

SCIENTIFIC REPORTS

OPEN

The progression pattern of male hyrax songs and the role of climactic ending

Vlad Demartsev^{1,2}, Amiyaal Ilany³, Arik Kershenbaum⁴, Yair Geva¹, Ori Margalit¹, Inbar Schnitzer³, Adi Barocas⁵, Einat Bar-Ziv⁶, Lee Koren³ & Eli Geffen¹

The study of animal vocal signals can either focus on the properties of distinct vocal elements or address the signal as a whole. Although some attention has been given to the continuous progression patterns of bird songs, such patterns in mammalian vocalisations have been largely overlooked. We examined temporal changes in structural and acoustic parameters in male rock hyrax songs. We found a gradual increase in call frequency and amplitude towards the song ending, as well as an abrupt increase in bout syntactic complexity, peaking in the last quintile of a song. In musical terms, such a pattern can be described as a crescendo (amplitude increase) with a terminal climax. In Western music, crescendos are used to maintain attention and direct the listeners towards a memorable highpoint of the musical piece. This structure may have an analogous function in animal communication, recruiting audience attention towards the climactic and potentially most informative part of the signal. Our playback experiments revealed that hyrax males tend to reply more to songs with a climactic ending, indicating that this progression pattern is important for hyrax communication. We suggest that animal vocal communication research can benefit from adding musical concepts to the analysis toolbox.

Although animal acoustic signals often form long and complex sequences¹, traditionally most of the research on animal song structure has focused on classifying the signal into distinct units (syllables). This allowed easy quantitative analysis of repertoire size, singing versatility and specific component production^{1,2}. However, such analysis can only reveal the first-order features of the signal and does not always account for such aspects as signal progression and dynamic structure, which might also bear information and influence the listeners. Dynamic features, such as formation of phrases containing a typical beginning, middle and end, stereotypic rhythms and pitch intervals, might bind seemingly discrete song elements into a single cohesive percept. Although there is an increasing interest in such signal properties and several attempts have been made to assess the continuous parameters of vocalisation^{3–6}, this approach is still far from being a common practice in animal vocalisation research.

Human music is able to induce emotional reactions and affect the behaviour of listeners⁷. Crescendo - a gradual increase in amplitude or intensity in a musical piece⁸, has been found to create an arousal effect in listeners⁹. An ending of a musical piece often consists of dynamic rhythmic changes or/and an amplitudinal climax, possibly to capture the audience's attention. The entire musical piece preceding the dramatic ending may funnel the listener towards that single highpoint at the end¹⁰, by establishing sender-receiver contact and creating expectations. Repetitions and innovations during the introduction build suspense and maintain listener attention¹¹ towards the powerful ending¹⁰.

Applying musical concepts to animal call features might reveal similar aspects of their effects on the behavioural states of listeners². The term “zoomusicology” established an analogy between non-human calls and human music several decades ago¹². The subject has mostly received attention in birds, after Charles Hartshorne¹³ suggested that every aspect of human music exists in bird songs: e.g. *accelerando* (gradual increase in tempo) in the field sparrow (*Spizella pusilla*) and ruffed grouse (*Bonasa umbellus*); *ritardando* (gradual decrease in tempo) in the yellow-billed cuckoo (*Coccyzus americanus*); *crescendo* (gradual increase in loudness) in Heuglin's robin

¹Dept. of Zoology, Tel Aviv University, Tel Aviv, 69978, Israel. ²Dept. of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, 8057, Switzerland. ³The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, 52900, Israel. ⁴Dept. of Zoology, University of Cambridge, Cambridge, UK. ⁵Dept. of Zoology and Physiology and Program in Ecology, University of Wyoming, Laramie, WY, 82071, USA. ⁶Mitrani Dept. of Desert Ecology, Ben-Gurion University of the Negev, Be'er Sheva, 8499000, Israel. Correspondence and requests for materials should be addressed to V.D. (email: demartsev@gmail.com)

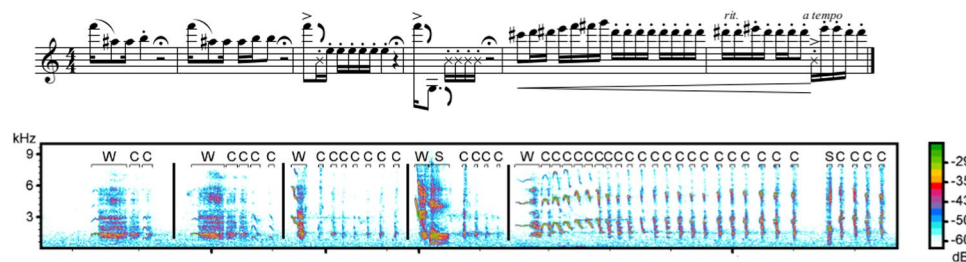


Figure 1. (a) Proposed musical notation of male song corresponding to the parts represented in the spectrogram. (b) Sample spectrogram of male rock hyrax song. Different vocal elements are marked and coded as W-Wail, C-Chuck and S-Snort. The representation order of singing bouts corresponds to natural temporal progression of a typical male hyrax song (beginning – low complexity, middle – low/intermediate complexity. end – high complexity). Vertical lines mark bout boundaries.

(*Cossypha heuglini*) and diminuendo (gradual decrease in loudness) in the Misto yellowfinch (*Sicalis luteola*)². In addition to the descriptive notion of musical features in animal calls, several attempts have been made to suggest a functional explanation for the phenomenon. For example, in the thrush nightingale (*Luscinia luscinia*) a song's temporal structure resembling an accelerando-like rhythm acceleration was found to progress through several song phrases towards a glissando (high-pitch sweep) finish. This gradual build-up towards culmination, manipulation of rhythmic timing and amplitude, could function to maintain the attention of receivers² and avoid signal habituation¹⁴.

Several mammals also produce songs with features similar to those used in music. Most famously, humpback whales (*Megaptera novaeangliae*) produce hours-long structured songs that involve highly styled repetitions of phrases and motifs¹⁵. Harbour seals (*Phoca vitulina*) produce long and complex vocal displays that, despite their inharmonious nature, have been classified as songs¹⁶. Among primates, the indri (*Indri*), tarsiers (*Tarsius*), titi monkeys (*Callicebus*) and gibbons (*Hylobates*) are known to sing. Siamang (*Hylabates syndactylus*) mated pairs produce long and stereotyped duets that include multiple different phrases and male-female transitions¹⁷.

In this work, we applied musical terminology to describe progression patterns of rock hyrax (*Procapra capensis*) vocalisation and discuss the theoretical parallelism between the functional effects of those patterns on the audience in both music and animal calls. In addition, we experimentally examined whether the climactic ending in hyrax songs has a communicational significance and affect the behaviour of conspecific listeners.

Adult rock hyrax males can be clearly identified as “bachelors”, who are mostly solitary and do not associate with any group¹⁸, and “residents” who reside with a group of females, routinely interact with them and are observed mate-guarding older and experienced females¹⁹. Both bachelors and residents produce long, complex vocalisations (songs), increasing in frequency towards the mating season. Male songs can be “spontaneous” (performed without any observable external trigger) or “induced” (performed following a specific external event; e.g. predator presence, conspecific alarm or distress calls, agonistic interaction, etc.)²⁰. Hyrax male songs consist of three vocal elements (wail, chuck and snort; Fig. 1) arranged in multiple bouts separated by distinct silent intervals^{21,22}, and are audibly recognized as progressively increasing in intensity (Fig. 1). Male calls (songs; Fig. 1) have been shown to reflect caller identity, age, social rank and body condition^{22–24}. Receivers are sensitive to both song structure²⁵ and to the signallers’ identity and individual traits²⁶. Signallers are attuned to the state of their audience and attempt to time their performance to periods of increased listener attention²⁷. Moreover, males increase the overall rhythm and complexity of their songs in the presence of an alert audience²⁰. Until now, the analysis of hyrax songs has been based on discrete measures of spectral parameters or on the composition of the different vocal elements within a song.

Here, we analysed hyrax songs from a different perspective. By viewing each song as a single unit, we addressed the song syntactic complexity and acoustic parameters on a wider scale that accounted for the dynamics of the signal progression. Demartsev, *et al.*²⁰ suggested that the syntactic complexity of a song might reflect the signalers' quality and play a role in male self-advertisement. We thus hypothesized that the complexity progression pattern would be more pronounced in the signals of high-quality males. Furthermore, songs that deviate from the general progression pattern might be less attractive to the audience and/or transmit signals of lower quality. We predicted that such songs would receive fewer replies from conspecific receivers in comparison to songs that reach a climactic ending.

Results

Natural song progression analysis. We analysed 140 previously recorded songs under natural conditions, performed by 24 adult males, out of 188 songs recorded²⁰. Forty eight songs, which were less than eight bouts long, were omitted from the analysis. Since the mean number of bouts/song (\pm SD) was 22.9 ± 19.7 bouts, we suspected that extremely short songs might have been terminated prematurely and were less suitable for the analysis of progression. Out of the 140 songs used, 60% showed an ascending progression pattern towards the song ending, 30.7% of songs did not show significant change, and 9.3% of songs showed a significant decrease in bout duration and/or entropy rate towards the song ending.

To test for progression in male hyrax songs, we used mixed models where male residency (resident or bachelor) and singing context (spontaneous or induced) were set as fixed effects, the relative temporal position of each vocal element in the song (hereafter proportion of song duration) was set as continuous predictor (covariate), the

| Term | Estimate | Wald χ^2 | P | Term | Estimate | Wald χ^2 | P |
|-----------------------------------|----------|---------------|------------------|-----------------------------------|----------|---------------|------------------|
| Wail | | | | Chuck | | | |
| Peak amplitude | | | | | | | |
| M _{RS} | 0.13 | 0.4 | 0.531 | M _{RS} | 0.13 | 0.4 | 0.545 |
| M _{SC} | −0.05 | 0.1 | 0.770 | M _{SC} | −0.23 | 2.2 | 0.139 |
| M _{SP} | 0.88 | 11.4 | 0.001 | M _{SP} | 0.54 | 8.6 | 0.003 |
| M _{RS} * M _{SP} | −0.27 | 0.4 | 0.524 | M _{RS} * M _{SP} | −0.23 | 0.3 | 0.560 |
| M _{SC} * M _{SP} | 0.12 | 0.1 | 0.716 | M _{SC} * M _{SP} | 0.42 | 2.1 | 0.151 |
| Peak frequency | | | | | | | |
| M _{RS} | −0.08 | 0.3 | 0.586 | M _{RS} | −0.07 | 0.1 | 0.732 |
| M _{SC} | 0.15 | 4.1 | 0.044 | M _{SC} | 0.32 | 5.2 | 0.022 |
| M _{SP} | 1.67 | 99.1 | <0.001 | M _{SP} | 1.61 | 56.7 | <0.001 |
| M _{RS} * M _{SP} | 0.14 | 0.2 | 0.664 | M _{RS} * M _{SP} | 0.16 | 0.2 | 0.681 |
| M _{SC} * M _{SP} | −0.27 | 3.1 | 0.079 | M _{SC} * M _{SP} | −0.60 | 5.6 | 0.018 |
| Fundamental frequency | | | | | | | |
| M _{RS} | 0.16 | 1.0 | 0.316 | M _{RS} | −0.02 | 0.0 | 0.932 |
| M _{SC} | 0.22 | 2.5 | 0.112 | M _{SC} | 0.05 | 0.1 | 0.762 |
| M _{SP} | 1.32 | 40.5 | <0.001 | M _{SP} | 0.89 | 35.7 | <0.001 |
| M _{RS} * M _{SP} | −0.33 | 1.0 | 0.310 | M _{RS} * M _{SP} | 0.04 | 0.0 | 0.891 |
| M _{SC} * M _{SP} | −0.42 | 2.2 | 0.139 | M _{SC} * M _{SP} | −0.10 | 0.1 | 0.734 |

Table 1. The effect of male residency (M_{RS}; resident or bachelor), singing context (M_{SC}; spontaneous or induced) and song progression (proportion of song duration; M_{SP}) on peak amplitude per bout, peak frequency per bout, and fundamental frequency per bout in the wail and chuck elements of hyrax male songs. $df = 1$ in all effect tests.

2-way interactions between fixed effects and proportion of song duration was included in the model, and male identity was set as a random effect. We tested the effect of the above explanatory variables on three vocal parameters of the chuck and wail elements (Table 1; Fig. 2) and on four bout characteristics in the male song (Table 2, Fig. 3).

A significant positive slope in the proportion of song duration (M_{sp}; Tables 1, 2) indicated an increase in the dependent variable along the song. Peak frequency, peak amplitude and fundamental frequency of elements all showed a clear trend of gradual increase during song progression (Table 1, Fig. 2). Except for peak frequency, this progression pattern was similar in both induced and spontaneous songs performed by males of both residency categories (Table 1). The increase in peak frequency was significantly different between spontaneous and induced songs for both the wail and the chuck elements (Table 1). Peak frequency in the wail element was higher in induced songs compared to spontaneous songs ($P = 0.044$; Table 1) but in both song types peak frequency increased along the song at equal rate (Table 1). For the chuck element, we detected a significant interaction between singing context and proportion of song duration ($P = 0.018$; Table 1). Peak frequency increased along the song at higher rate in spontaneous songs (GEE, slope = 1.64, Wald $\chi^2 = 114.5$, $df = 1$, $P < 0.001$) compare with induced songs (GEE, slope = 1.05, Wald $\chi^2 = 19.2$, $df = 1$, $P < 0.001$). In contrast to the gradual increase in frequencies, bout duration, number of chucks and snorts per bout, and entropy rate all showed a significant abrupt increase towards the end of the song, whereas the initial 50–80% of the song was maintained relatively constant (Fig. 3). This increase in bout complexity along the song was independent of male residency or singing context as both of these fixed effects and their interaction were insignificant in all tests (Table 2).

Playback analysis. In Set A we compared the rate of reply between playbacks of natural ascending songs (i.e. control; unaltered songs with increase in entropy and bout length) to playbacks of synthetic monotonous or descending songs. Playback type was set as the fixed effect and social group identity and track number as the random effects. The reply rate was significantly different between treatments (GEE, Wald $\chi^2 = 34.9$, $df = 2$, $P < 0.001$, $n = 83$). The synthetic monotonous songs, which lacked the increase in bout length and bout entropy towards the end, were replied to at a similar rate to that of natural songs (control, multiple comparisons by sequential Bonferroni $P = 0.568$). However, the reply rate to synthetic descending songs, with decreased bout length and entropy at the final stages of the song, was significantly lower, showing only ~30% probability of reply in comparison to ~50% in naturally ascending ($P = 0.007$) and synthetic monotonous ($P < 0.001$) songs (Fig. 4a).

The complementary Set B trials revealed a similar trend of natural descending songs receiving the lowest reply rate (~40%) and synthetic ascending songs receiving the highest reply rate (~70%). Synthetic monotonic songs received an intermediate ~50% reply rate (Fig. 4b). Although this trend of linear increase in reply rate with the increase in bout duration and bout entropy was strongly pronounced, it was not statistically significant (GEE, Wald $\chi^2 = 2.1$, $df = 2$, $P = 0.351$, $n = 92$).

To accommodate the effect of both male and track identities in the trials on reply rate, we repeated the above analysis after setting both male and track identities as random effects. The reply rate between the treatments of set A was significantly different (GEE, Wald $\chi^2 = 7.7$, $df = 2$, $P = 0.021$, $n = 83$). The reply rate to the synthetic

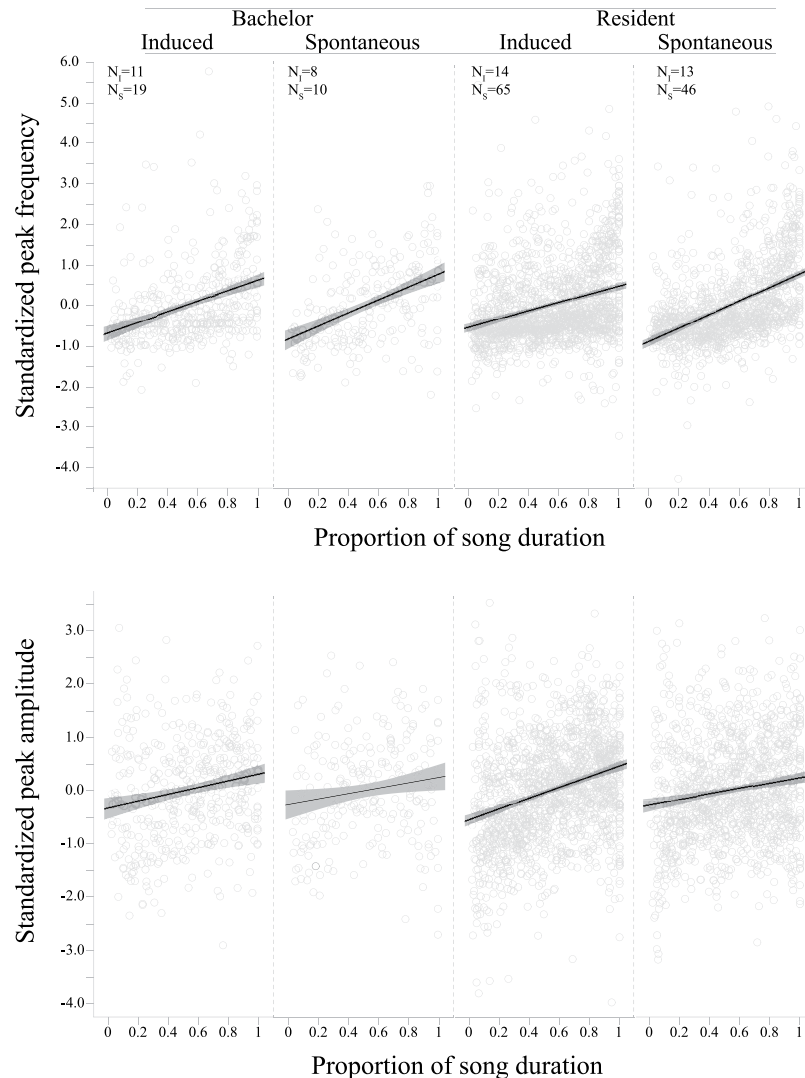


Figure 2. Progression plots of (a) standardised peak frequency and (b) standardised peak amplitude of vocal element and their proportional location through the signal. The trend lines show linear slope estimates with CI as shaded areas. **Bachelor** songs were performed by males without stable association with a female group. **Resident** songs were performed by males that were regularly observed associating with a female group. **Induced** songs were performed following an external trigger event. **Spontaneous** songs were performed without any observable trigger. N_i denotes number of individual males and N_s denotes number of songs in a corresponding category.

descending songs was significantly lower than to the natural songs ($P = 0.016$). The reply rate between the treatments of set B was insignificant as above (GEE, Wald $\chi^2 = 2.7$, $df = 2$, $P = 0.254$, $n = 92$).

Further, we tested all 24 different (natural and synthetic) playback tracks for the effect of bout duration and bout entropy rate on the conspecific reply rate. We used the regression slope between the proportion of song duration and bout duration (Fig. S1) or bout entropy rate as a predictor for reply rate. In this model, both regression slopes were the independent variables, reply rate was the dependent, and both male and track identities were set as the random effects. The reply rate significantly increased with the rise in the slope of bout entropy rate (GEE, Wald $\chi^2 = 5.1$, $df = 1$, $P = 0.023$, $n = 175$). A similar trend, although insignificant, was observed for bout duration (Wald $\chi^2 = 2.9$, $df = 1$, $P = 0.087$, $n = 175$). The effect of the interaction between bout entropy rate and bout duration on the reply rate was insignificant (Wald $\chi^2 = 1.0$, $df = 1$, $P = 0.321$). These results suggest that males tend to reply significantly more to songs that end up with higher complexity.

Last, in order to control for possible effects of audio track manipulation, we have performed a pairwise comparison of reply rate between natural ascending and synthetic ascending tracks, and a complementary comparison of reply rate between natural descending and synthetic descending tracks. No significant difference in reply rate was found between the natural and synthetic ascending (GEE, Wald $\chi^2 = 0.453$, $df = 1$, $P = 0.501$, $n = 60$) or descending (Wald $\chi^2 = 1.9$, $df = 1$, $P = 0.167$, $n = 56$) tracks.

| Term | Slope | Wald χ^2 | P |
|-----------------------------------|-------|---------------|--------|
| <i>Bout duration</i> | | | |
| M _{RS} | −0.17 | 1.1 | 0.289 |
| M _{SC} | 0.17 | 2.7 | 0.099 |
| M _{SP} | 0.99 | 43.9 | <0.001 |
| M _{RS} * M _{SP} | 0.32 | 1.1 | 0.298 |
| M _{SC} * M _{SP} | −0.32 | 2.7 | 0.098 |
| <i>Number of chucks/bout</i> | | | |
| M _{RS} | 0.10 | 0.5 | 0.498 |
| M _{SC} | 0.17 | 2.9 | 0.091 |
| M _{SP} | 2.30 | 250.8 | <0.001 |
| M _{RS} * M _{SP} | −0.20 | 0.5 | 0.474 |
| M _{SC} * M _{SP} | −0.33 | 2.9 | 0.088 |
| <i>Number of snorts/bout</i> | | | |
| M _{RS} | −0.05 | 0.3 | 0.566 |
| M _{SC} | −0.12 | 1.4 | 0.241 |
| M _{SP} | 1.29 | 215.6 | <0.001 |
| M _{RS} * M _{SP} | 0.09 | 0.3 | 0.574 |
| M _{SC} * M _{SP} | 0.22 | 1.4 | 0.243 |
| <i>Bout entropy rate</i> | | | |
| M _{RS} | −0.00 | 0.1 | 0.742 |
| M _{SC} | 0.01 | 0.3 | 0.595 |
| M _{SP} | 0.43 | 35.3 | <0.001 |
| M _{RS} * M _{SP} | −0.01 | 0.4 | 0.523 |
| M _{SC} * M _{SP} | 0.00 | 0.0 | 0.861 |

Table 2. The effect of male residency (M_{RS}; resident or bachelor), singing context (M_{SC}; spontaneous or induced) and song progression (proportion of song duration; M_{SP}) on bout duration, the number of chucks and snorts per bout, and bout entropy rate in hyrax male songs. *df* = 1 in all effect tests.

Discussion

Our analysis of natural song progression revealed an increase in all the measured parameters throughout the signal. However, the increase rates were not identical. Peak frequency, peak amplitude and fundamental frequency increased gradually, whereas bout duration and entropy rate remained constant throughout most of the song duration, until a steep increase at the final part of the signal. Our playback experiments verified that the pattern of increased bout length and complexity towards a songs' ending is significant for hyrax communication, as the receivers' reply rate was higher in response to songs with a more complex ending in both experimental sets.

The functional importance of vocal temporal parameters has been shown in several mammalian species. Gibbon males produce a progressively monotonic call sequence in order to demonstrate their level of readiness and to serve as an invitation for a female to join in a duet²⁸. However, in the presence of a predator, gibbons perform a “crescendo”, starting with soft notes that rapidly become louder²⁹. The progression pattern across hyrax songs might be aimed at drawing audience attention, which is not easily gained, by generating expectation of a climactic signal ending. Males often initiate singing following external events that promote alertness among neighbouring individuals and exploit this heightened alertness to deliver the signal to a wider and more attentive audience²⁰. In a similar manner, the gradual increase in hyrax song amplitude may attract listeners and maintain their attention. In other systems, higher amplitude signals have been found to draw more attention³⁰ and to be robust to habituation or withdrawal response³¹. For example, male mice (*Mus musculus*) perform an ultrasonic loud, modulated, and syntactically complex “calling song” when they are exposed to the scent of a female, in order to attract her closer. However, when the female is already present and there is no more need to attract her attention, males switch to a less energetic song with consistent pitch and more stereotyped structure³².

Hyrax songs have a long crescendo build-up with a gradual increase in fundamental frequency that may be similar to an “upwards glissando” (a gradual increase in pitch)⁸. In this work, we did not experimentally assess the effects of song frequency and song amplitude on conspecific attention and reply rate. A potential follow-up work, focused on manipulating these parameters, would lead to a better understanding of their function in the context of hyrax songs. At this stage, we can only hypothesize that a crescendo structure with a gradual upward frequency change may attract a larger audience and maintain higher attentiveness towards the end of the song.

Additionally, an acoustic climax in animal calls can be associated with an elevated vocal effort that is sensitive to fatigue, aging and disease³³. As such, the terminal call climax can serve as an honest signal of the performers' condition near the point of exertion, consistent with the handicap principle³⁴. Such “on the edge”³³ performance might be a burden for low-quality individuals³⁵, which may become exhausted after the long introductory stage and produce a less complex ending. In this case, the long warm-up may serve as a handicap to emphasize the performers' ability. Our results, however, showed no differentiation between resident and bachelor males, as both male categories are able to produce songs with a complex ending. Singing behaviour is generally cheap for male

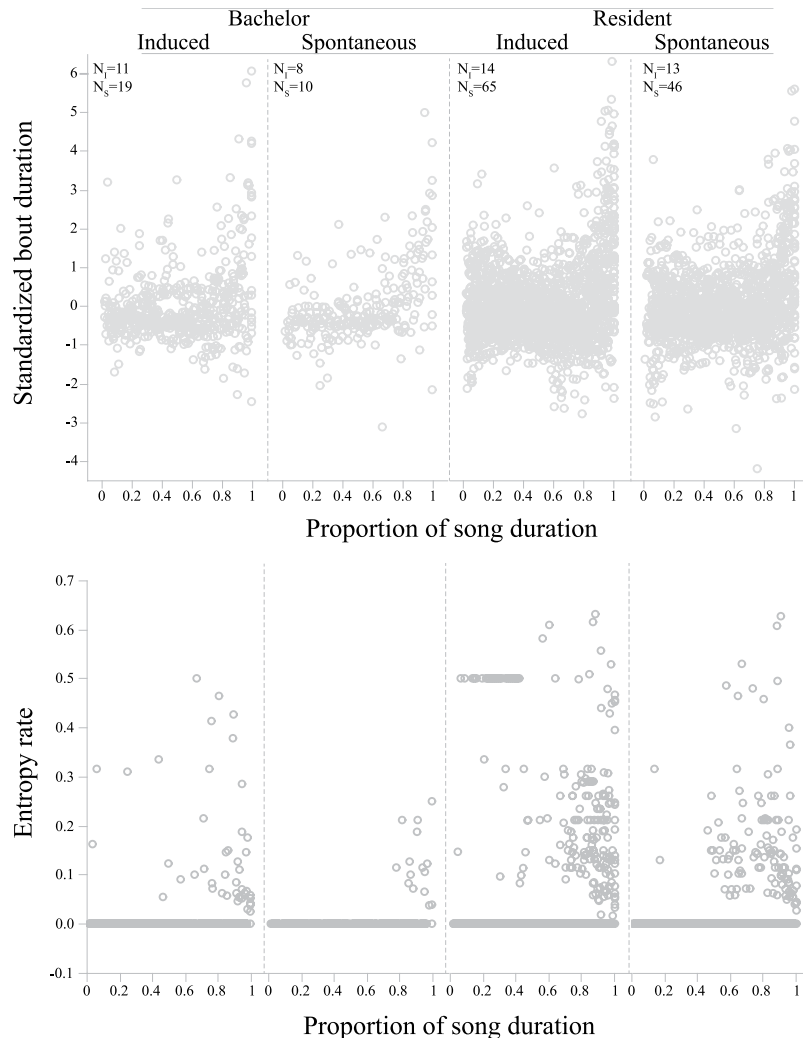


Figure 3. Progression plots of (a) standardised bout duration and (b) bout entropy rate and their proportional location through the signal. **Bachelor** songs were performed by males without stable association with a female group. **Resident** songs were performed by males that were regularly observed associating with a female group. **Induced** songs were performed following an external trigger event. **Spontaneous** songs were performed without any observable trigger. N_i denotes number of individual males; N_s denotes number of songs in a corresponding category.

hyraxes, at least in terms of energetic investment³⁶; however, the relative cost might be higher for individuals in poor condition.

The results of our playback experiments demonstrate that the listeners can detect the climactic ending of the songs, as conspecifics tended to reply more to climactic songs in comparison to their descending variants (Fig. 4). This supports our notion of a complex ending advertising higher quality and/or drawing more attention, as higher intensity signals are often followed by an increase in the response rate^{37–39}, although opposite patterns have also been reported^{40,41}.

The gain of audience attention and the handicap principle are not mutually exclusive explanations for the signal progression pattern that we observed in hyrax songs. In fact, the song's progression pattern possibly achieves both goals: keeping the audience attentive until the end by increasing amplitude and demonstrating singer capabilities (i.e. the handicap principle) by performing a complex ending. We can find a similar pattern in music, especially in opera performances and rock music. In many opera sessions and rock songs, a gradual build-up phase is followed by a climactic finale, which involves singing performance at the edge of human vocal capabilities and extremely difficult electric guitar solos, respectively. Such a dramatic ending of a musical piece creates an exciting and memorable moment for the listeners and at the same time demonstrates the quality of the performing artists.

The parallelism between features of animal calls and human music deserves serious consideration when analysing sequential animal calls from the perspective of their communicational value. An analogy between animal and human songs has repeatedly been suggested^{42–44}. Indeed, it does seem that some of the principles that govern musical structures and make it attractive and interesting to human listeners might have a similar function in non-human vocal communication. As demonstrated in this work, assessment of the signal as a whole has the

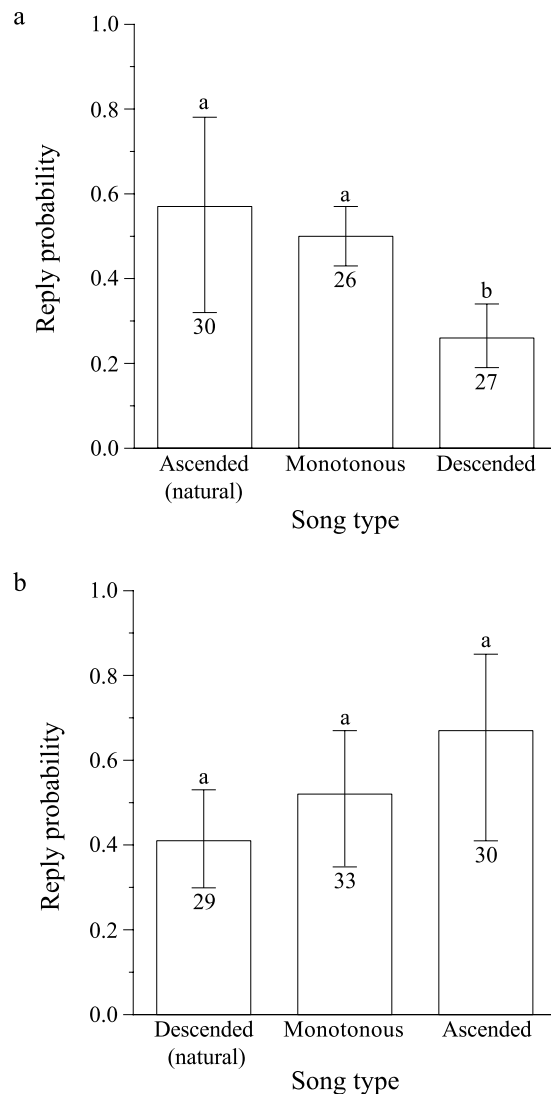


Figure 4. Reply rates ($\pm 95\%$ Wald Confidence Intervals) to playback experiments. **(a)** SetA (descending playbacks) – reply rates to natural, ascending control songs (N_{ASC}) and synthetic monotonous (S_{MON}) and synthetic descending (S_{DSC}) versions. **(b)** SetB (ascending playbacks) – reply rates to natural, descending control songs (N_{DSC}) and synthetic monotonous (S_{MON}) and synthetic ascending (S_{ASC}) versions. Letters above error bars denote statistical significance ($P < 0.05$) and values below error bars denote number of playbacks.

potential to reveal novel information channels that can be easily overlooked by analysis that focuses mainly on distinct units. The use of musical concepts in this context could aid in explaining the potential communicative significance of such novel channels. In addition, different signal units may in fact be perceptually bound together¹. Thus, treating the signal as a whole may be more informative and enhance our understanding of the way signals are experienced by conspecific receivers.

Materials and Methods

Ethical statement. This study was conducted under permits from the Israeli Nature and Parks Authority (NPA), which is the government agency responsible for supervising all wildlife research in Israel. All field procedures were in accordance with NPA guidelines and regulations for trapping, sampling and handling wild hyraxes, as well as for conducting playback. Yearly reports detailing all performed procedures and experiments involving animal subjects were submitted to NPA for assessment. A consecutive permit was granted following the approval of all performed activities by NPAs' Permit Department. Annual permit numbers are (2002/14674, 2003/14674, 2004/17687, 2005/17687, 2007/27210, 2008/31138, 2009/32871, 2010/37520, 2011/38061, 2012/38400, 2013/38803, 2014/40185, 2015/40768 and 2016/41174). Throughout the entire course of our study, no long-term stress or interference effects were detected in the individual animals or in the population. Both the general population size and the integrity of the specific social groups in the research area remained stable.

Field protocol. The study was conducted at the Ein Gedi Nature Reserve in Israel (31°28'N, 35°24'E) as part of a long-term project. The data for the current study were collected continuously from 2002 to 2016.

Field procedures followed previously published protocols^{18, 22, 23, 36}. Briefly, hyraxes were observed during the morning activity hours, about 4 hours daily, using 10X42 binoculars (Monarch, Nikon) and a telescope with up to X75 zoom magnification (Fieldscope ED82, Nikon). Rock hyraxes were trapped using live box traps (Tomahawk Live Trap Co, Tomahawk, WI, USA) baited with cabbage and kohlrabi. The traps were set at dawn, inspected after 3–4 hours, and locked open until the next trapping session. Trapped animals were anaesthetised by intramuscular injection of ketamine hydrochloride (0.1 ml/kg). Each hyrax was individually marked with a subcutaneous transponder (DataMars SA) and either an ear tag (~0.25 grams per tag) or a light numbered collar (~5 grams). Captured hyraxes were weighed and measured. Following anaesthesia recovery (at least 120 min), the animals were released back at their capture sites and resumed full normal activity. All treatments were performed in the shade to avoid overheating.

Residency status. Each year, male hyraxes' residency status (i.e. bachelor or resident) was determined according to the social network algorithms described in Barocas, *et al.*¹⁸. Resident males were observed in a stable association with a group of females, sharing sleeping dens and feeding sites. Bachelor males showed no stable association with other individuals and were observed only in brief interactions with females during the mating season.

Vocalisation recording and analysis. Hyrax long-range vocalisations were recorded from a distance of 10–50 m with a Sennheiser ME 67 shotgun microphone (frequency response 50–20,000 Hz \pm 2.5 dB) powered by a Sennheiser K6 module, and covered with a Sennheiser MZW70-1 blimp windscreen (Sennheiser Electronic GmbH & Co. K. G., Wedemark, Germany). The microphone was hand-held using an MZS20-1 shock-mount with a pistol grip. Vocalisations were recorded in mono (Tascam HD-P2 digital audio recorder; TASCAM Corporation, Montebello, CA, USA), with a sampling frequency of 48 kHz and a sampling width of 24 bits²⁷.

We analysed 140 previously recorded songs performed under natural conditions by 24 adult males. Songs were classified using the performer's residence status (bachelor or resident) for the relevant year. For each male category, we then classified songs into two states: "spontaneous" or "induced"²⁰.

For each song, a spectrogram was generated using Avisoft SAS LabPro software version 5.2.07 (Avisoft Bioacoustics, Berlin, Germany). Spectrograms were measured at 512 FFT length, 100% frame, using a Hamming window. All the vocal elements were identified and manually marked from the sonograms using the Avisoft SASLabPro cursors. For each element, we measured the peak frequency, peak amplitude, fundamental frequency, start time, end time and duration using SASLabPro automatic spectrogram parameters function. Singing bouts were defined by measuring the silent intervals between vocal elements. Intervals above 1 second separated elements into consecutive bouts. Bout duration was calculated by subtracting start time of the first bout element from the end time of the last bout element. To standardize for differences in song length, we converted the time passed from the beginning of the song for each bout element into a percentage of the total song length. This percentage was specified as an explanatory variable in subsequent models.

Amplitude measurements are sensitive to distance and to body orientation of an animal relative to the microphone⁴⁵. During singing, male hyraxes stay at a single location, usually on the top of a raised rock or a tree branch, and keep their position until singing is complete. Changes in body orientation during singing are rare and mostly a result of a distraction, in most cases causing the termination of a singing session. Disturbed songs or interrupted songs were not used in the analysis.

Bout complexity analysis. Male hyrax songs have been shown to contain a syntactic structure²¹. The relative rate of transitions between different vocal elements can be expressed as a 5×5 Markov transition matrix (wail, chuck and snort, with the addition of "bout start" and "bout end" markers), and we used the non-uniformity of this transition matrix to quantify syntactic diversity⁴⁶. We calculated syntax diversity using the weighted sum of the entropies of the transition matrix probabilities, which was shown by Shannon, *et al.*⁴⁷ to be an estimate of the entropy rate of an ergodic Markov chain⁴⁸. We set the "start-end" and "end-start" blocks as zero on all matrices because these transitions are biologically meaningless. The diversity in the rate of element change was calculated as $S = -\sum_i p_i \sum_j p_{i,j} \log_n p_{i,j}$ where $p_{i,j}$ is the probability of transition from element i to element j , taken from n possible elements ($n = 5$), and p_i represents the stationary probability of element i . We calculated the diversity of song bouts throughout the length of every song.

Playback experiments. Playback experiments were performed using a remotely-activated FoxPro Scorpion X1B speaker with a TX200 wireless remote controller (FOXPRO Inc., Lewistown, PA, USA), following our previously published protocols^{20, 27}. The speaker was placed before dawn in one of 10 concealed spots in an area frequently visited by focal hyraxes. The speaker was activated once we had observed and positively identified at least one stationary male hyrax or a group of at least four unidentified hyraxes, within a 20–30 m radius. If natural hyrax vocalisation had been heard, the playback initiation was postponed for at least 5 min. The mean \pm SE length of songs used for playback experiments was 103 ± 30 sec, as this is a common duration for hyrax singing²⁰. The songs were played according to their original duration, with no repetition/looping in any single playback trial. The amplitude of playbacks was calibrated by preliminary trials to match the normal level of hyrax singing (about 80 dB)³⁶. We restricted the number of playbacks to two per day/per site in order to minimize disturbance to hyrax daily routine and natural vocal interactions, to prevent hyraxes from becoming accustomed to the speaker location and to avoid desensitization. To minimize any specific site effects, the receivers for playback trials were selected from social groups located throughout the research area. To eliminate receiver familiarity with the singer, we used songs recorded several years previously and/or at a distance of at least 3 km from the playback sites. All behavioural responses of the individuals visible within the 30 m radius of the speaker during playbacks were noted. We considered a singing response as any reply by at least one of the present males within 2.5 min from the

end of playback. The 2.5 min threshold was set according to our previously published protocols²⁰ and based on the frequency of male hyrax singing³⁶. In cases of several males responding, only the first reply was considered for the playback analysis.

We conducted two sets of playback experiments (Appendix 1):

Set A (descending) - In this set, we sought to evaluate the effect of reducing the complexity of song ending on the probability of reply. We selected recordings from our natural song library based on sound quality. Songs were randomly used throughout the playback sessions (the number of repetitions for each song is indicated in Appendix 1). Five natural songs, each performed by a different adult male and with an increased complexity ending (natural ascending - N_{ASC} , Appendix 1), were used as the control and as a template for complexity reduction manipulation.

To determine whether decrease in bout length and bout complexity in the last ~20% of the song would also decrease receivers' reply rate in comparison to the natural control, N_{ASC} templates were digitally manipulated using Avisoft SAS LabPro software version 5.2.07 (Avisoft Bioacoustics, Berlin, Germany). Each song bout was manipulated separately so the original songs bout number was maintained. The shortening of the last bout slightly affected the overall song duration, however, the mean difference in song duration between natural and manipulated songs was negligible (mean \pm SE change 6.6 ± 4.8 sec; ~3.5% of the mean song length).

Two synthetic versions of each original recording were created (Figs S1, S2):

Synthetic Monotonic - S_{MON} : Singing bouts were shortened by replacing terminal bout elements with background noise from the same recording and the number of transitions between different elements was reduced to cancel out the strong rise in syntactic complexity towards the end of the song (monotonic tracks MON1- MON5, Appendix 1). The manipulated songs had the same bout number and approximately the same duration as the N_{ASC} tracks but lacked the rise in bout duration and complexity towards the end.

Synthetic Descending - S_{DSC} : Singing bouts were shortened and the number of element transitions was reduced so that the syntactic complexity would decrease in comparison to the middle of the song (descending tracks DSC1-DSC5, Appendix 1). The resulting songs had the same bout number and approximately same duration as the N_{ASC} tracks but had shorter bouts of reduced complexity towards the end.

The songs in Set A were played 83 times during the 2015–2016 field seasons, at 6 sites, with a mean of 13.8 ± 3.6 playbacks per site.

Set B (ascending) - Using this set we evaluated the effect of increasing the complexity of song ending on the probability of reply. We selected recordings from our natural song library based on sound quality. Songs were randomly used throughout the playback sessions (the number of repetitions for each song is indicated in Appendix 1). The three natural songs, each performed by a different adult male, which had a descending ending (natural descending - N_{DSC} , Appendix 1), were used as the control and as a template for complexity reduction manipulation.

To determine whether increase in bout length and bout complexity in the last ~20% of the song would also increase receivers' reply rate in comparison to the natural control, two synthetic versions of each original recording were created (Figs S1, S2):

Synthetic Monotonic - S_{MON} : Singing bouts were lengthened and the number of element transitions was increased so that the syntactic complexity would remain similar throughout the whole song (monotonic tracks MON6-MON8, Appendix 1). The added elements originated from the manipulated recording. They were duplicated and arranged while maintaining natural interval between consecutive elements. The resulting songs had the same bout number and approximately same duration as the N_{DSC} tracks but had constant bout length and complexity.

Synthetic Ascending - S_{ASC} : Singing bouts were lengthened and the number of element transitions was increased so that the syntactic complexity would increase in comparison to the beginning and the middle of the song (ascending tracks ASC6-ASC8, Appendix 1). The resulting songs had the same bout number and only slightly longer duration than the N_{DSC} tracks but had longer bouts of higher complexity towards the end.

The songs in Set B were played 92 times during the 2015–2016 field seasons, at 11 sites, with a mean \pm SE of 8.4 ± 5.3 playbacks per site.

Statistical analysis. Values of the listed dependent vocal variables (i.e. peak frequency, peak amplitude, fundamental frequency, bout duration, and number of chucks and snorts per bout), measured from male songs, were standardised within song prior to the analyses using the conventional equation $(X_i - \bar{X}_{\text{song}})/SD_{\text{song}}$ in order to eliminate large-scale differences in values between songs. To test for the effects of male residency (resident or bachelor), singing context (spontaneous or induced) and proportion of song duration (independent variables) on the dependent variables, we used the generalized estimating equation approach (GEE). The variable proportion of song duration scaled the position of each of the focal elements in the song relative to the beginning of the song. GEE are an extension of generalized linear models (GLM) for correlated data (i.e. mixed model), and specifically designed for repeated measures within the same subjects⁴⁹. We set individual as the random effect in all GEE analyses. The Wald χ^2 statistic was used for testing the significance of each of the effects and their interaction. Multiple comparisons were conducted using the sequential Bonferroni correction. GEE model fitting was done in SPSS (version 23, SPSS Inc.).

References

1. Kershenbaum, A. *et al.* Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews* **91**, 13–52, doi:10.1111/brv.12160 (2016).
2. Rothenberg, D., Roeske, T. C., Voss, H. U., Naguib, M. & Tchernichovski, O. Investigation of musicality in birdsong. *Hear. Res.* **308**, 71–83, doi:10.1016/j.heares.2013.08.016 (2014).
3. Cholewiak, D. M., Sousa-Lima, R. S. & Cerchio, S. Humpback whale song hierarchical structure: Historical context and discussion of current classification issues. *Mar. Mamm. Sci.* **29**, E312–E332, doi:10.1111/mms.12005 (2013).

4. Dahlin, C. R. & Wright, T. F. Does syntax contribute to the function of duets in a parrot, *Amazona auropalliata*? *Anim. Cogn.* **15**, 647–656, doi:10.1007/s10071-012-0493-y (2012).
5. Cody, M. L., Stabler, E., Castellanos, H. M. S. & Taylor, C. E. Structure, syntax and “small-world” organization in the complex songs of California Thrashers (*Toxostoma redivivum*). *Bioacoustics-the International Journal of Animal Sound and Its Recording* **25**, 41–54, doi:10.1080/09524622.2015.1089418 (2016).
6. Holy, T. E. & Guo, Z. S. Ultrasonic songs of male mice. *PLoS Biol.* **3**, 2177–2186, doi:10.1371/journal.pbio.0030386 (2005).
7. Eggermann, H., Pearce, M. T., Wiggins, G. A. & McAdams, S. Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience* **13**, 533–553, doi:10.3758/s13415-013-0161-y (2013).
8. Apel, W. *Harvard dictionary of music*. (Harvard University Press, 1953).
9. Schubert, E. Modeling perceived emotion with continuous musical features. *Music Perception* **21**, 561–585, doi:10.1525/mp.2004.21.4.561 (2004).
10. Osborn, B. Subverting the verse-chorus paradigm: Terminally climactic forms in recent rock music. *Music Theory Spectrum* **35**, 23–47, doi:10.1525/mts.2013.35.1.23 (2013).
11. Attas, R. Form as Process: The buildup introduction in popular music. *Music Theory Spectrum* **37**, 275–296, doi:10.1093/mts/mtv020 (2015).
12. Doolittle, E. & Gingras, B. Zoomusicology. *Curr. Biol.* **25**, R819–R820, doi:10.1016/j.cub.2015.06.039 (2015).
13. Hartshorne, C. Born to sing: Interpretation and world survey of bird song (Indiana University Press, 1973).
14. Hughes, A. L. Stereotyped and non-stereotyped features of the temporal patterning of singing sessions in the ovenbird *Seiurus auricapillus*. *Behav. Process.* **87**, 165–170, doi:10.1016/j.beproc.2011.02.010 (2011).
15. Payne, R. S. & McVay, S. Songs of humpback whales. *Science* **173**, 585–8, doi:10.1126/science.173.3997.585 (1971).
16. Fitch, W. T. The biology and evolution of music: A comparative perspective. *Cognition* **100**, 173–215, doi:10.1016/j.cognition.2005.11.009 (2006).
17. Geissmann, T. Gibbon songs and human music from an evolutionary perspective. (The MIT Press, 2000).
18. Barocas, A., Ilany, A., Koren, L., Kam, M. & Geffen, E. Variance in centrality within rock hyrax social networks predicts adult longevity. *Plos One* **6**, e22375, doi:10.1371/journal.pone.0022375 (2011).
19. Bar Ziv, E. *et al.* Individual, social, and sexual niche traits affect copulation success in a polygynandrous mating system. *Behav. Ecol. Sociobiol.* **1**–12, doi:10.1007/s00265-016-2112-4 (2016).
20. Demartsev, V. *et al.* Male hyraxes increase song complexity and duration in the presence of alert individuals. *Behav. Ecol.* **25**, 1451–1458, doi:10.1093/beheco/aru155 (2014).
21. Kershenbaum, A., Ilany, A., Blaustein, L. & Geffen, E. Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B-Biological Sciences* **279**, 2974–2981, doi:10.1098/rspb.2012.0322 (2012).
22. Koren, L. & Geffen, E. Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. *Behav. Ecol. Sociobiol.* **63**, 581–590, doi:10.1007/s00265-008-0693-2 (2009).
23. Koren, L., Mokady, O. & Geffen, E. Social status and cortisol levels in singing rock hyraxes. *Horm. Behav.* **54**, 212–216, doi:10.1016/j.yhbeh.2008.02.020 (2008).
24. Koren, L. & Geffen, E. Individual identity is communicated through multiple pathways in male rock hyrax (*Procavia capensis*) songs. *Behav. Ecol. Sociobiol.* **65**, 675–684, doi:10.1007/s00265-010-1069-y (2011).
25. Demartsev, V. *et al.* Harsh vocal elements affect counter-singing dynamics in male rock hyrax. *Behav. Ecol.* doi:10.1093/beheco/arw063 (2016).
26. Demartsev, V. *et al.* A mixed strategy of counter-singing behavior in male rock hyrax vocal competitions. *Behav. Ecol. Sociobiol.* **1**–9, doi:10.1007/s00265-016-2222-z (2016).
27. Ilany, A., Barocas, A., Koren, L., Kam, M. & Geffen, E. Do singing rock hyraxes exploit conspecific calls to gain attention? *Plos One* **6**, doi:10.1371/journal.pone.0028612 (2011).
28. Haraway, M. M. & Maples, E. G. Flexibility in the species-typical songs of gibbons. *Primates* **39**, 1–12, doi:10.1007/bf02557739 (1998).
29. Clarke, E., Reichard, U. H. & Zuberbuehler, K. The syntax and meaning of wild gibbon songs. *Plos One* **1**, doi:10.1371/journal.pone.0000073 (2006).
30. Tárano, Z. & Carballo, L. Call intercalation in dyadic interactions in natural choruses of Johnstone’s whistling frog *Eleutherodactylus johnstonei* (Anura: *Eleutherodactylidae*). *Behav. Process.* **126**, 55–63, doi:10.1016/j.beproc.2016.03.004 (2016).
31. Chan, A. A. Y.-H. *et al.* Increased amplitude and duration of acoustic stimuli enhance distraction. *Anim. Behav.* **80**, 1075–1079, doi:10.1016/j.anbehav.2010.09.025 (2010).
32. Chabout, J., Sarkar, A., Dunson, D. B. & Jarvis, E. D. Male mice song syntax depends on social contexts and influences female preferences. *Frontiers in Behavioral Neuroscience* **9**, doi:10.3389/fnbeh.2015.00076 (2015).
33. Riede, T., Arcadi, A. C. & Owren, M. J. Nonlinear acoustics in the pant hoots of common chimpanzees (*Pan troglodytes*): Vocalizing at the edge. *J. Acoust. Soc. Am.* **121**, 1758–1767, doi:10.1121/1.2427115 (2007).
34. Terleph, T. A., Malaivijitnond, S. & Reichard, U. H. Age related decline in female lar gibbon great call performance suggests that call features correlate with physical condition. *BMC Evol. Biol.* **16**, doi:10.1186/s12862-015-0578-8 (2016).
35. Zahavi, A. Cost of honesty - (Further remarks on handicap principle). *J. Theor. Biol.* **67**, 603–605, doi:10.1016/0022-5193(77)90061-3 (1977).
36. Ilany, A., Barocas, A., Kam, M., Ilany, T. & Geffen, E. The energy cost of singing in wild rock hyrax males: Evidence for an index signal. *Anim. Behav.* **85**, 995–1001, doi:10.1016/j.anbehav.2013.02.023 (2013).
37. Behr, O., Knorrich, M. & von Helversen, O. Territorial counter-singing in male sac-winged bats (*Saccopteryx bilineata*): low-frequency songs trigger a stronger response. *Behav. Ecol. Sociobiol.* **63**, 433–442, doi:10.1007/s00265-008-0677-2 (2009).
38. Kitchen, D. M. *et al.* Male baboon responses to experimental manipulations of loud “wahoo calls”: testing an honest signal of fighting ability. *Behav. Ecol. Sociobiol.* **67**, 1825–1835, doi:10.1007/s00265-013-1592-8 (2013).
39. Clutton-Brock, T. H. & Albon, S. D. Roaring of red deer and the evolution of honest advertisement. *Behaviour* **69**, 145–8, doi:10.1163/156853979x00449 (1979).
40. Cramer, E. R. A. & Price, J. J. Red-winged blackbirds *Agelaius phoeniceus* respond differently to song types with different performance levels. *J. Avian Biol.* **38**, 122–127, doi:10.1111/j.2006.0908-8857.03839.x (2007).
41. García, M., Wyman, M. T., Charlton, B. D., Fitch, W. T. & Reby, D. Response of red deer stags (*Cervus elaphus*) to playback of harsh versus common roars. *Naturwissenschaften* **101**, 851–854, doi:10.1007/s00114-014-1217-8 (2014).
42. Darwin, C. *The descent of man, and selection in relation to sex*. (John Murray, Albemarle Street, 1871).
43. Marler, P. Origins of music and speech: Insights from animals. (MIT Press, 2000).
44. Fitch, W. T. *The evolution of language*. (Cambridge University Press, 2010).
45. Gerhardt, H. C. In *Animal acoustic communication: Sound analysis and research methods* (eds Steven, L. Hopp, Michael, J. Owren & Christopher, S. Evans) 1–25 (Springer Berlin Heidelberg, 1998).
46. Lameira, A. *et al.* Predator guild does not influence orangutan alarm call rates and combinations. *Behav. Ecol. Sociobiol.* **67**, 519–528, doi:10.1007/s00265-012-1471-8 (2013).
47. Shannon, C. E., Weaver, W., Blahut, R. E. & Hajek, B. *The mathematical theory of communication*. (University of Illinois Press, 1949).

48. Kershenbaum, A. Entropy rate as a measure of animal vocal complexity. *Bioacoustics* **23**, 195–208, doi:[10.1080/09524622.2013.850040](https://doi.org/10.1080/09524622.2013.850040) (2014).
49. Overall, J. E. & Tonidandel, S. Robustness of generalized estimating equation (GEE) tests of significance against misspecification of the error structure model. *Biometrical Journal* **46**, 203–213, doi:[10.1002/bimj.200210017](https://doi.org/10.1002/bimj.200210017) (2004).

Acknowledgements

We are grateful to the Ein Gedi Nature Reserve staff for their logistic support, to the Nature and Park Authority for permission to work in the Ein Gedi Nature Reserve and to the Ein Gedi Field School for their hospitality. We thank all project students, field assistants, and guests for their valuable help in the field, and to N. Paz for her editorial services. We especially appreciate the assistance of Prof G. Koren and R. Koren in providing the musical notation of hyrax singing. This study was supported by four grants from the Israel Science Foundation (577/99, 488/05, 461/09, 550/14).

Author Contributions

V.D., A.I., Y.G., O.M., I.S., A.B., E.B.Z. and L.K. collected the data, V.D., A.K. and E.G. analysed the data, V.D. and E.G. wrote the main text, E.G. prepared the figures and all authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-03035-x](https://doi.org/10.1038/s41598-017-03035-x)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017