Current Biology
Report

Frogs Exploit Statistical Regularities in Noisy Acoustic Scenes to Solve Cocktail-Party-like Problems

Norman Lee,[1,7,8,*] Jessica L. Ward,[1,2] Alejandro Vélez,[1,5] Christophe Micheyl,[3,6] and Mark A. Bee[1,4]
[1]Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN 55108, USA
[2]Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, Saint Paul, MN 55108, USA
[3]Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA
[4]Graduate Program in Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA
[5]Present address: Department of Biology, San Francisco State University, San Francisco, CA 94132, USA
[6]Present address: Starkey Hearing Research Center, Berkeley, CA 94704, USA
[7]Present address: Biology Department, St. Olaf College, Northfield, MN 55057, USA
[8]Lead Contact
*Correspondence: lee33@stolaf.edu
http://dx.doi.org/10.1016/j.cub.2017.01.031

SUMMARY

Noise is a ubiquitous source of errors in all forms of communication [1]. Noise-induced errors in speech communication, for example, make it difficult for humans to converse in noisy social settings, a challenge aptly named the "cocktail party problem" [2]. Many nonhuman animals also communicate acoustically in noisy social groups and thus face biologically analogous problems [3]. However, we know little about how the perceptual systems of receivers are evolutionarily adapted to avoid the costs of noise-induced errors in communication. In this study of Cope's gray treefrog (*Hyla chrysoscelis*; Hylidae), we investigated whether receivers exploit a potential statistical regularity present in noisy acoustic scenes to reduce errors in signal recognition and discrimination. We developed an anatomical/physiological model of the peripheral auditory system to show that temporal correlation in amplitude fluctuations across the frequency spectrum ("comodulation") [4–6] is a feature of the noise generated by large breeding choruses of sexually advertising males. In four psychophysical experiments, we investigated whether females exploit comodulation in background noise to mitigate noise-induced errors in evolutionarily critical mate-choice decisions. Subjects experienced fewer errors in recognizing conspecific calls and in selecting the calls of high-quality mates in the presence of simulated chorus noise that was comodulated. These data show unequivocally, and for the first time, that exploiting statistical regularities present in noisy acoustic scenes is an important biological strategy for solving cocktail-party-like problems in nonhuman animal communication.

RESULTS

In biological systems, noise-induced errors can impose dire fitness consequences for signalers and receivers [7, 8]. Such errors select for the optimization of signal structures [9, 10], signaling strategies [11, 12], and the sensory, perceptual, and cognitive mechanisms for processing signals [9, 10]. The potential for errors in communication is greatest when receivers must respond to signals produced in a complex milieu of competing signals having similar physical properties. Such mixtures of signals constitute significant sources of noise for many animals. The raucous acoustic scenes associated with large groups of conspecifics, such as a human cocktail party [2], a communal songbird roost [6], or choruses of insects [13] and frogs [14], represent social environments where perceptual adaptations for coping with such noise would be particularly advantageous. An emerging view in sensory ecology is that auditory systems, much like visual systems [15, 16], are evolutionarily optimized to process statistical regularities present in natural scenes [4, 17–19]. One statistical regularity of many natural acoustic scenes, like those characteristic of noisy social gatherings, derives from the physical properties of natural sounds, which exhibit slow fluctuations in amplitude through time [19, 20]. In many instances, these fluctuations are correlated across the frequency spectrum (i.e., comodulated) [4–6]. The extent to which comodulation in noisy acoustics scenes is exploited to solve complex communication problems has been a contentious issue in studies of human speech communication [21–23] but remains largely untested in other animals [6, 24].

Natural Statistics of a Noisy Acoustic Scene

Males of Cope's gray treefrog form dense choruses in which they produce loud, pulsatile advertisement calls to attract females (Figure 1). Individual calls (Figure 1A; Audio S1) are produced at high sound pressure levels (SPLs) reaching 85 to 90 dB at 1 m [25], and sustained noise levels in choruses commonly range between 70 and 80 dB SPL. Chorus noise has a frequency spectrum matching that of the call (Figure 1B; Audio S2) and exhibits slow fluctuations in amplitude [26]. Previous studies of auditory

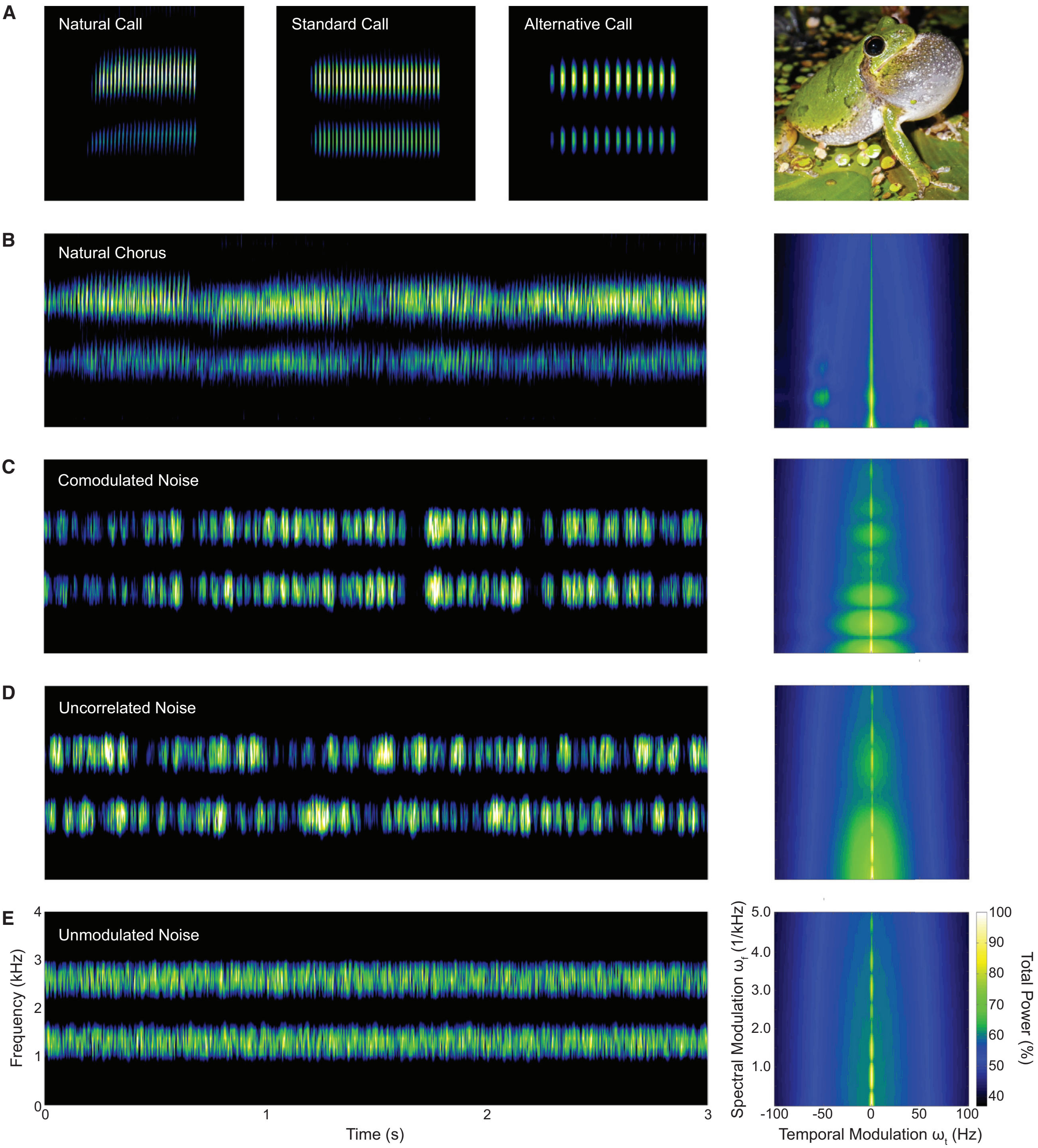

Figure 1. Depictions of Natural and Synthetic Signals and Noise

(A) Spectrograms of a natural advertisement call (left; Audio S1), the synthetic standard call (center), and one of several different synthetic alternative calls (right) used in this study. The depiction of a natural call illustrates two acoustic properties mimicked by synthetic stimulus calls: this signal's pulsatile structure and bimodal frequency spectrum, with spectral peaks near 1.3 and 2.6 kHz. The synthetic standard call was used as a stimulus in experiments 1–4. The alternative call depicted here was used in two-alternative choice tests in experiment 3 and differs from the standard call (50 pulses/s) in having a slower pulse rate (20 pulses/s). Oscillograms of standard and alternative calls used to create differences in call effort in experiment 4 are illustrated in Figure S1A. Also shown here (far right) is a photograph of a calling male of Cope's gray treefrog, used with permission from J.C. Tanner.

(B) An illustrative spectrogram (left) and the mean modulation power spectrum (right) of the natural noise generated by choruses of Cope's gray treefrog (Audio S2). The spectrogram of the natural chorus illustrates the two spectral bands of background noise in choruses arising from the mixture of vocalizations produced

(legend continued on next page)

masking in hylid treefrogs have shown that chorus noise reduces signal active space [27–30], impairs species discrimination [31, 32], interferes with sound localization [33], and constrains choices of preferred mates [29, 34]. Thus, females face an evolutionarily significant "cocktail-party-like problem" in choruses because the large number of signalers in a chorus creates noisy listening conditions that induce costly communication errors.

We tested the hypothesis that chorus noise is comodulated across frequencies of biological relevance to the frogs themselves. To do so, we passed acoustic recordings of Cope's gray treefrog choruses through an anatomical/physiological model of the species' peripheral auditory system and quantified the degree of comodulation across the frequency spectrum (see Supplemental Experimental Procedures). Acoustic recordings were made from within choruses on nights and at times of high calling activity during the species's breeding season. Our model implemented a bank of gammatone filters (Figure 2) to simulate spectral processing by the two sensory papillae in the frog inner ear that are sensitive to airborne sounds [36]. The tonotopically organized amphibian papilla was modeled as six adjacent filters with center frequencies between 238 Hz and 1.3 kHz. The basilar papilla was modeled as a single filter centered at 2.6 kHz. Thus, two of the filters (1.3 and 2.6 kHz) were centered on the spectral peaks emphasized in conspecific calls (Figure 1A). Filter bandwidths were determined from a meta-analysis of published tuning curves measured electrophysiologically from frog auditory nerve fibers (Figure 2A; see Supplemental Experimental Procedures). Adjacent filters simulating the amphibian papilla were spaced to overlap at frequencies 3 dB above threshold. The relative gain of the center frequency of each filter was adjusted to match the corresponding frequency from the species' audiogram [37]. To simulate the half-wave rectification reflected in the coding of amplitude-modulated sounds in the auditory nerve [38], we analytically determined the Hilbert envelope of the output of each frequency filter. Each Hilbert envelope thus preserved the temporal modulations in amplitude present only in the corresponding range of acoustic frequencies passing through each specified filter.

Comodulation was evaluated by computing the cross-covariance between Hilbert envelopes across all pairwise combinations of frequency filters (Figures 2 and S2A). This cross-covariance procedure allowed us to assess the extent to which sound amplitude in different regions of the frequency spectrum varied together through time (i.e., increasing and decreasing together on a moment-to-moment basis). Consistent with our hypothesis, mean cross-covariance values always significantly exceeded null expectations and were highest for frequencies emphasized in conspecific calls (Figure S2B). The mean value comparing output from the two filters centered on 1.3 and 2.6 kHz was 26.7 standard deviations greater than expected by chance (Figures 2B and S2). Our anatomically and physiologically inspired analyses of natural chorus noise, therefore, confirmed that comodulation is a prominent statistical regularity present in the acoustic scenes of Cope's gray treefrog choruses.

We next tested the hypothesis that receivers exploit comodulation in background noise to improve listening performance in ecologically relevant communication tasks. In four psychophysical experiments, we evaluated female mating decisions by quantifying phonotaxis in response to synthetic advertisement calls (Figure 1A) in the presence of artificial "chorus-shaped" noises (Figures 1C–1E). Each experimental noise was constructed by adding two narrow-band noises centered on the two spectral peaks present in Cope's gray treefrog calls (1.3 and 2.6 kHz). The temporal envelopes of the two noise bands were manipulated so that they were (1) comodulated (Figure 1C; Audio S3), (2) modulated but uncorrelated (Figure 1D; Audio S4), or (3) unmodulated (Figure 1E; Audio S5) (see Supplemental Experimental Procedures). A control experiment confirmed that these noises were behaviorally neutral and did not, by themselves, influence phonotaxis (Figure S3). If comodulation in ambient background noise can be exploited to mitigate noise-induced communication errors, we expected to find relatively better performance in comodulated conditions.

Signal Recognition in Comodulated Noise

In two experiments we estimated "signal recognition thresholds" (SRTs) [30], which are conceptually analogous to the "speech reception threshold" measured in studies of masked speech perception in humans [22]. Compared with quiet, the presence of all three experimental noises introduced errors that were functionally equivalent to missed mating opportunities. That is, in the presence of noise, subjects failed to respond to an attractive signal (the "standard call" in Figure 1A) when it was presented at sound levels that were nevertheless sufficiently high to elicit phonotaxis in quiet. Consistent with our hypothesis, however, subjects responded to signals at lower thresholds in comodulated noise compared to other noise conditions.

In experiment 1 (Figures 3A and S1B), we presented signals at each of five signal-to-noise ratios (SNRs; −12, −6, 0, +6, and +12 dB). The proportion of subjects responding varied significantly as a function of SNR ($\chi^2 = 93.8$, degrees of freedom [df] = 1, $p < 0.001$) and noise condition ($\chi^2 = 7.0$, df = 2, $p = 0.031$) and was significantly higher in the comodulated condition compared with both the uncorrelated ($p = 0.040$) and unmodulated ($p = 0.011$) conditions. We determined SRTs as the lowest signal level at which fitted response proportions exceeded 0.5. In the comodulated condition, SRTs were 2.7 dB and 3.7 dB lower than in the uncorrelated and unmodulated conditions, respectively (Figure 3A).

---

by calling males. The mean modulation power spectrum [20] illustrates the prominence of temporal fluctuations in amplitude (x axis) occurring at slow rates (e.g., <5–10 Hz). The mean depicted here was determined from an ensemble of 26 chorus recordings (see Supplemental Experimental Procedures). Each recording was truncated into 90 1-s segments, and a Gaussian spectrogram (Gaussian window bandwidth: 32 Hz; window size: 1,316) was computed for each segment. A 2D FFT was computed for each Gaussian spectrogram, and real values were averaged across all segments from all recordings to give an average modulation power spectrum.

(C–E) Spectrograms (left) and modulation power spectra (right) of the three artificial chorus-shaped noises used in experiments 1–4: the comodulated noise (C) (Audio S3), the uncorrelated noise (D) (Audio S4), and the unmodulated noise (E) (Audio S5). During each behavioral test of a subject, a specified noise was broadcast continuously to simulate the ambient background noise of a chorus while one or more specified signals were broadcast periodically to simulate individual calling males.

Additional details on the speaker arrangements used in behavioral experiments are provided in Figure S1B and in the Supplemental Experimental Procedures.

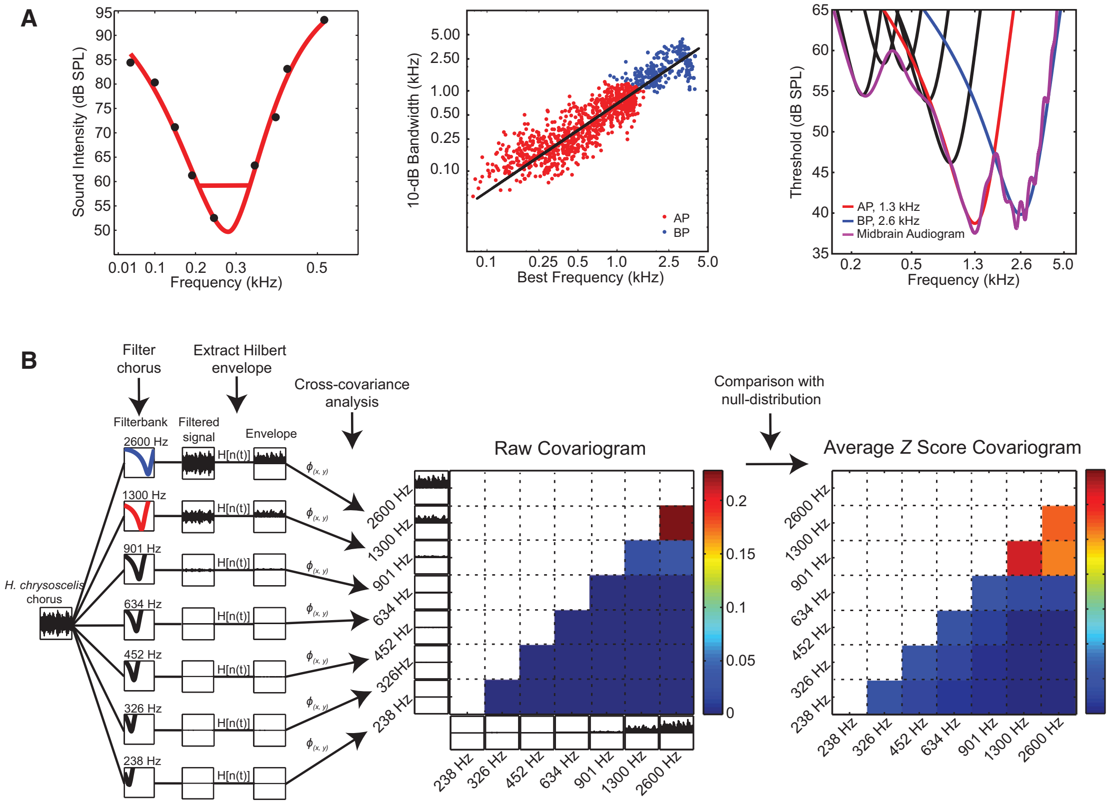

Figure 2. Biologically Inspired Analyses of Chorus Noise Reveal Significant Comodulation

An anatomical/physiological model was used to determine the degree of comodulation present in natural chorus sounds.

(A) The model consisted of a bank of auditory filters fitted to the audiogram of Cope's gray treefrog. Left: each filter was modeled using parameters from VIIIth nerve frequency tuning curves (FTCs) measured in previous studies of frogs (see Supplemental Experimental Procedures). An example of a previously published VIIIth nerve FTC from a frog [35] showing the rounded-exponential function (red curve) used to determine its best frequency (BF), threshold, and bandwidth (BW) 10 dB above threshold (10-dB BW). Center: scatterplot showing the positive relationship between 10-dB BW and BF obtained from a meta-analysis of 1,071 FTCs from seven species of frogs across ten different published studies (see Supplemental Experimental Procedures). Units are classified as innervating either the amphibian papilla (shown in red) or the basilar papilla (shown in blue). Right: diagram showing the model auditory filterbank, with the gain of each filter adjusted to the sensitivity of the midbrain audiogram (purple curve). Filters centered on the 1.3 and 2.6 kHz peaks of the advertisement call are shown in red and blue, respectively.

(B) Cross-covariance analyses were conducted to quantify the magnitude of comodulation in chorus noise. Left: chorus recordings (n = 26 choruses) of 1.5-min duration were filtered using the model filterbank depicted in the right panel of (A). Center: pairwise comparisons between the Hilbert envelope of the output of each frequency filter were made using cross-covariance, as illustrated here by the raw covariogram for a representative chorus recording. Below the diagonal in the raw covariogram shows peak cross-covariance magnitudes for different envelope comparisons plotted as a heatmap. Right: the mean *Z* score covariogram depicts the mean cross-covariance values, averaged across all 26 chorus recordings, as *Z* scores relative to null distributions based on comparing the envelopes of frequency filter outputs across different choruses (Figure S2; Supplemental Experimental Procedures). Colors indicate the number of standard deviations beyond the mean of the null distribution. The high degree of comodulation revealed by these analyses could not be explained as merely resulting from overlap between adjacent auditory filters in the model (Supplemental Experimental Procedures).

In experiment 2 (Figures 3B and S1B), we estimated SRTs using an adaptive tracking procedure to determine the lowest SNR that reliably elicited phonotaxis. SRTs varied significantly across the three noise conditions ($F_{2,57} = 23.9$, $p < 0.001$) and were elevated compared to thresholds measured in quiet. The mean threshold was significantly lower in the comodulated condition compared with both the uncorrelated and unmodulated conditions (Figure 3B). In the comodulated condition, subjects experienced, on average (±SEM), 2.6 ± 1.0 dB and 6.9 ± 4.0 dB of masking release compared with the uncorrelated and unmodulated conditions, respectively. Thresholds were also significantly lower in the uncorrelated condition compared with the unmodulated condition (Figure 3B).

Signal Discrimination in Comodulated Noise

Two additional experiments investigated signal discrimination in the contexts of species discrimination (experiment 3) and sexual selection (experiment 4) using two-alternative choice tests. The presence of noise introduced errors in discrimination performance compared with quiet. Consistent with our hypothesis, however, subjects made fewer discrimination errors in comodulated noise.

In experiment 3, subjects chose between the standard call and an alternative call differing in pulse rate (Figure 1A), which is the primary cue females use to discriminate between conspecific males (40–65 pulses/s) and males of a morphologically indistinguishable sister species, *Hyla versicolor* (17–35 pulses/s) [39]. In quiet, subjects preferentially selected calls with conspecific

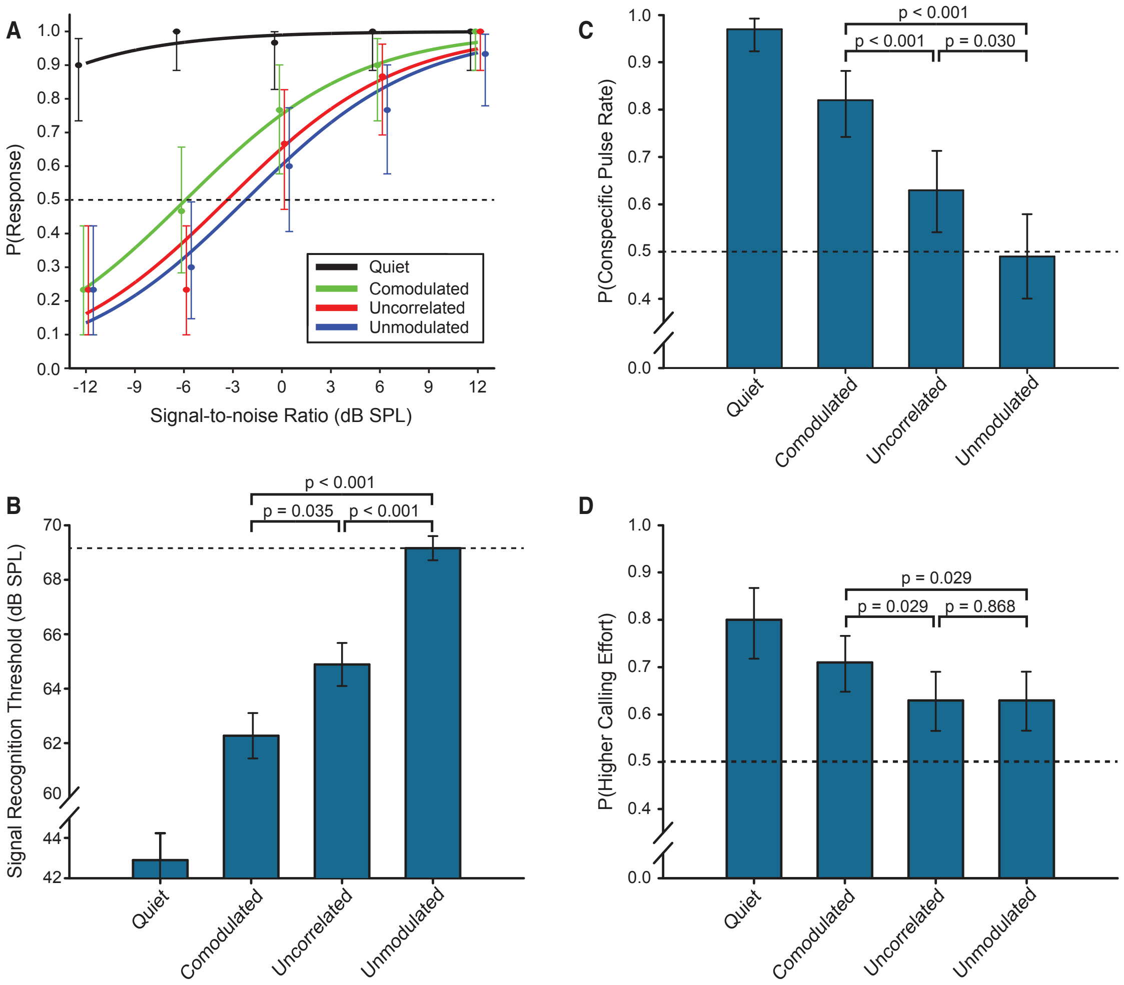

Figure 3. Comodulated Noise Improves Performance in Several Key Communication Tasks Relative to Uncorrelated Noise and Unmodulated Noise

The three artificial chorus-shaped noises used in experiments 1–4 were behaviorally neutral and did not, by themselves, influence phonotaxis (Figure S3).

(A and B) Experiments 1 (A) and 2 (B) consisted of single-stimulus (no-choice) tests and revealed lower signal recognition thresholds (SRTs) in comodulated noise.

(A) Points depict the proportion (±95% exact binomial confidence intervals) of subjects responding at each of five signal-to-noise ratios (−12, −6, 0, +6, and +12 dB, or equivalent signal levels in quiet); solid lines represent fitted functions from generalized estimating equations. The horizontal dashed line represents the criterion (0.5) for determining SRTs.

(B) Bars depict the mean (±SEM) SRTs determined using an adaptive tracking procedure. The horizontal dashed line in (B) indicates the level of performance relative to the condition with the highest threshold.

(C and D) Experiments 3 (C) and 4 (D) consisted of two-alternative choice tests and revealed better discrimination of sound patterns in comodulated noise.

(C) Bars depict the proportion (±95% exact binomial confidence intervals) of subjects choosing stimuli with conspecific pulse rates (P(Conspecific Pulse Rate)). Horizontal dashed line in (C) depicts the level of performance expected by chance (0.5) in a two-alternative choice test.

(D) Bars depict the proportions (±95% exact binomial confidence intervals) of subjects choosing stimuli with relatively higher calling efforts (P(Higher Calling Effort)). Horizontal dashed line in (D) depicts the level of performance expected by chance (0.5) in a two-alternative choice test.

pulse rates (all $p < 0.001$; Figure 3C). Compared with quiet, noise-induced errors in mating decisions were reflected in a reduction in the proportion of subjects selecting calls with conspecific pulse rates. However, error rates differed significantly across noise conditions ($\chi^2 = 26.2$, df = 2, $p < 0.001$). Subjects were significantly more likely to correctly select a conspecific pulse rate in the comodulated condition compared with both the uncorrelated and unmodulated conditions (Figure 3C). They were also better at doing so in the uncorrelated condition compared with the unmodulated condition.

In experiment 4, subjects chose between the standard call (Figure 1A) and an alternative with either a higher or a lower "calling effort," an acoustic property that is a joint function of call rate and call duration (Figure S1A). As in many other animals [40], females of Cope's gray treefrogs prefer males that produce more energetically costly, "high effort" signals [41]. Experiment 4 thus simulated an intraspecific mate choice between two males differing in the quality of their sexual displays. In quiet, females preferentially selected stimuli with relatively higher calling efforts (all $p < 0.001$; Figure 3D). In the presence of noise, the proportions of subjects choosing stimuli with relatively higher calling efforts were reduced but varied significantly as a function of noise condition ($\chi^2 = 6.0$, df = 2, $p = 0.049$). Subjects were significantly more likely to correctly choose a simulated caller with a relatively higher calling effort in the comodulated condition compared with both the uncorrelated and unmodulated conditions (Figure 3D). Responses in the uncorrelated and unmodulated conditions were not different.

DISCUSSION

Our results provide strong evidence that receivers can take advantage of statistical regularities in noisy acoustic scenes to solve cocktail-party-like communication problems. An anatomical/physiological model of frog auditory processing revealed that comodulation across biologically relevant frequencies is a prominent statistical regularity in the noise generated by large social aggregations. Behavioral experiments demonstrated that, while noise induces errors in the mating decisions of females compared with those made in quiet, these errors are generally reduced in modulated noise (compared with unmodulated noise) and are further reduced in comodulated noise (compared with modulated but uncorrelated noise). Moreover, improved performance in comodulated noise was not context specific. It extended to recognition of signals at lower SNRs, as well as to better discrimination of sound patterns that mediate species discrimination and assessments of display quality in intraspecific mate choice. These behavioral contexts represent fundamental communication tasks of critical evolutionary importance for frogs and many other animals. Results from this study, therefore, substantially extend earlier work on humans [5, 21–23], frogs [42], and other animals [6, 43, 44] by showing that comodulation is a statistic of natural acoustic scenes that can be exploited to mitigate costly, noise-induced errors in communication.

Our findings cast significant new light on neglected features of environmental noise that are likely of biological relevance in many animal communication systems. Studies of acoustic communication, for example, typically consider the source of noise (e.g., biotic, abiotic, or anthropogenic), as well as its average frequency spectrum, its average amplitude, and how these two static features impact decisions made by signalers [7, 8]. In stark contrast, statistical regularities that describe the dynamic nature of noisy acoustic scenes and how these regularities are exploited by receivers to improve signal reception, as investigated in the present study, have received almost no attention [6, 24]. Given the central importance of noise as a source of selection on animal communication systems [1, 7, 8], we should not be surprised to find in future studies that many animals, similar to humans in some contexts [18], are exquisitely sensitive to the statistical regularities present in natural acoustic scenes. Future empirical studies of natural scene statistics are thus needed to fully and more accurately assess the impacts of noise on the evolution of animal communication systems. More broadly, results from this study have important ramifications for the influential "receiver psychology" paradigm [45], which emphasizes that psychological mechanisms for signal detection and discrimination are potent sources of selection on signal evolution. The prevalence of perceptual adaptations for processing statistical regularities in natural acoustic scenes and how they may ultimately impact the structure of signals remain important and unexplored questions about receiver psychology and signal evolution.

Findings from this study also have implications beyond animal communication. In many frogs, such as Cope's gray treefrog, the separate high-frequency and low-frequency spectral peaks present in signals and noise are transduced by physically distinct sensory papillae in the inner ear [36], potentially after taking different biophysical routes to the middle ear [46]. Thus, the frog auditory system processes comodulation not only across frequency channels but also across separate peripheral end organs and physical transmission pathways. Hence, the biophysical and neurosensory mechanisms frogs use to exploit comodulated noise are potentially distinct from those operating in other vertebrates. Efforts to discover the biological strategies used by a diversity of species to extract communication signals from noise have potential to uncover evolutionarily novel mechanisms that might be harnessed to improve hearing prosthetics and speech recognition systems. Compared to people with healthy auditory systems, people with hearing loss experience added difficulty communicating in a crowd, and hearing aids and cochlear implants provide limited benefits in such environments [47]. Computer algorithms for automated speech recognition also yield higher error rates in the presence of noise generated by competing speech [48]. However, evolution by natural selection has solved biological analogs of the human cocktail party problem numerous times [3, 8]. Moreover, the sense of hearing had multiple evolutionary origins [49], and even within vertebrates, key auditory mechanisms have arisen multiple times independently and differ among lineages [50]. Consequently, there is almost certainly diversity in evolved solutions to cocktail-party-like problems. Deeper knowledge of the potential diversity in hearing mechanisms could shed light on how evolution has attempted to solve complex communication problems that continue to challenge biomedical and computer engineers.

In summary, our data indicate that exploiting statistical regularities in natural acoustic scenes may be a common signal-processing strategy that has evolved to mitigate noise-induced errors in animal communication. Yet, the mechanisms underlying this strategy may differ across species given the fascinating evolutionary history of hearing. Detailed investigation into these strategies at the perceptual, biophysical, and neuronal levels in a diversity of animals not only will deepen our understanding of the mechanisms and evolution of animal communication, but could ultimately help to improve human health and technology.

ACCESSION NUMBERS

The raw data and MATLAB scripts for the model of the frog auditory periphery have been deposited in the University of Minnesota Digital Conservancy and can be accessed through the following DOI: https://doi.org/10.13020/D6ZP4H.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and five audio clips and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.01.031.

AUTHOR CONTRIBUTIONS

N.L., J.L.W., A.V., and M.A.B. designed experiments and collected psychophysical data. N.L., C.M., and M.A.B. designed the model of the auditory periphery. N.L. and M.A.B. wrote the paper. All authors discussed results and commented on the manuscript.

ACKNOWLEDGMENTS

We thank C. Miller, K. Schrode, J. Tumulty, J. Tanner, and three anonymous reviewers for feedback on earlier versions of the manuscript. This research was funded by a grant from the National Institute on Deafness and Other Communication Disorders (R01 DC 009582) to M.A.B. Animals were treated according to protocols reviewed and approved by the University of Minnesota's Institutional Animal Care and Use Committee (1202A10178 and 1401-31258A).

Received: November 8, 2016
Revised: January 13, 2017
Accepted: January 18, 2017
Published: February 23, 2017

REFERENCES

1. Wiley, R.H. (2015). Noise Matters: The Evolution of Communication (Harvard University Press).
2. McDermott, J.H. (2009). The cocktail party problem. Curr. Biol. *19*, R1024–R1027.
3. Bee, M.A., and Micheyl, C. (2008). The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J. Comp. Psychol. *122*, 235–251.
4. Nelken, I., Rotman, Y., and Bar Yosef, O. (1999). Responses of auditory-cortex neurons to structural features of natural sounds. Nature *397*, 154–157.
5. Verhey, J.L., Pressnitzer, D., and Winter, I.M. (2003). The psychophysics and physiology of comodulation masking release. Exp. Brain Res. *153*, 405–417.
6. Klump, G.M. (2016). Perceptual and neural mechanisms of auditory scene analysis in the European starling. In Psychological Mechanisms in Animal Communication, M.A. Bee, and C.T. Miller, eds. (Springer), pp. 57–88.
7. Brumm, H., and Slabbekoorn, H. (2005). Acoustic communication in noise. Adv. Stud. Behav. *35*, 151–209.
8. Brumm, H., ed. (2013). Animal Communication and Noise (Springer).
9. Endler, J.A. (1993). Some general comments on the evolution and design of animal communication systems. Philos. Trans. R. Soc. Lond. B Biol. Sci. *340*, 215–225.
10. Feng, A.S., Narins, P.M., Xu, C.H., Lin, W.Y., Yu, Z.L., Qiu, Q., Xu, Z.M., and Shen, J.X. (2006). Ultrasonic communication in frogs. Nature *440*, 333–336.
11. Peters, R.A., Hemmi, J.M., and Zeil, J. (2007). Signaling against the wind: modifying motion-signal structure in response to increased noise. Curr. Biol. *17*, 1231–1234.
12. Hotchkin, C., and Parks, S. (2013). The Lombard effect and other noise-induced vocal modifications: insight from mammalian communication systems. Biol. Rev. Camb. Philos. Soc. *88*, 809–824.
13. Römer, H. (2013). Masking by noise in acoustic insects: problems and solutions. In Animal Communication and Noise, H. Brumm, ed. (Springer), pp. 33–63.
14. Vélez, A., Schwartz, J.J., and Bee, M.A. (2013). Anuran acoustic signal perception in noisy environments. In Animal Communication and Noise, H. Brumm, ed. (Springer), pp. 133–185.
15. Simoncelli, E.P., and Olshausen, B.A. (2001). Natural image statistics and neural representation. Annu. Rev. Neurosci. *24*, 1193–1216.
16. Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. Annu. Rev. Psychol. *59*, 167–192.
17. Garcia-Lazaro, J.A., Ahmed, B., and Schnupp, J.W. (2006). Tuning to natural stimulus dynamics in primary auditory cortex. Curr. Biol. *16*, 264–271.
18. McDermott, J.H., Schemitsch, M., and Simoncelli, E.P. (2013). Summary statistics in auditory perception. Nat. Neurosci. *16*, 493–498.
19. Theunissen, F.E., and Elie, J.E. (2014). Neural processing of natural sounds. Nat. Rev. Neurosci. *15*, 355–366.
20. Singh, N.C., and Theunissen, F.E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. J. Acoust. Soc. Am. *114*, 3394–3411.
21. Grose, J.H., and Hall, J.W., 3rd. (1992). Comodulation masking release for speech stimuli. J. Acoust. Soc. Am. *91*, 1042–1050.
22. Festen, J.M. (1993). Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice. J. Acoust. Soc. Am. *94*, 1295–1300.
23. Kwon, B.J. (2002). Comodulation masking release in consonant recognition. J. Acoust. Soc. Am. *112*, 634–641.
24. Ronacher, B., and Hoffmann, C. (2003). Influence of amplitude modulated noise on the recognition of communication signals in the grasshopper *Chorthippus biguttulus*. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *189*, 419–425.
25. Gerhardt, H.C. (1975). Sound pressure levels and radiation patterns of vocalizations of some North American frogs and toads. J. Comp. Physiol. *102*, 1–12.
26. Vélez, A., and Bee, M.A. (2011). Dip listening and the cocktail party problem in grey treefrogs: signal recognition in temporally fluctuating noise. Anim. Behav. *82*, 1319–1327.
27. Gerhardt, H.C., and Klump, G.M. (1988). Masking of acoustic signals by the chorus background noise in the green treefrog: a limitation on mate choice. Anim. Behav. *36*, 1247–1249.
28. Wollerman, L. (1999). Acoustic interference limits call detection in a Neotropical frog *Hyla ebraccata*. Anim. Behav. *57*, 529–536.
29. Wollerman, L., and Wiley, R.H. (2002). Background noise from a natural chorus alters female discrimination of male calls in a Neotropical frog. Anim. Behav. *63*, 15–22.
30. Bee, M.A., and Schwartz, J.J. (2009). Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. J. Acoust. Soc. Am. *126*, 2788–2801.
31. Bee, M.A. (2008). Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. Anim. Behav. *75*, 1781–1791.
32. Ward, J.L., Buerkle, N.P., and Bee, M.A. (2013). Spatial release from masking improves sound pattern discrimination along a biologically relevant pulse-rate continuum in gray treefrogs. Hear. Res. *306*, 63–75.
33. Caldwell, M.S., and Bee, M.A. (2014). Spatial hearing in Cope's gray treefrog: I. Open and closed loop experiments on sound localization in the presence and absence of noise. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *200*, 265–284.
34. Schwartz, J.J., Buchanan, B.W., and Gerhardt, H.C. (2001). Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments. Behav. Ecol. Sociobiol. *49*, 443–455.
35. Feng, A.S., Narins, P.M., and Capranica, R.R. (1975). Three populations of primary auditory fibers in the bullfrog (*Rana catesbeiana*): their peripheral origins and frequency sensitivities. J. Comp. Physiol. *100*, 221–229.
36. Simmons, D.D., Meenderink, S.W.F., and Vassilakis, P.N. (2007). Anatomy, physiology, and function of the auditory end-organs in the frog inner ear. In Hearing and Sound Communication in Amphibians, *Volume 29*, P.A. Narins, A.S. Feng, R.R. Fay, and A.N. Popper, eds. (Springer), pp. 184–220.
37. Hillery, C.M. (1984). Detection of amplitude-modulated tones by frogs: implications for temporal processing mechanisms. Hear. Res. *14*, 129–143.

38. Feng, A.S., Hall, J.C., and Siddique, S. (1991). Coding of temporal parameters of complex sounds by frog auditory nerve fibers. J. Neurophysiol. *65*, 424–445.

39. Schul, J., and Bush, S.L. (2002). Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. Proc. Biol. Sci. *269*, 1847–1852.

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

41. Ward, J.L., Love, E.K., Vélez, A., Buerkle, N.P., O'Bryan, L.R., and Bee, M.A. (2013). Multitasking males and multiplicative females: dynamic signalling and receiver preferences in Cope's grey treefrog. Anim. Behav. *86*, 231–243.

42. Goense, J.B.M., and Feng, A.S. (2012). Effects of noise bandwidth and amplitude modulation on masking in frog auditory midbrain neurons. PLoS ONE *7*, e31589.

43. Branstetter, B.K., and Finneran, J.J. (2008). Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). J. Acoust. Soc. Am. *124*, 625–633.

44. Fay, R.R. (2011). Signal-to-noise ratio for source determination and for a comodulated masker in goldfish, *Carassius auratus*. J. Acoust. Soc. Am. *129*, 3367–3372.

45. Guilford, T., and Dawkins, M.S. (1991). Receiver psychology and the evolution of animal signals. Anim. Behav. *42*, 1–14.

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

47. Moore, B.C.J. (2003). Introduction to the Psychology of Hearing, Fifth Edition (Elsevier).

48. Cooke, M., Hershey, J.R., and Rennie, S.J. (2010). Monaural speech separation and recognition challenge. Comput. Speech Lang. *24*, 1–15.

49. Webster, D.B., Fay, R.R., and Popper, A.N., eds. (1992). The Evolutionary Biology of Hearing (Springer).

50. Carr, C.E., and Christensen-Dalsgaard, J. (2015). Sound localization strategies in three predators. Brain Behav. Evol. *86*, 17–27.