J Comp Physiol A
DOI 10.1007/s00359-017-1149-9

ORIGINAL PAPER

Schema vs. primitive perceptual grouping: the relative weighting of sequential vs. spatial cues during an auditory grouping task in frogs

Hamilton E. Farris[1,2,3] · **Michael J. Ryan**[4,5]

Received: 10 August 2016 / Revised: 19 January 2017 / Accepted: 20 January 2017

Abstract Perceptually, grouping sounds based on their sources is critical for communication. This is especially true in túngara frog breeding aggregations, where multiple males produce overlapping calls that consist of an FM 'whine' followed by harmonic bursts called 'chucks'. Phonotactic females use at least two cues to group whines and chucks: whine-chuck spatial separation and sequence. Spatial separation is a primitive cue, whereas sequence is schema-based, as chuck production is morphologically constrained to follow whines, meaning that males cannot produce the components simultaneously. When one cue is available, females perceptually group whines and chucks using relative comparisons: components with the smallest spatial separation or those closest to the natural sequence are more likely grouped. By simultaneously varying the temporal sequence and spatial separation of a single whine and two chucks, this study measured between-cue perceptual weighting during a specific grouping task. Results show that whine-chuck spatial separation is a stronger grouping cue than temporal sequence, as grouping is more likely for stimuli with smaller spatial separation and non-natural sequence than those with larger spatial separation and natural sequence. Compared to the schema-based whine-chuck sequence, we propose that spatial cues have less variance, potentially explaining their preferred use when grouping during directional behavioral responses.

Keywords Cocktail party problem · Auditory stream · Auditory scene analysis · Mate choice · Phonotaxis · *Physalaemus pustulosus*

Introduction

During propagation, sound waves combine and reach receivers as a single combination wave. Thus, for female frogs performing phonotaxis in a multi-source environment (e.g., multi-male chorus), the ability to choose between potential mates requires deconstructing the acoustic mixture, so that its component sounds can be grouped and assigned to their correct sources. Although several cues enable grouping, they can be classified into two general categories: primitive and schema-based cues (Bregman 1990; Bee and Micheyl 2008). Primitive cues are considered to be stimulus driven, in which perceptual groups are based on, for example, similarities in acoustic parameters, such as spectrum, temporal envelope, and spatial location (Darwin and Carlyon 1995; Moore and Gockel 2002, 2012; Winkler et al. 2009). These primitive cues may be associated with any sounds, including arbitrary and experimental ones that do not function in communication. In contrast, schema-based grouping cues are not necessarily based on relationships between particular acoustic parameters, as grouping may depend on matching stimuli to an innate or learned stimulus 'template', such as innate preferences

---

✉ Hamilton E. Farris
hfarri@lsuhsc.edu

[1] Neuroscience Center, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA

[2] Department of Cell Biology and Anatomy, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA

[3] Department of Otorhinolaryngology, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA

[4] Department of Integrative Biology, University of Texas, 1 University Station C0930, Austin, TX 78712, USA

[5] Smithsonian Tropical Research Institute, Balboa, Panama

for particular sequences of communication sounds (Farris and Ryan 2011) or expectations based on experience or knowledge of a stimulus, including melodies and speech (Bregman 1990; Bey and McAdams 2002; Darwin 2008; Devergie et al. 2010; Kidd et al. 2014). In different taxa, the conditions in which one set of grouping cues dominates over the other are not fully understood. Using an animal model with a robust grouping response, this study addressed this gap by presenting stimuli that gave subjects a choice between primitive-based and schema-based perceptual groups. The experiments here use a specific stimulus case to test which cue is more heavily weighted during phonotactic mate choice, a functionally significant behavior in which errors in grouping and source assignment are expected to be costly.

In túngara frogs, males produce complex calls consisting of an FM ‘whine’ followed by 0–7 harmonic bursts called ‘chucks’ (Ryan 1985) (Fig. 1). Farris et al. (2002, 2005) showed that these call components are perceptually distinct. The whine, which can be produced alone, is necessary and sufficient to elicit and direct female phonotaxis (i.e., identity and location decisions, respectively). In contrast, due to laryngeal morphology in males, chucks cannot be produced alone, as they are constrained to follow the production of a whine (Ryan and Drewes 1990; Gridi-Papp et al. 2006). Experimental presentation of a chuck by itself

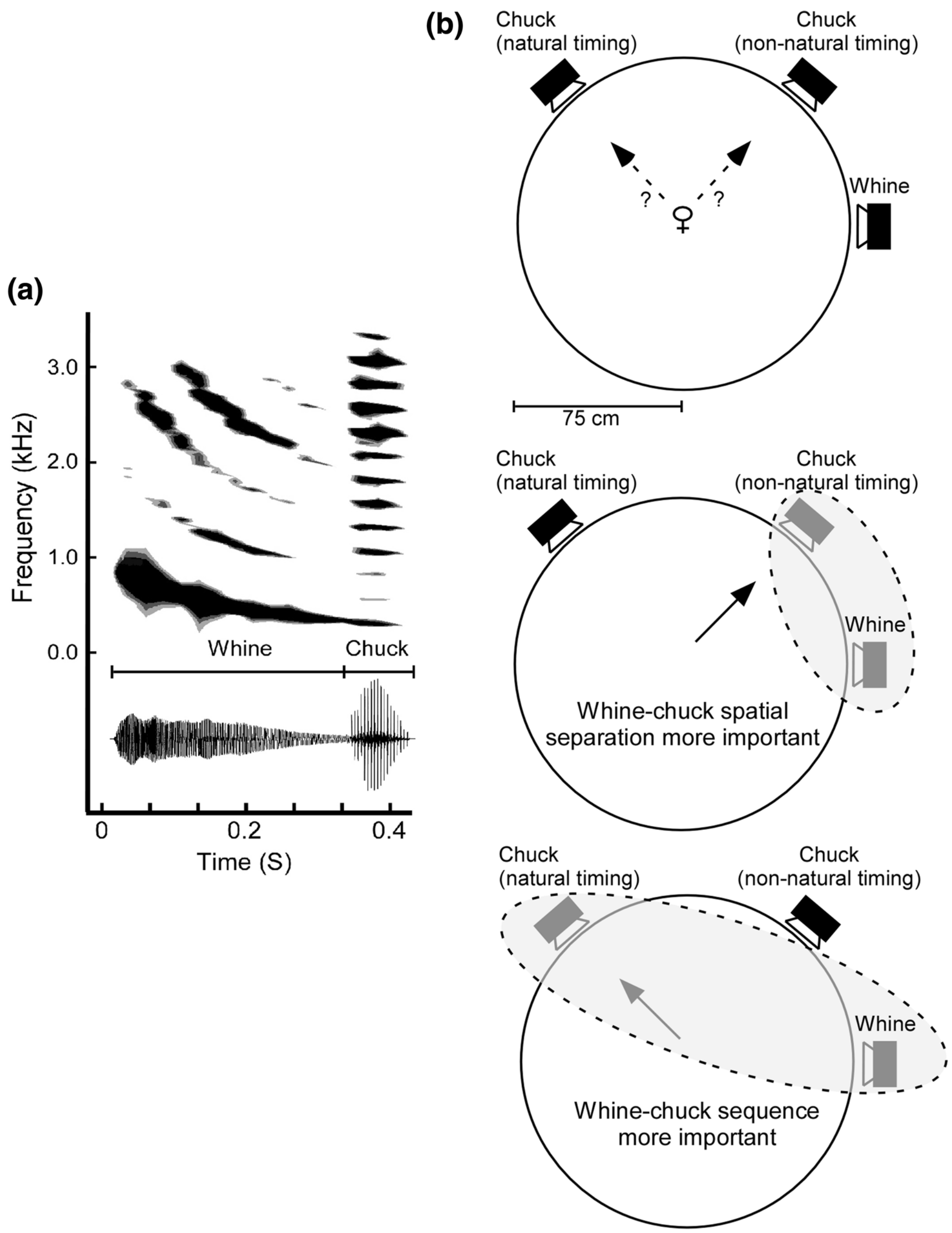

Fig. 1 **a** Spectrogram (*upper panel*) and time waveform of the standard whine and chuck used in this study. The whine and chuck are shown in their natural sequence. **b** Diagram of test arena and hypothetical grouping responses of phonotactic females. *Top panel* females are given a choice for whine-chuck grouping. *Middle panel* Phonotaxis to the chuck with simultaneous onset to the whine, but 45° spatial separation indicates that spatial cues are stronger in grouping decisions. *Bottom panel* phonotaxis to the chuck with natural timing but separated by 135° from the whine indicates that schema-based temporal cues are stronger in grouping decisions. *Shaded areas* represent the null hypotheses for grouping (*arrow* is phonotactic direction) using either of the two cues

does not elicit phonotaxis (i.e., no response or random movement). However, the experimental presentation of a whine with a chuck that is spatially separated elicits robust phonotaxis directed to the chuck, revealing its function in location decisions only. Farris et al. (2002, 2005) concluded that this conditional response to the chuck satisfied the logical criteria for perceptual grouping behavior: different responses were measured for the two components presented alone and in combination. Farris and Ryan (2011) showed that when multiple chucks are presented with a single whine, whine-chuck grouping is based on relative comparisons within primitive and schema-based cues. For example, with respect to a primitive cue, chucks with the smallest spatial separation from the whine are more likely to be grouped. With respect to schema-based grouping, due to the constrained whine-then-chuck sequence, grouping errors may be avoided by employing the sequential schema to not group chucks that start simultaneously with whines. Although use of this schema is not evident when a single whine and chuck are presented (i.e., a single chuck can be grouped with a whine in many whine-chuck sequences), presentation of multiple chucks reveals that chucks closest in time to the natural whine-chuck sequence are often more likely grouped with the whine (Farris and Ryan 2011). These relative comparisons within cues, in which a subset of available sounds are included in perceptual groups based on their relative similarities within one acoustic cue, lead us to test here how stimulus comparisons are made between cues. Specifically, when multiple cues are available, is there evidence for different perceptual weighting of primitive vs. schema-based cues when determining groups?

Materials and methods

Subjects

Phonotactic responses were tested using females collected in amplexus in Gamboa, Panama. All behavioral procedures were licensed and approved by Smithsonian Tropical Research Institute (IACUC permit: 2011-0825-2014-02).

General behavioral procedure

The experimental procedure follows that of our previous work (Farris et al. 2002, 2005; Farris and Ryan 2011). Phonotactic responses were measured to whines and chucks presented alone or in combination with varying spatial separation and temporal sequence. Under infrared illumination, females were placed beneath a mesh cone (10 cm diameter) at the center of a circular (75 cm radius) array of speakers in a light-proof acoustic isolation chamber (2.75 × 1.83 m). The chamber was lined with additional anechoic foam (Sonex, 1.5 inch; NRC 0.8) along the bottom 0.6 m of each wall. Following 3 min of exposure to the stimuli, the cone was removed allowing subjects to move freely, while the stimuli continued. Using an infrared camera, frog position throughout the trial (±5°) was recorded. A successful trial ended when a female exited the circular array within 15 min of release. “No-choice” trials were scored if females (1) failed to leave the 10 cm center circle within 5 min of removing the cone; (2) remained stationary anywhere within the circular array for 2 min; or (3) did not exit the perimeter within 15 min of release. To ensure that “no-choice” trials 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. Twenty-seven females completed testing without consecutive ‘no-choice’ trials. Based on the variance in the direction of phonotactic responses in the previous studies (Farris et al. 2002, 2005; Farris and Ryan 2011), this sample exceeded that necessary (i.e., 20) to enable statistical analysis of the distribution of exit angles from the circular arena. Video recordings of each trial confirmed these angles.

Acoustic stimuli

Recorded from a male at the collection site in Gamboa, Panama, the standard whine and chuck stimuli are those from a call closest to the population mean of 14 acoustic variables (Ryan and Rand 2003). For consistency, this is the same standard call used in the previous studies of whine-chuck grouping (Farris et al. 2002, 2005; Farris and Ryan 2011). Stimulus period was 2 s. Stimuli were amplified (Crown XLS 202) and broadcast from broadband speakers (Radioshack # 40-1040) positioned along the perimeter of the 75 cm radius arc on the floor of the chamber. Prior to each night’s tests, the peak amplitudes of the stimuli were calibrated (±1.0 dB) at the release point at the center of the arena. Chuck amplitude was 6 dB re. whine amplitude [90 dB SPL re. 20 µPa, or the amplitude of a whine at 50 cm, (Ryan 1985)].

Experimental design: spatial vs. sequential grouping cues

Phonotactic responses were measured for the following stimuli. There were four control stimuli that either included the whine or chuck presented alone or a single whine and single chuck, together. The control stimulus conditions were as follows: (1) whine alone; (2) chuck alone; (3) single whine and single chuck with natural sequence but spatially separated by 135°; (4) single whine and single chuck with simultaneous onset (i.e., non-natural sequence) but spatially separated by 45°. These stimuli reconfirm the distinct responses to each component and that there is grouping for a single whine and chuck under these conditions (Farris et al. 2002, 2005). The one experimental stimulus combined the relative chuck parameters for stimuli 3 and 4. Thus, stimulus (5) included a single whine presented with a 45° separated chuck with simultaneous onset, and a 135° separated chuck with natural timing. The hypothesis is that grouping of the former reveals stronger use of spatial cues, whereas grouping of the latter reveals stronger schema-based grouping based on sequence (Fig. 1). It is important to note that although stimulus number 5 tests the hypothesis of whether primitive or schema-based grouping cues are more heavily weighted in whine-chuck grouping, as a single experimental stimulus, it may not test how the two cues compete in other (not tested) competitive conditions.

Statistical analysis

Response distributions of female exit angles were analyzed using three circular statistical tests (Farris et al. 2002, 2005; Farris and Ryan 2011). A Rayleigh test determined whether exit angles were different from a random distribution. Two tests were then used to confirm a grouping response to a chuck. First, a Hotelling test (for paired samples of angles) compared individual responses in each whine-chuck stimulus to that to the whine alone (Zar 1999). This confirmed that the response distribution had shifted direction away from the whine even when chucks are spatially close. Significance levels (α) for the Hotelling test were corrected for the three comparisons using a Bonferroni correction ($\alpha = 0.05/3 = 0.01667$). Second, a *V* test determined whether the shifted responses were, indeed, localized at a particular chuck position (Zar 1999).

Results

Female phonotaxis in response to control stimuli reconfirmed that the whine and chuck are perceptually distinct. A whine alone elicited highly directed phonotaxis in all subjects (Fig. 2a), whereas a chuck alone did not elicit stereotypical phonotaxis (Fig. 2b). That is, for the chuck alone, only 14 of the 27 individuals exited the arena and the exit angles were randomly directed (Fig. 2b; Table 1). For control stimuli in which a single spatially separated chuck was paired with a whine, grouping was demonstrated in both conditions. Even though there is no response to a chuck alone, the presentation of a whine and chuck elicited the conditional response to the chuck: exit angles were significantly shifted from that of the whine alone, while also being significantly localized at the position of the single chuck (Fig. 2c, d; Table 1). These data match those from our previous work (Farris et al. 2002, 2005) showing that females grouping a single whine and chuck are permissive with respect to spatial and temporal cues, grouping a chuck with simultaneous onset to the whine, but separated by 45° (Fig. 2d); and a chuck with natural timing but widely separated at 135° (Fig. 2c). However, these data reconfirm only that grouping is possible under spatially and temporally non-optimal conditions, but do not indicate which cue, if either, is preferentially utilized in grouping. This question is addressed by using a single stimulus that pairs both types of non-optimal chucks with a single whine (i.e., a temporally incorrect but spatially proximal chuck and a spatially distant but temporally correct chuck). Figure 2e shows, for such a stimulus, that female exit angles are significantly localized at the position of the chuck closest to the whine albeit with unnatural simultaneous onset. In contrast, there is no evidence for significant grouping at the naturally timed chuck that was positioned at a 135° whine-chuck spatial separation, even though it was grouped when it was the only chuck presented (Fig. 2e; Table 1).

Discussion

Across taxa, many aspects of acoustic communication are similar, including the acoustic structure of the signals, certain mechanisms of audition, and the acoustic environment in which communication occurs (e.g., multiple sources of sound) (Hauser 1996; Bradbury and Vehrencamp 1998; Gerhardt and Huber 2002). Consequently, we would expect the ability to use and compare different cues during auditory grouping to exhibit similarities in many species. However, because most of our understanding of this ability comes from experiments with humans, it is not known how widespread and similar such processing is and whether similar mechanisms are employed (see Bee and Micheyl 2008; Bee 2012 for review). Thus, our data set based on the robust grouping responses in túngara frogs adds to a limited comparative literature. The present data indicate that whine-chuck grouping can occur in the absence of both spatial and temporal coherence. By presenting these non-optimal stimuli simultaneously in a single stimulus condition, we tested which parameter is more heavily weighted in auditory grouping. Because phonotaxis was localized to the closest, yet unnaturally timed chuck, the responses are consistent with the hypothesis that in this limited case, one cue (primitive: spatial separation) is perceptually more important than another (schema-based: whine-chuck sequence) when grouping these complex communication sounds. While the relative advantage of one grouping cue over another is not unique to túngara frogs (see below), it raises the comparative question of which cue is expected to be preferred in different taxa. More specifically, based on differences in auditory mechanisms of cue processing and /

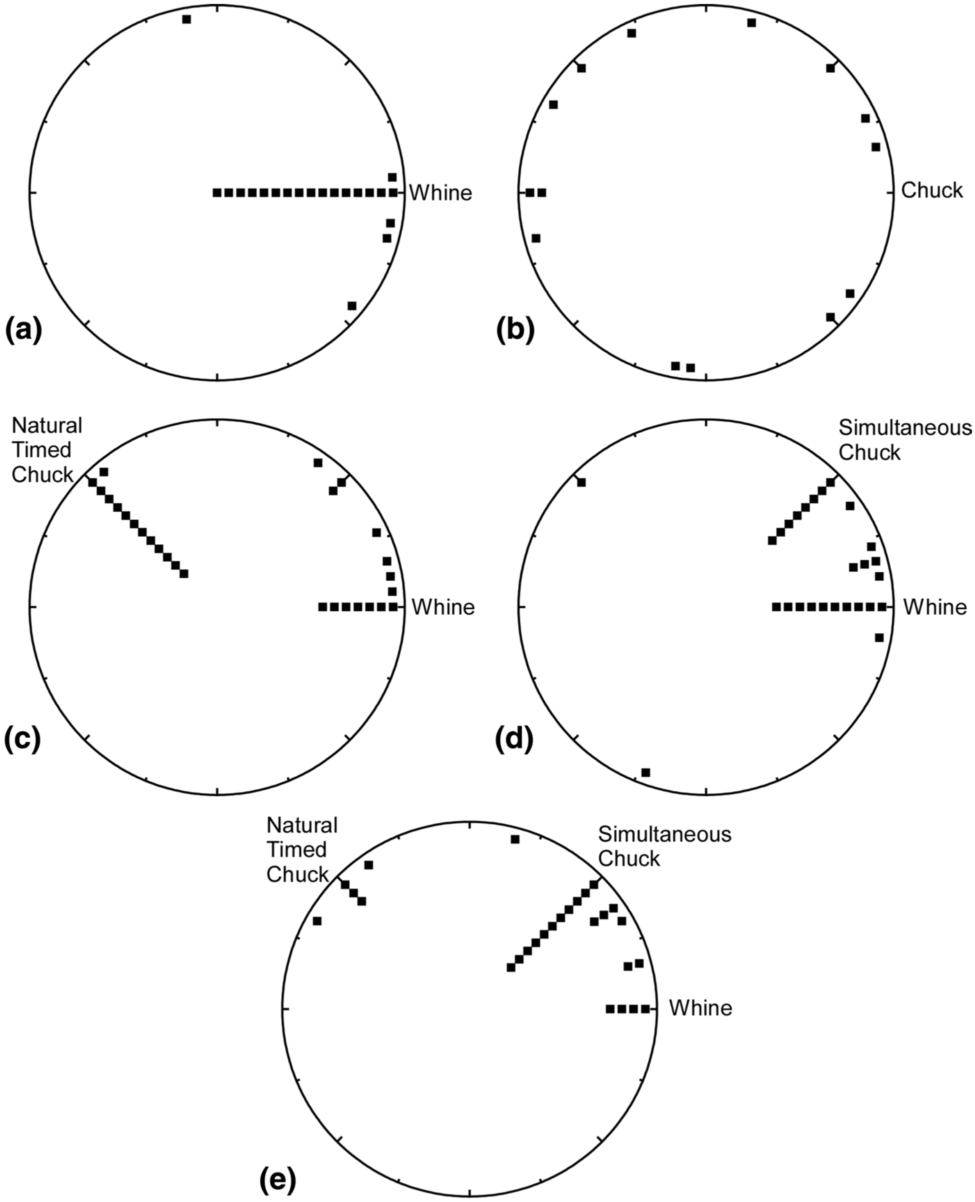

Fig. 2 Squares are the exit angles from the arena indicating the phonotactic responses for individual females. **a** Control broadcast of whine alone elicits directed phonotaxis localized at the speaker. **b** Control broadcast of chuck alone elicits random or no response. **c** Confirmation of grouping response to a single naturally timed chuck at 135° spatial separation. **d** Confirmation of grouping response to a single chuck with simultaneous onset to the whine at 45° spatial separation. **e** Exit angles for stimuli with three call components: whine; simultaneous onset chuck at 45°; natural-timed chuck at 135°. There is a larger grouping response to the 45° separated chuck. See Table 1 for responses distributions and analyses

or the functional contexts in which grouping is performed, should certain taxa (like túngara frogs) be predicted to favor spatial cues over other cues?

As mentioned above, the relatively recent interest in comparative models of perceptual grouping (Bee and Micheyl 2008; Bee 2012, 2015) means experiments measuring potential preference for a particular grouping cue are primarily limited to humans. These experiments often include spatial cues due to their importance in sound localization and subsequent role in signal detection by reducing masking (e.g., binaural masking level difference). Because localization and spatial unmasking are related to sound sorting, the influence of spatial cues on grouping is potentially greater than that for other cues, like frequency or timing relationships. This potential is based on the fact that sounds from the same source commonly have greater variance in their spectral content than in their spatial cues. Such experiments in humans are similar in design to those here, in which incoherent grouping cues (e.g., spatial vs. frequency) are 'competed' against one another to test which determines the sorting of complex sounds (Deutsch 1979), including speech. Data from these experiments show that the evidence for spatial cue use is mixed, however, and depends on the type of competing stimuli (Darwin and Hukin 2000; Darwin 2008; Bremen and Middlebrooks 2013). For example, Culling and Summerfield (1995) found a little evidence that listeners used interaural timing cues when grouping simultaneous sounds in different frequency channels (e.g., vowel identification). In contrast, under stimulus conditions closer to the free-field, ITD

Table 1 Analysis of phonotaxis direction

Stimuli Position and timing Re. whine			*n*	Mean exit angle (±95% CI)	Vector (*r*)	Raleigh test (*P*)	Exit angles vs. whine only response (*P*)	Exit angles localized at Chuck 1 position (*P*)
Whine	Chuck 1	Chuck 2						
Control								
1. –	Chuck 0°		14	312° (–)	0.091	0.893	–	>0.25
2. Whine 0°	–		27	0° (8)	0.946	<0.0001	Null *Ho*	–
3. Whine 0°	Natural-timed Chuck 135°		27	73° (45)	0.48	<0.001	<0.00001	<0.05
4. Whine 0°	Simultaneous Chuck 45°		27	19° (15)	0.829	<0.0001	<0.00351	<0.0005
								Exit angles localized at Chuck 135° position
Space vs. time								
5. Whine 0°	Simultaneous Chuck 45°	Natural-timed Chuck 135°	27	48° (20)	0.751	<0.0001	<0.00001	<0.0005

Columns are the stimulus condition (the presence, position, and timing of the stimuli are indicated in the first three columns); sample size of responding females; mean exit angle; vector strength of the exit angle distribution; Raleigh test for a random distribution; Hotelling test for paired samples of angles which compared individual responses to that for the whine alone; *V* test of whether response distributions are localized at the position of chuck 1 (Zar 1999). Significance in the Hotelling test required Bonferroni alpha correction ($\alpha = 0.05/3 = 0.0167$). Note that the confidence intervals for the random response to the chuck alone are undefined. In addition, only 14 females exited the arena for this stimulus; 13 exhibited no response

and IID (interaural time and interaural intensity difference) do appear to have an effect on segregation (Drennan et al. 2003). For speech, there is large improvement in reception thresholds when spatial cues are available (Bronkhorst 2000). In addition, ITDs may play a role in grouping sequences of sounds if the listener is given preceding sounds which cue direction (Hukin and Darwin 1995), and potentially focus the direction of subsequent attention (Darwin and Hukin 1999). Going forward, however, more tests with túngara frogs are needed to determine if they too show mixed use of cues in various stimulus conditions.

Note that the last hypothesis regarding the use of sound sequence in grouping could be relevant to the use of spatial grouping cues in túngara frogs shown here, as the whine-chuck natural sequence offers the possibility for cueing target direction. That is, the use of binaural cues to localize the whine would then cue the auditory system to favor grouping chucks closer to the whine's location. This certainly could explain grouping by spatial cues when the whine-chuck is presented in the natural sequence. However, this cannot explain grouping of the closest chuck in the present experiments, as the closest chuck had a simultaneous onset to the whine. Note that this temporal coincidence, in which the more likely grouped chuck was simultaneous with the portion of the whine with the highest amplitude, eliminates concerns that masking could have played a role in the grouping response. The results are thus more consistent with the interpretation that there is a relative comparison of chuck spatial separation from the whine, as we have previously shown (intra-cue relative comparisons; Farris and Ryan 2011). Furthermore, the present data indicate that this relative comparison of spatial separation supersedes the comparisons of sequence.

Although we have determined only one condition in which spatial cues are more heavily weighted than other cues in grouping decisions, a preference for spatial cues in grouping would appear to be evolutionarily adaptive during phonotaxis in the multi-source acoustic environment of breeding aggregations. When evaluating communication signals, especially those for mating, proper source assignment is critical to information transfer (Bregman 1990; Fay 2008): errors in source assignment are costly, potentially leading to failures in species recognition, failures in choosing high quality mates, and poor evaluation of competitors (Maynard Smith and Harper 2003). Given such costs, selection would be expected to favor the use of grouping cues with low probabilities of misidentifying the sources of sexual signals. In túngara frogs, the predictability of schema-based temporal cues (whine-chuck sequence) may not be as high as spatial cues. Whereas the relative timing and number of chucks vary between individuals and, more importantly, between calls within individuals (Goutte et al. 2010), there is no variance in whine-chuck spatial separation for calls produced by the same individual. Thus, because of their predictability (Winkler et al. 2009), spatial cues may be more reliable for grouping. An alternative, but not exclusive, hypothesis for spatial cue use in our tests of túngara frogs is based on the behavior it mediates. Phonotaxis is an inherently spatial task, requiring identification of male location. Thus, from a functional point of view, the results here generate the hypothesis that the relative weighting of grouping cues may depend on the behavioral function grouping serves, in this case, directional phonotaxis. Testing such a hypothesis is an example in which use of the comparative approach could elucidate the source of variance in psychophysical phenomenon (Farris and Taylor 2017).

With respect to grouping mechanisms, because there are still few comparative data on grouping, it is not yet clear how interspecific variance in non-neural (e.g., outer ear and conduction structures) and neural mechanisms contribute to the use of one cue more than another. For example, it is possible that non-neural mechanisms for spatial processing specific to certain frogs could contribute to the preferential use of spatial grouping cues. Binaural mechanisms in frogs are quite different from those in humans (Bee and Christensen-Dalsgaard 2016), including the lack of outer ears and the multiple sound paths that allow stimuli to travel through the head, reaching both the inner and outer surfaces of the tympanic membranes. This creates directional sensitivity at the tympanic membrane through the different pressure gradients produced by stimuli from different directions (Rheinlaender et al. 1981). For neural mechanisms, although spatial cue predictability is known to be coded on rapid time scales (milliseconds) in the frog midbrain (Ponnath et al. 2013), use of behavioral tests and free-field stimuli did not enable determination of the particular underlying neural mechanisms employed in comparing spatial and temporal grouping cues. However, one conclusion can be made based on our study organism. The perceptual competition between grouping cues, including between the primitive and schema-based cues presented here, does not require the circuitry of the mammalian cortex, as these decisions were measured in an 'acortical' animal, a frog. It is well known that the vertebrate auditory system processes aspects of frequency, space, and time at the early ascending nodes (including the periphery and brainstem) (Webster et al. 1992; Winer and Schreiner 2005). The results here suggest that these early nodes are capable of processing and comparing grouping cues when making phonotactic decisions. Such complexity in processing, including grouping across frequency channels and comparing grouping cues in different acoustic dimensions, raises a fundamental question about the auditory system: to what extent do higher processing stages that show non-nuclear organization (e.g., mammalian cortex) increase processing capabilities? The túngara frog model system reveals that non-cortical, nuclear based circuitry is sufficient for quite complex tasks, including the comparing of primitive and schema-based grouping cues shown here. This inference adds independent evidence to direct findings of sub-cortical spatial stream segregation (Middlebrooks and Bremen 2013).

Together with the previous work (Farris and Ryan 2011), the behavioral assays here show that grouping decisions in túngara frogs, both within and between cues, appear categorical. That is, all responses were directed to one of the available chucks, with no responses directed toward an 'average' or intermediate position in the arena (Farris and Ryan 2011). Here, for the between-cue comparison, there was also no evidence for intermediate responses (e.g., halfway between the two chucks), as the exit angles were significantly localized at the position of one of the chucks, with most at those closest to the whine. Such 'winner takes all' categorical responses which are consistent with the hypothesis that these two cues, primitive (space) and schema (sequence), are processed separately and that the comparisons are ranked, with the hierarchy favoring the primitive cue. In other words, for the test condition here the cues do not appear to be evaluated together, which would have led to a single decision that considered information in both domains. Going forward, additional experimentation (e.g., more stimulus conditions and electrophysiological assays; Hahne et al. 2002) would be needed to determine the extent to which these two grouping mechanisms are independent of one another.

Understanding of the competition between primitive and schema-based cues when sorting the auditory scene is still of limited experimental interest, especially in comparative models. Although this limitation is likely due to the few model organisms that easily demonstrate this behavior in experimental settings, it is, nevertheless, unfortunate, as recent studies with non-human taxa have shown great promise in addressing fundamental questions in auditory scene analysis that may be difficult to test in humans (Bee and Klump 2004; Bee and Micheyl 2008). For example, data from túngara frogs and other species have addressed the "argument in psychology about whether any schemas can be innate" (Bregman 1990). Indeed, evidence for the grouping preference of the whine-chuck natural sequence shows that schemas need not be learned (Farris and Ryan 2011). Bregman's statement that "we do not know whether we give more weight to the grouping decisions provided by" primitive or schema-based strategies is almost certainly limited to humans, even though it need not be. Our paper addresses this issue, as we show evidence that such 'perceptual competitions' are carried out in other taxa. These results help to begin to build a framework for understanding the fundamental (i.e., necessary and sufficient) mechanisms mediating such perception.

Acknowledgements R. Taylor, R. Rosencrans, and N. Bazan generously provided advice and/or equipment for the project. The manuscript was improved by comments from two reviewers. HEF was supported by NIH Grant P20RR016816 (N. Bazan, PI). MJR was supported by NSF Grant (IOS 1120031; R. Taylor and R. Page Co-PIs). All behavioral procedures were licensed and approved by Smithsonian Tropical Research Institute (IACUC permit: 2011-0825-2014-02).

Compliance with ethical standards

All behavioral procedures were licensed and approved by Smithsonian Tropical Research Institute (IACUC permit: 2011-0825-2014-02).

Conflict of interest We have no competing interests.

References

Bee MA (2012) Sound source perception in anuran amphibians. Curr Opin Neurobiol 22(2):301–310

Bee MA (2015) Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. Int J Psychophysiol 95(2):216–237

Bee MA, Christensen-Dalsgaard J (2016) Sound source localization and segregation with internally coupled ears: the treefrog model. Biol Cybern 110(4–5):271–290

Bee MA, Klump GM (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. J Neurophysiol 92(2):1088–1104

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122(3):235–251

Bey C, McAdams S (2002) Schema-based processing in auditory scene analysis. Percept Psychophys 64(5):844–854

Bradbury JW, Vehrencamp SL (1998). Principles of animal communication. Sinauer Assoc. Inc., Sunderland

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

Bremen P, Middlebrooks JC (2013) Weighting of spatial and spectro-temporal cues for auditory scene analysis by human listeners. PLoS One 8(3):e59815

Bronkhorst AW (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. Acta Acust United Ac 86:117–128

Culling JF, Summerfield Q (1995) Perceptual separation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay. J Acoust Soc Am 98(2 Pt 1):785–797

Darwin CJ (2008) Listening to speech in the presence of other sounds. Philos Trans R Soc Lond B Biol Sci 363(1493):1011–1021

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

Darwin CJ, Hukin RW (1999) Auditory objects of attention: the role of interaural time differences. J Exp Psychol Hum Percept Perform 25(3):617–629

Darwin CJ, Hukin RW (2000) Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. J Acoust Soc Am 107(2):970–977

Deutsch D (1979) Binaural integration of melodic patterns. Percept Psychophys 25(5):399–405

Devergie A, Grimault N, Tillmann B, Berthommier F (2010) Effect of rhythmic attention on the segregation of interleaved melodies. J Acoust Soc Am 128(1):EL1–EL7

Drennan WR, Gatehouse S, Lever C (2003) Perceptual segregation of competing speech sounds: the role of spatial location. J Acoust Soc Am 114(4 Pt 1):2178–2189

Farris HE, Ryan MJ (2011) Relative comparisons of call parameters enable auditory grouping in frogs. Nat Commun. doi:10.1038NCOMMS1417

Farris HE, Taylor RC (2017) Mate searching animals as model systems for understanding perceptual grouping. In: Bee MA, Miller CT (eds) Psychological mechanisms in animal communication. Springer, New York, pp 89–118

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(3):181–188

Farris HE, Rand AS, Ryan MJ (2005) The effects of time, space and spectrum on auditory grouping in túngara frogs. J Comp Physiol A 191(12):1173–1183

Fay RR (2008) Sound source perception and stream segregation in nonhuman vertebrate Animals. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer, New York, pp 307–323

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans, University of Chicago, Chicago

Goutte S, Kime NM, Argo TF, Ryan MJ (2010) Calling strategies of male túngara frogs in response to dynamic playback. Behaviour 147(1):65–83

Gridi-Papp M, Rand AS, Ryan MJ (2006) Animal communication: complex call production in the túngara frog. Nature 441(7089):38

Hahne A, Schroger E, Friederici AD (2002) Segregating early physical and syntactic processes in auditory sentence comprehension. Neuroreport 13(3):305–309

Hauser MD (1996) The evolution of communication. MIT, Cambridge

Hukin RW, Darwin CJ (1995) Effects of contralateral presentation and of interaural time differences in segregating a harmonic from a vowel. J Acoust Soc Am 98:1380–1387

Kidd G Jr, Mason CR, Best V (2014) The role of syntax in maintaining the integrity of streams of speech. J Acoust Soc Am 135(2):766–777

Maynard Smith J, Harper DGC (2003) Animal signals. Oxford University Press, Oxford

Middlebrooks JC, Bremen P (2013) Spatial stream segregation by auditory cortical neurons. J Neurosci 33(27):10986–11001

Moore BCJ, Gockel H (2002) Factors influencing sequential stream segregation. Acta Acust United Ac 88(3):320–333

Moore BC, Gockel HE (2012) Properties of auditory stream formation. Philos Trans R Soc Lond B Biol Sci 367(1591):919–931

Ponnath A, Hoke KL, Farris HE (2013) Stimulus change detection in phasic auditory units in the frog midbrain: frequency and ear specific adaptation. J Comp Physiol A 199(4):295–313

Rheinlaender J, Walkowiak W, Gerhardt HC (1981) Directional hearing in the green treefrog: a variable mechansim? Naturwissenschaften 67:430–431

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 Lond 40(1):37–52

Ryan MJ, Rand AS (2003) Sexual selection in female perceptual space: how female túngara frogs perceive and respond to complex population variation in acoustic mating signals. Evol Int J org Evol 57(11):2608–2618

Webster DB, Fay RR, Popper AN (1992) The evolutionary biology of hearing, Springer, New York.

Winer JA, Schreiner CE (2005). The inferior colliculus, Springer, New York

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 13(12):532–540

Zar JH (1999) Biostatistical analysis. Prentice-Hall, Inc., Upper Saddle River