines and chucks, spatial resolution during presentation of a pair of chucks is more acute (45°) than that for grouping a single chuck ($\leq 135^\circ$)^{30,31}. Also, grouping responses to multi-chuck stimuli appeared categorical or limited to one chuck at a time, as the responses shown in Figure 3 reveal that exit angles are not located at a position that is an average of the two chucks. Taken together, these data generate the hypothesis that increased spatial resolution during the multi-chuck stimulus prevents detection of the non-grouped chuck. To test this, the grouping experimental paradigm is combined with that of spatial masking by presenting a whine, a single chuck and a burst of white noise with the same envelope and temporal parameters of the chuck (that is, replacing the second chuck). The experiment tested if during whine-chuck grouping a third, non-grouped, spatially separated stimulus is indeed detected by observing whether the noise burst could mask the chuck.

A premise of this experiment is that the noise is a masker and does not elicit a phonotactic or grouping response on its own. Thus, control broadcasts first established that a broadband noise with same envelope structure of a chuck is not grouped (Fig. 4). Like chucks broadcast alone, the noise alone does not elicit phonotaxis and exit angles are not significantly different from random ($r = 0.16$; $P = 0.652$). When the noise is presented with the whine (90° spatial separation), although there is a significant shift in mean exit angle away from that generated from presenting a whine alone (mean angle: whine 94.3° vs whine-noise 60.8°; $F = 15.7$; $P < 0.0003$; Fig. 4),

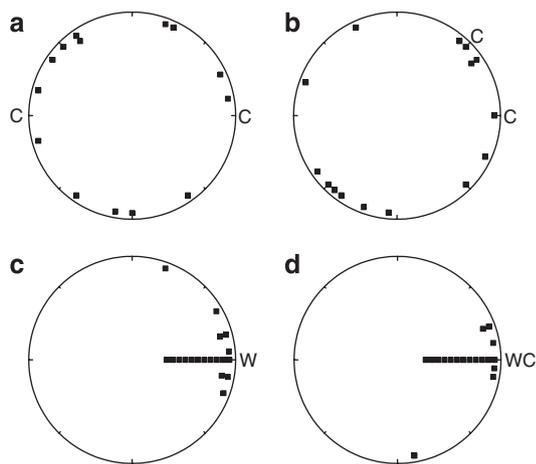


Figure 2 | Exit angles for control stimuli. (a) Two chucks with 180° spatial separation; (b) Two chucks with 45° spatial separation; (c) Whine presented alone; (d) Whine-chuck with normal timing and 0° spatial separation. C, chuck; W, whine.

Table 1 | Analysis of phonotaxis direction for a whine or chucks alone or in natural combination.

Control stimuli	<i>n</i>	Mean angle (±95%)	Vector (<i>r</i>)	Raleigh test (<i>P</i>)	Chuck resp.	Non-chuck resp.	No resp.	Expect resp. (<i>P</i>)	V-test (<i>P</i>)
Two chucks (45°)	23	297° (180)	0.155	0.705	5	10	8	0.1289	
Two chucks (180°)	21	140° (180)	0.196	0.592	1	13	7	0.3278	
Whine	19	4° (9)	0.952	<0.0001					<0.0001
Whine-Chuck (0°)	18	358° (9)	0.946	<0.0001					<0.0001

Columns are: the stimuli (with stimulus separation); number of females tested (*n*); mean (±95% confidence interval) exit angle; vector strength (*r*); Raleigh test for response different from random; number of responses (±10°) to a chuck; number of non-chuck responses; number of females exhibiting no response to chuck only stimuli; Fisher exact test comparing number of responses to chucks in the two chuck conditions to those when presented a single chuck²⁷; V-test showing mean exit angles are localized at the position of the stimulus (0°).

this shift is significantly less than that for a grouped chuck, as the whine-chuck response is significantly closer to the chuck position (mean angle: whine-noise 60.8° vs whine-chuck 28.0°; $F=9.0$;

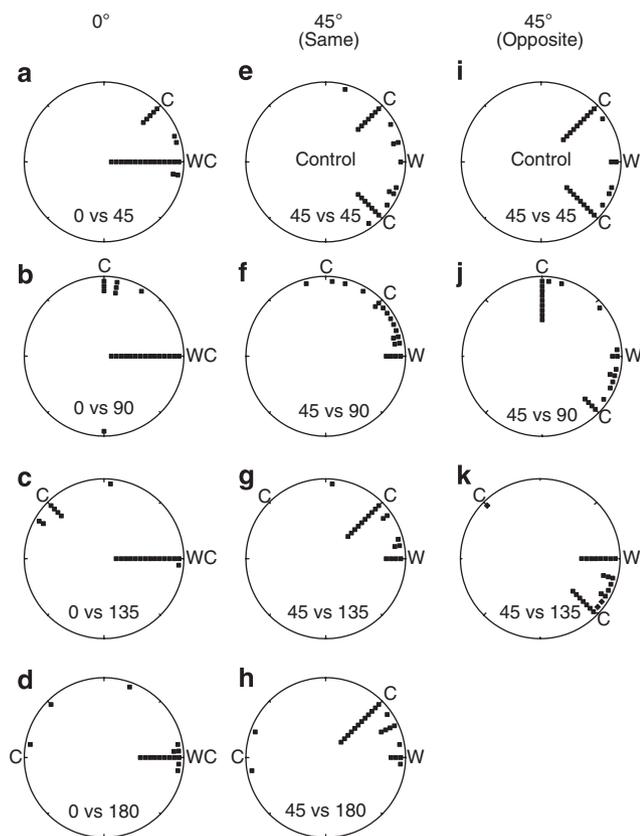


Figure 3 | Exit angles in response to a whine presented with two competing chucks. Each chuck has a natural timing but at varying whine-chuck separations. (a–d) Closest chuck has 0° whine-chuck separation. (e–h) Closest chuck is 45° from the whine, with both chucks on same side of whine. (i–k) Closest chuck is 45° from the whine, but on opposite sides of whine. (e, i) Controls with chucks at 45° on each side of the whine were run for each experiment. C, chuck; W, whine.

$P<0.005$). Furthermore, in categorical analysis, the number of direct responses to the noise ($\pm 10^\circ$) broadcast with the whine is not significantly different from that when the noise is broadcast alone (whine-noise 4/25 vs noise alone 3/25; $P=0.5$; Fig. 4). And, this number is significantly less than that measured for attraction to a chuck broadcast with a whine (whine-noise 4/25 vs whine-chuck 16/25; $P<0.0007$). Thus, the broadcast of a spatially separated noise at the natural temporal position of a chuck does not elicit the grouping response. These data allow for the noise to be used as a masker.

When the noise is presented simultaneously with the chuck, each at 90° from the whine and 180° from each other, the chuck is no longer grouped (Fig. 4). Direct responses to the chuck are significantly reduced (whine-chuck 16/25 vs whine-chuck + noise 1/25; $P<0.00005$) and the mean exit angle is nearly identical to that for the whine alone (whine 94.3° vs whine-chuck + noise 93.5°; $F=0.0043$; $P=0.948$). In addition to these analyses showing the effect on grouping by adding noise to a single whine-chuck stimulus (Fig. 4), the effect is also illustrated when comparing responses to a multi-chuck stimulus (Fig. 4d vs Fig. 5b). Thus, the response is consistent with masking. And, although we found a 45° resolution between multiple chucks during the grouping response (Fig. 3), this resolution does not appear to be based on limits of spatial unmasking, which was evident for our only chuck-noise separation at 180°. Note that no other (smaller) separations were tested because acuity was determined at the largest angle.

Schema-based grouping. The morphologically constrained whine-then-chuck sequence (Fig. 1) provides a cue for grouping and source assignment. Because the cue is based on stored information on the part of the receiver and not on source acoustics, results here address whether grouping in túngara frogs is schema-driven. Following the letters designating stimuli in Table 3, control stimuli (stims. a, b) first reconfirm that a naturally timed chuck at 90° spatial separation from the whine is grouped (Table 3; Fig. 5a,b). The addition of a second simultaneous chuck (equal separation, opposite side of the arena) showed that there is no bias in the arena: either chuck may be grouped, which is reflected in the bimodal distribution of the exit angles and their high dispersion (r). Also, the distribution is not different from random (stim. b; Fig. 5b).

In four of the stimulus conditions, the grouping response favored the chuck timed to the natural sequence. First, when the natural-timed chuck is broadcast with chucks that could not be produced by single

Table 2 | Analysis of phonotaxis direction for a whine with two spatially separated chucks in natural temporal sequence.

Chuck angles (re. whine 0°)	<i>n</i>	Mean exit angle ($\pm 95\%$ CI)	Vector (<i>r</i>)	Raleigh test (<i>P</i>)	Directed to closest chuck position (<i>P</i>)	Compared with grouping single chuck (<i>P</i>)
<i>Closest</i>	<i>Distant</i>					
0°	None	18	358° (9)	0.946	<0.0001	<0.0001
0°	45°	24	10° (8)	0.947	<0.0001	<0.0001
0°	90°	23	20° (22)	0.741	<0.0001	<0.0001
0°	135°	22	24° (45)	0.525	<0.0001	<0.001
0°	180°	17	8° (26)	0.75	<0.0001	<0.0001
<i>Same</i>						
45°	None	24	35° (16)	0.844	<0.0001	<0.0001
45°	90°	20	31° (14)	0.877	<0.0001	<0.0001
45°	135°	20	31° (10)	0.927	<0.0001	<0.0001
45°	180°	24	34° (17)	0.809	<0.0001	<0.0001
<i>Opposite</i>						
45°	90°	25	29° (35)	0.570	<0.001	>0.1
45°	135°	20	338° (13)	0.873	<0.0001	<0.0005

Columns are the spatial separation of single or a pair of chucks (re. whine); number of females (*n*); mean ($\pm 95\%$ confidence interval) exit angle; vector strength of the exit distribution; Raleigh test for random distribution; V-test of whether the exit angle distribution is localized to the position of the chuck closest to the whine; Watson-Williams test for difference between the exit distribution for single whine-chuck at 0° or 45° separation vs that for a pair of chucks with closest chuck at 0° or 45°. The exit angle distribution for the null hypothesis at 45° was previously collected²⁷. The spatially separated chucks are either on the same or opposite sides of the whine (Fig. 3).

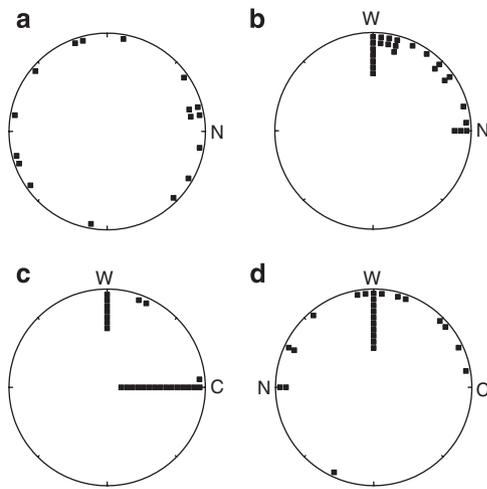


Figure 4 | Spatial resolution for masking and grouping are not the same. (a) Broadband noise (N) with the same envelope of the chuck does not evoke orientation by females. (b) Noise broadcast with a whine shifts response orientation, but does not elicit grouping response like that for a whine-chuck (c). (d) Noise broadcast simultaneously with a naturally timed chuck prevents grouping at 180° spatial separation. C, chuck; W, whine.

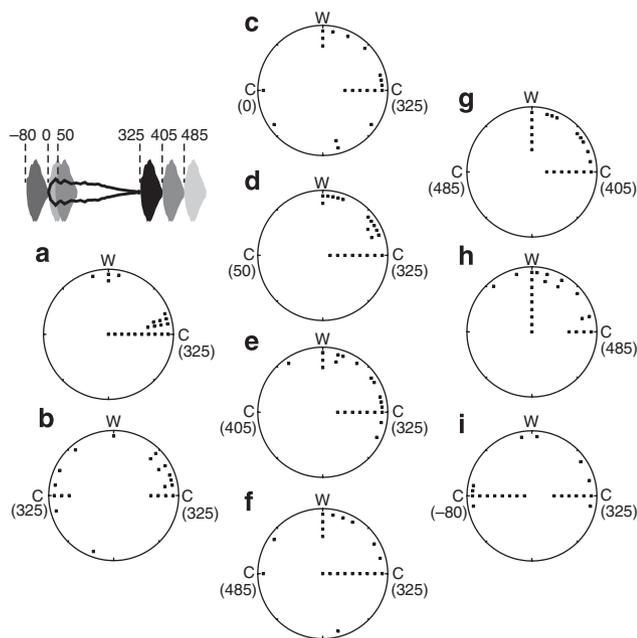


Figure 5 | The effects of relative whine-chuck sequence on grouping. Squares are exit angles in response to two competing chucks with the same separations (90° re. whine), but different timings (in parentheses, in ms relative to whine onset). Natural chuck timing is 325 ms. Letter for each plot (a–i) corresponds to stimulus designation in Table 3. Upper left is a diagram of the whine (open shape; onset time is 0 ms) and different chuck onset times to show the extent of temporal overlap.

males because they overlap the whine (stims. c, d; Fig. 5c,d), the mean exit angles are localized at the natural-timed chuck and the response distributions do not differ from that for a whine followed by a single natural-timed chuck (stim. a, Fig. 5a vs c,d). Failure to group overlapping chucks is not based on a ‘go to the last chuck’ rule, however, as the natural-timed chuck is also more likely to be grouped than those that follow it (stims. e, f). Unlike the chucks

overlapping the whine, the chucks following the natural-timed chuck could be produced by a male.

Stimulus condition g tests whether the decision rule is particular to the natural temporal position or just chucks timed closest to it (Table 3). Here, two chucks were presented which could be produced by a male, but neither is timed in the natural position of a single chuck: both are delayed and occupy times when a male would produce second and third chucks. The mean exit angle was not different from the speaker broadcasting the chuck timed closest to natural sequence (that is, the position of the 405 ms onset chuck; Fig. 5g), which is consistent with the hypothesis that grouping favors chucks closest to the natural time. However, the responses were significantly less directed to that chuck compared to the response to the naturally timed (325 ms onset) chuck presented alone with a whine ($P=0.03$, Table 3). This suggests an importance of the natural timing window. Furthermore, the failure of the 405 ms onset chuck to completely match responses to a natural-timed chuck occurred even though the last chuck (485 ms onset) did not elicit the same grouping response as the single chuck control (stim. h, Fig. 5h). Taken together the data suggest that under these stimulus conditions the natural time of a first chuck is weighted more heavily during grouping.

It is important to note that there was one striking exception to the dominance of the whine-chuck natural sequence, however. When a chuck precedes the whine, there is no longer significant grouping of the natural sequenced chuck (stim. i, Fig. 5i), as responses are nearly identical to those in control stimulus condition b, with both chucks eliciting grouping (Table 3; Fig. 5).

Discussion

With respect to spatial cues, a general similarity in both humans and frogs is that simple auditory spatial resolution is characterized by greater acuity than that used in grouping. In humans, for example, grouping can occur over larger stimulus separations (180°^{9,23,32,33} than the minimum audible angle³⁴. Indeed, the weakness of spatial or binaural cues in grouping is revealed by the fact that other cues can override them^{35–39}. Likewise in frogs, assays of behavior, ear biomechanics, and neurophysiology show that auditory spatial resolution varies with source separations of 90° (ref. 27) or much less^{40–43}. In contrast, during sequential pattern recognition similar to auditory streaming, songs may still be grouped or combined when separated by 120–180° (refs 44,45). Our previous data in *P. pustulosus* are consistent with this trend, as a single whine and chuck may be grouped over 135° (refs 30,31) (Fig. 1). Our present study (and data from humans and other frogs) shows, however, that use of spatial cues in grouping is not necessarily weak, in that acuity may be stimulus dependent^{9,11,15,32}. For example, whereas segregation of sound sequences by some treefrogs is tolerant of large separations, Bee⁴⁶ recently showed that in the context of processing simultaneous sounds, separation of call components by <10° can reduce attractiveness. This difference in spatial acuity between simultaneous and sequential tasks is consistent with that exhibited by humans⁴⁷. Our study adds another stimulus context to such results showing greater acuity in grouping. Importantly, we show that for sounds that would otherwise be grouped over wide spatial separations on their own, greater spatial acuity is revealed with the addition of more sources: use of relative separation for grouping occurs over 45° compared to 135° for a single whine and chuck^{30,31}. But this greater acuity is limited to certain conditions and stimuli. First, greater spatial acuity during grouping is not uniform for different stimulus orientations. Acuity is reduced (that is, more separation is required for grouping) when chucks are on either side of a whine as opposed to the same side. Second, acuity is limited to the addition of more chucks and not noise, which produced masking over the largest possible separation (180°).

The results are important not only for the reduction in acuity, but also because they reveal a strategy that is not based on absolute processing criteria. That is, use of spatial cues in grouping appears

Table 3 | Analysis of phonotaxis direction for a whine with two chucks at equal separation, but timed differently relative to the onset of the whine.

Stim.	Timing of chucks (ms) at spatial positions of 0° and 180°		n	Mean exit angle	Vector (r)	Raleigh test (P)	V-test at 0° (P)	Comparison with Null Ho (P)
	0°	180°						
a.	325	None	22	18°	0.846	<0.0001	<0.0001	Null Ho
b.	325	325	21	47°	0.293	0.165	0.1	
c.	325	0	20	17°	0.511	<0.0001	<0.001	1.0
d.	325	50	20	32°	0.837	<0.0001	<0.0005	0.323
e.	325	405	23	33°	0.75	<0.0001	<0.0005	0.324
f.	325	485	21	35°	0.597	<0.0001	<0.001	0.467
g.	405	485	22	44°	0.786	<0.0001	<0.001	0.03
h.	485	None	25	68°	0.799	<0.0001	<0.025	<0.0001
i.	325	-80	22	130°	0.162	0.568	>0.1	

Columns are the stimulus designation from Figure 5; the timing of the chucks relative to the onset of the whine at 0 ms (whine spatial position is 90°, chuck position is 0° or 180°); number of females (n); mean exit angle; vector strength (r) showing dispersion of the exit distribution; Raleigh test for random distribution; V-test showing probability that the exit distribution is different from the stimulus at 0°; Watson-Williams test for difference between the mean exit angle and that for single whine-chuck in natural sequence at 90° separation (Null Ho). Distributions without significant vectors (stims. **b** and **i**) are not used in the Watson-Williams test.

flexible. Often, studies of spatial effects on grouping test the extent to which spatial separation still allows for grouping. For example, the spatial separation of a component or set of components in a multi component stimulus is increased until it 'pops out' or is perceived as a separate set of objects. Here, we test a special case where a subject is given a stimulus in which all components are within the spatial separation for grouping. The subject is then asked whether there is equal likelihood of grouping as a function of relative separation. This new context of relative comparisons notwithstanding, our data further demonstrate the stimulus dependence of spatial grouping by showing that responses vary with stimulus orientation and the presence of a masker. Thus, four stimulus contexts are now known to affect spatial acuity when grouping whines and chucks: separation, relative separation, orientation and the presence of a masker.

Perception of auditory objects may be based on stored information that is comprised of numerous auditory attributes⁶. Although in túngara frogs there is (so far) no evidence that grouping varies with instructed or learned attributes, there are a plethora of data showing that stored information is employed to evaluate sound. Several studies have shown, for example, that both gross and fine FM characteristics of the whine enable call recognition and conspecific versus heterospecific discrimination^{48,49}; working memory varies with signal complexity⁵⁰; and, whines and chucks are processed differently in recognition and location decisions^{30,31}. Prior to this study, however, tests of call recognition, preference or grouping of a single whine and chuck revealed little evidence for the use of call sequence in phonotactic decisions^{31,51}. We now show that information contained in the call sequence can be used to determine call sources when multiple sources are presented. In particular, in all but one case the data implicate a receiver strategy that includes information regarding the production of the whine and chuck to determine their source. That is, chucks occurring at times that are either impossible or less likely to be produced by a single male are not grouped with the whine when a natural-timed chuck is present. For example, chucks overlapping the onset of the whine, although grouped when presented as a single chuck with a whine, are no longer grouped in favor of a second chuck that is timed to the natural sequence. There are no source-specific acoustic reasons for this strategy (for example, location, harmonic relationship, etc.), leading to the conclusion of schema-based grouping that uses relative comparisons. Our data show that there is variance in this strategy, however, as grouping of a chuck preceding the onset of the whine is as likely as a chuck following the whine in the natural sequence. This exception could be explained if single males used 100% duty cycle to produce a series of calls, placing the chuck from a previous call just before the whine of a subsequent call. This is unlikely, however, as call period is

approximately 1.5 s too long⁵². Thus, failure to bias grouping to the natural sequence in this case could still lead to whine-chuck grouping errors in a chorus. Such inaccurate grouping could be a result of preferences for particular male traits, however. Although studied in a preference rather than grouping paradigm, there is evidence that changes in the relative timing of stimuli can alter strong preferences for particular call attributes. For example, in *Hyla cinerea* a signal that is nearly ignored when presented simultaneously with another, receives the majority of phonotaxis responses when presented in a leading or preceding temporal position⁵³. This effect of timing has been shown across taxa^{54,55}, including túngara frogs. Females prefer preceding calls and males compete to time their calls so that they precede others⁵⁶. Although our data suggest that this preference for the leading signal is manifested in auditory grouping as well as during call choice, the reason for this exception in grouping responses is still unclear.

This study reached three conclusions regarding auditory object formation in mate-searching frogs. When multiple sources create the potential for multiple auditory groups: túngara frogs use relative spatial separation to make grouping decisions; spatial acuity for grouping is not explained by spatial acuity for masking; and schema-based cues from call sequence can be used to include or exclude calls from auditory groups. These are important because they: reveal the perceptual mechanisms that enable females to assign sexual signals to males, a process inextricably linked to mate choice; and elucidate common cognitive processing that are shared by widely disparate taxa in solving similar problems.

Methods

Subjects. Females were collected in amplexus in Gamboa, Panama approximately 3 h after sunset in June 2007–2009. Animals were marked (toe clipped) to prevent retesting. All procedures were performed in accordance with the guidelines established by the Smithsonian Tropical Research Institute, the University of Texas and LSUHSC Animal Care and Use Committees.

Experimental procedure. The experimental procedure is the same as that used previously^{30,31}. Phonotactic responses were measured to whines and chucks presented alone or in combination at various spatial separations. Each trial began by placing a gravid female *P. pustulosus* beneath a mesh cone (10 cm diameter) at the center of a circular array of 5 speakers on the floor of the sound chamber. Following 3 min of exposure to the stimuli, the cone was remotely removed allowing a subject to move freely while the stimuli continued to play. Using an infrared camera, we (minimum of 3 observers) recorded the position (in 5° intervals) at which females exited the full perimeter. Video of all trials was recorded 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 were not tested again. Females were tested only once per stimulus. Subjects controlled their orientation with respect to the speakers at all times during the experiments (that is, in or out of the cone). Although this design differs from that with headphones or head holders in humans, it maintains the ethological relevance of the task, as females must not only determine sound source location, they must also walk to it. All trials were performed within 12 h of sunset under infrared illumination. Ambient temperature was $\sim 27^{\circ}\text{C}$. Within the natural range produced by males, the peak amplitude of all chucks was 6 dB re. whine amplitude (90 dB SPL re. 20 μPa ; call amplitude at 50 cm^{57}). For each experiment, stimuli were presented in random order. Females remain responsive for up to ~ 18 different stimulus presentations. Because the experiments here are comprised of > 30 different stimuli, the experiments were run on three separate sample populations over three summers.

Stimuli. For consistency, the whine and chuck were the same as those used by Farris *et al.*^{30,31} (Fig. 1). These call components were those of an average call from Gamboa, where we collected females. The average was calculated from 14 acoustic variables from 250 complex calls of 50 males (five calls per male)⁵⁸. Call period for all stimuli was 2 seconds.

Stimuli were generated using an Audigy 2 ZS (Creative Labs) 7 channel sound card; amplified using Crown (XLS 202) amplifiers; and broadcast from Radioshack 4 inch broadband speakers (# 40–1040) speakers positioned along the perimeter of a 75 cm radius arc inside an Acoustic Systems sound chamber (2.75 \times 1.83 m) lined with additional anechoic foam along the bottom 0.6 m of each wall (Sonex, 1.5 inch). Prior to each night's trials, the peak amplitude of the whine and chuck stimuli were calibrated using a 500 Hz tone, a GenRad 1982 sound level meter (Fast, linear weighting) and 0.5 inch microphone placed 3 cm above the floor of the arena at the center of the arc. To control for potential position effects and to reduce any phonotactic bias due to potential speaker variability, the broadcast orientation (in the array) was randomized. Furthermore, the speakers were switched three times during an experiment (after ~ 7 females).

Effect of relative spatial separation experiment. Chucks cannot be produced by themselves due to morphological constraint²⁹. Thus, whine-chuck spatial coherence is a potential cue for grouping. However, previous data showed that for a single whine and a single chuck, absolute spatial separation had very little effect on grouping: significant grouping occurred up to 135° whine-chuck separation^{30,31} (Fig. 1). With the addition of a second chuck, experiments here used 3 stimuli to test whether the relative whine-chuck separation affects grouping. Using the circular phonotaxis arena described above, a single whine was presented in natural sequence with two identical chucks that varied independently in spatial separation from the whine (0 – 180°). Because of the importance of spatial orientation in the acuity of auditory spatial processing³⁴, two permutations were used. The first varied the relative whine-chuck spatial separation of the two chucks so that they were on the same side of the whine. This means that when the whine is at 0° normal to the front of the frog the two chucks are lateralized in the same hemisphere. The second permutation used the same relative whine-chuck separations, but placed the chucks on opposite sides of the whine, allowing for tests of whether relative spatial cues are affected by lateralization.

The direction of female phonotaxis (that is, the exit angle from the arena) was measured to test the null hypothesis that grouping is directed towards the chuck closest to the whine, showing that object formation results from comparisons between spatial separations.

Spatial acuity of masking and grouping experiment. The categorical grouping responses in experiment 1 (see results) suggested that only one of the spatially separated chucks may be detected at a time, raising the possibility that there is high spatial acuity around a single chuck (see results). Thus, different from the grouping task, this experiment examined the spatial resolution for chuck detection using a masking paradigm (cf. spatial release from masking). Here, a single whine and chuck were presented in the natural sequence at a constant 90° spatial separation, well within the grouping angle for single whines and chucks^{30,31}. Band-limited (50 Hz–10 kHz; 98 dB/octave skirts) noise with the chuck's envelope structure (Hilbert transform⁵⁹) was presented simultaneously with the chuck from a third speaker. The bandwidth was chosen to encompass the spectrum of the entire chuck, as full spectrum chucks are more likely grouped than filtered ones³¹. The experiment measured the smallest chuck-noise separation which still enabled the conditional response to the chuck (that is, where the chuck was not masked and still grouped with the whine).

Effect of relative sequence experiment. The morphologically limited whine-then-chuck syntax provides a schema-based cue for grouping²⁹. For example, chucks and whines with simultaneous onsets cannot be produced by the same male. Thus, this experiment used 3 stimuli: a whine and two chucks at equal spatial separation (90° re. whine). Relative to that of the whine, the timing of the two chucks is varied to test whether grouping is more likely for the chuck that was broadcast closest to the time of the natural sequence.

Analysis. Using females' 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 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 (a 20° arc; 13 cm). The second category included 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 assessed whether the number of responses in each category differed between stimulus conditions or expected response numbers generated by previous data^{30,31}.

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 exit angles were randomly distributed. Second, a V-test determined whether responses were localized at a particular exit angle⁶⁰. The V-test, however, is most powerful when the distribution of exit angles is unimodal. Because phonotaxis is also directed to the whine in some trials, the V-test may lack resolution for smaller spatial separations. Thus, differences in mean exit angles between various broadcast conditions were analyzed using an additional method, the Watson-Williams test⁶⁰ with alpha correction for multiple comparisons. By comparing distributions of exit angles between different stimulus conditions, this test reveals whether the exit angles were affected by altering the stimuli. 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 control broadcasts (for example, whine alone; whine and single chuck). Watson-Williams tests are not performed with distributions that are not different from random, however, as the mean of such a distribution will have 180° variance, causing the test to show that it has a similar mean to any other distribution. Unlike the categorical analysis, circular analyses were restricted to trials in which the females exited the perimeter, excluding 'no choice' responses.

References

- Shamma, S. A. & Micheyl, C. Behind the scenes of auditory perception. *Curr. Opin. Neurobiol.* **20**, 361–366 (2010).
- Mellinger, D. K. & Mont-Reynaud, B. M. in *Auditory Computation* (eds Popper, A. N. & Fay, R. R.) 271–331 (Springer, 1996).
- Carlyon, R. P. How the brain separates sounds. *Trends Cogn. Sci.* **8**, 465–471 (2004).
- Yost, W. A. & Sheft, S. in *Human Psychophysics* (eds Yost, W. A., Popper, A. N., & Fay, R. R.) 193–236 (Springer Verlag, 1993).
- Darwin, C. J. Listening to speech in the presence of other sounds. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **363**, 1011–1021 (2008).
- Ciocca, V. The auditory organization of complex sounds. *Front. Biosci.* **13**, 148–169 (2008).
- Bregman, A. S. *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA, 1990).
- Bradbury, J. W., & Vehrencamp, S. L. *Principles of Animal Communication* (Sinauer Assoc. Inc., Sunderland, MA, 1998).
- Darwin, C. J. & Carlyon, R. P. in *Hearing* (ed. Moore, B. C.) 387–424 (Academic Press, 1995).
- Grimault, N. *et al.* Perceptual auditory stream segregation of sequences of complex sounds in subjects with normal and impaired hearing. *Br. J. Audiol.* **35**, 173–182 (2001).
- Moore, B. C. J. & Gockel, H. Factors influencing sequential stream segregation. *Acta Acust. United Ac* **88**, 320–333 (2002).
- Roberts, B., Glasberg, B. R. & Moore, B. C. Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *J. Acoust. Soc. Am.* **112**, 2074–2085 (2002).
- Grimault, N., Bacon, S. P. & Micheyl, C. Auditory stream segregation on the basis of amplitude-modulation rate. *J. Acoust. Soc. Am.* **111**, 1340–1348 (2002).
- McCabe, S. L. & Denham, M. J. A model of auditory streaming. *J. Acoust. Soc. Am.* **101**, 1611–1621 (1997).
- O'Connor, K. N. & Sutter, M. L. Global spectral and location effects in auditory perceptual grouping. *J. Cogn. Neurosci.* **12**, 342–354 (2000).
- Bregman, A. S. & Rudnick, A. I. Auditory segregation: stream or streams? *J. Exp. Psychol. Hum. Percept. Perform.* **1**, 263–267 (1975).
- Ciocca, V. & Bregman, A. S. Perceived continuity of gliding and steady-state tones through interrupting noise. *Percept. Psychophys.* **42**, 476–484 (1987).
- Darwin, C. J. Simultaneous grouping and auditory continuity. *Percept. Psychophys.* **67**, 1384–1390 (2005).
- Darwin, C. J. Perceiving vowels in the presence of another sound: constraints on formant perception. *J. Acoust. Soc. Am.* **76**, 1636–1647 (1984).
- Bee, M. A. & Micheyl, C. The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *J. Comp. Psychol.* **122**, 235–251 (2008).
- Feng, A. S. & Ratnam, R. Neural basis of hearing in real-world situations. *Annu. Rev. Psychol.* **51**, 699–725 (2000).
- Fay, R. R. in *Auditory Perception of Sound Sources* (eds Yost, W. A., Popper, A. N., & Fay, R. R.) 307–323 (Springer, 2008).
- Cherry, E. C. Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am.* **25**, 975–979 (1953).

24. Gerhardt, H. C., & Huber, F. *Acoustic communication in insects and anurans* (University of Chicago, Chicago, 2002).
25. Gerhardt, H. C. & Klump, G. M. Masking of acoustic-signals by the chorus background-noise in the green tree frog - A Limitation on Mate Choice. *Anim. Behav.* **36**, 1247–1249 (1988).
26. Wollerman, L. & Wiley, R. H. Background noise from a natural chorus alters female discrimination of male calls in a neotropical frog. *Anim. Behav.* **63**, 15–22 (2002).
27. Bee, M. A. Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Anim. Behav.* **74**, 549–558 (2007).
28. Bee, M. A. Finding a mate at a cocktail party: Spatial release from masking improves acoustic mate recognition in grey treefrogs. *Anim. Behav.* **75**, 1781–1791 (2008).
29. Ryan, M. J. & Drewes, R. C. Vocal morphology of the *Physalaemus-pustulosus* species group (Leptodactylidae) - morphological response to sexual selection for complex calls. *Biol. J. Linn. Soc. Lond.* **40**, 37–52 (1990).
30. Farris, H. E., Rand, A. S. & Ryan, M. J. The effects of spatially separated call components on phonotaxis in túngara frogs: evidence for auditory grouping. *Brain Behav. Evol.* **60**, 181–188 (2002).
31. Farris, H. E., Rand, A. S. & Ryan, M. J. The effects of time, space and spectrum on auditory grouping in túngara frogs. *J. Comp. Physiol. A* **191**, 1173–1183 (2005).
32. Buell, T. N. & Hafter, E. R. Combination of binaural information across frequency bands. *J. Acoust. Soc. Am.* **90**, 1894–1900 (1991).
33. Cherry, E. C. & Weary, D. M. Some further experiments upon the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* **26**, 554–559 (1954).
34. Mills, A. W. On the minimum audible angle. *J. Acoust. Soc. Am.* **30**, 237–246 (1958).
35. Barsz, K. Auditory pattern perception: the effect of tone location on the discrimination of tonal sequences. *Percept. Psychophys.* **50**, 290–296 (1991).
36. Best, V. *et al.* The influence of spatial separation on divided listening. *J. Acoust. Soc. Am.* **120**, 1506–1516 (2006).
37. Darwin, C. J. & Hukin, R. W. Auditory objects of attention: the role of interaural time differences. *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 617–629 (1999).
38. Deutsch, D. Binaural integration of melodic patterns. *Percept. Psychophys.* **25**, 399–405 (1979).
39. Gockel, H. & Carlyon, R. P. Effects of ear of entry and perceived location of synchronous and asynchronous components on mistuning detection. *J. Acoust. Soc. Am.* **104**, 3534–3545 (1998).
40. Jorgensen, M. B. & Gerhardt, H. C. Directional hearing in the gray tree frog *Hyla versicolor*: eardrum vibrations and phonotaxis. *J. Comp. Physiol. A* **169**, 177–183 (1991).
41. Lin, W. Y. & Feng, A. S. Free-field unmasking response characteristics of frog auditory nerve fibers: comparison with the responses of midbrain auditory neurons. *J. Comp. Physiol. A* **187**, 699–712 (2001).
42. Schwartz, J. J. & Gerhardt, H. C. Spatially mediated release from auditory masking in an anuran amphibian. *J. Comp. Physiol. A* **166**, 37–41 (1989).
43. Christensen-Dalsgaard, J. in *Sound Source Localization* (eds. Popper, A. N. & Fay, R. R.) 67–123 (Springer, 2005).
44. Schwartz, J. J. & Gerhardt, H. C. Directionality of the auditory system and call pattern recognition during acoustic interference in the gray tree frog, *Hyla versicolor*. *Audit. Neurosci.* **1**, 195–206 (1995).
45. Bee, M. A. & Riemersma, K. K. Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs? *Anim. Behav.* **76**, 831–843 (2008).
46. Bee, M. A. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (*Hyla chrysoscelis*). *J. Comp. Psychol.* **124**, 412–424 (2010).
47. Darwin, C. J. in *Auditory perception of sound sources*, (eds. Yost, W. A. Popper, A. N. and Fay, R. R.) 215–232 (Springer, 2008).
48. Ryan, M. J. & Rand, A. S. Phylogenetic influence on mating call preferences in female túngara frogs, *Physalaemus pustulosus*. *Anim. Behav.* **57**, 945–956 (1999).
49. Ryan, M. J. & Rand, A. S. Species recognition and sexual selection as a unitary problem in animal communication. *Evolution* **47**, 647–657 (1993).
50. Akre, K. L. & Ryan, M. J. Complexity increases working memory for mating signals. *Curr. Biol.* **20**, 502–505 (2010).
51. Wilczynski, W., Rand, S. A. & Ryan, M. J. Female preferences for temporal order of call components in the túngara frog: a Bayesian analysis. *Anim. Behav.* **58**, 841–851 (1999).
52. Pauly, G. B. *et al.* The vocal sac increases call rate in the túngara frog *Physalaemus pustulosus*. *Physiol. Biochem. Zool.* **79**, 708–719 (2006).
53. Hobel, G. Interaction between signal timing and signal feature preferences: causes and implications for sexual selection. *Anim. Behav.* **79**, 1257–1266 (2010).
54. Marshall, V. T. & Gerhardt, H. C. A precedence effect underlies preferences for calls with leading pulses in the grey treefrog, *Hyla versicolor*. *Anim. Behav.* **80**, 139–145 (2010).
55. Greenfield, M. D. & Snedden, W. A. Selective attention and the spatio-temporal structure of orthopteran choruses. *Behaviour* **140**, 1–26 (2003).
56. Greenfield, M. D. & Rand, A. S. Frogs have rules: Selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). *Ethology* **106**, 331–347 (2000).
57. Ryan, M. J. *The Túngara Frog, a Study in Sexual Selection and Communication* (University of Chicago Press, Chicago, IL, 1985).
58. Ryan, M. J. & Rand, A. S. Sexual selection in female perceptual space: how female túngara frogs perceive and respond to complex population variation in acoustic mating signals. *Evolution* **57**, 2608–2618 (2003).
59. Hartmann, W. H. *Signals, Sound, and Sensation* (AIP Press, New York, 1998).
60. Zar, J. H. *Biostatistical Analysis* (Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1999).

Acknowledgements

Helpful input was provided by: T. Forrest, W. Wilczynski, X. Bernal, A. Baugh, E. Golob, A. Lea and S. Rand. N. Lessios, A. Dapper, L. Zeigler, S. Ozeroff and J. Saunders assisted in collecting frogs and running experiments. This research was supported by NSF IBN 98 16564 to MJR. HEF was supported by NIH P20RR016816 (to N. Bazan). We thank the Smithsonian Tropical Research Institute for their hospitality and logistical support.

Author contributions

HEF conceived of the study, collected and analyzed data and wrote the paper. MJR conceived of the study, oversaw the project and wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Farris, H.E., & Ryan, M.J. Relative comparisons of call parameters enable auditory grouping in frogs. *Nat. Commun.* **2**:410 doi: 10.1038/ncomms1417 (2011).