Wolf howls encode both sender- and context-specific information

Stuart K Watson[1] *, Simon W Townsend[1,2,3] * and Friederike Range[4,5]

[1]Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland
[2] Department of Psychology, University of Warwick, Warwick, U.K.
[3] Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland
[4] Wolf Science Center, Domestication Lab, Konrad Lorenz Institute of Ethology, University of Veterinary Medicine Vienna, Savoyenstraße 1a, A-1160 Vienna, Austria
[5] Comparative Cognition Unit, Messerli Research Institute, University of Veterinary Medicine Vienna, Medical University of Vienna and University of Vienna, Vienna, Austria

Received 11 December 2017
Initial acceptance 21 February 2018
Final acceptance 15 August 2018
MS number 17-00974R

*Joint first author

Correspondence: S. K. Watson and S. W. Townsend, Department of Comparative Linguistics, Plattenstrasse 54, CH-8032 Zürich, Switzerland.
E-mail address: stuart.watson@uzh.ch (S. K. Watson), simon.townsend@uzh.ch (S. W. Townsend)

Keywords: context specificity, individual differences, long-distance call, vocalizations

Abstract

Loud, long-distance calls serve varied functions across animal species including marking territory, attracting mates and signalling one´s identity. Here, we examined the types of sender- and context-specific information encoded in the howls of captive timber wolves, *Canis lupus*. We analysed 913 howls from nine individuals across three packs and investigated whether howl structure varied consistently as a function of phenotypic factors (age class, sex, pack and identity of the caller) in addition to the context in which the call was produced: specifically, whether the call was produced in a 'spontaneous' context just after sunrise or was 'elicited' by the absence of a group member. Calls were correctly classified by individual identity and production context, but not by any other factors. Principal components analyses indicated that individual differences were primarily associated with frequency-based measures, whereas acoustic variation between production contexts was associated with a variety of frequency-, intensity- and energy-based measures. Recognition of individual differences in vocalizations is likely to be important for navigating social relationships in wolves and further work is required to determine which life history factors may shape these individual differences. Differences resulting from production context are suggestive that these howl variants may serve different functions. The extent to which these individual and contextual differences are understood by receivers remains an open question.

Given their often obvious and striking nature, the long-distance vocalizations of animals have received considerable empirical research interest over the years (Hauser, 1996; Bradbury & Vehrencamp, 1998; Gustison & Townsend, 2015). From the infrasonic rumbles of African elephants, *Loxodonta africana*, to the songs of whales or birds, long-distance or ‘loud calls’ have been shown to serve a range of mating and territorial functions. For example, the loud calls of gibbons (*Hylobytes* spp.) play a role in negotiating and advertising territory among male–female pairs (Geissman, 2002), while the songs of many bird species are important in attracting females and even stimulating ovulation (Catchpole & Slater, 2003). The loud calls of social mammals, such as lions, *Panthero leo*, and chimpanzees, *Pan troglodytes*, have also been shown to serve multiple adaptive functions, such as signalling territories while maintaining contact and mediating cohesion with group members (Grinnell & McComb, 2001; Notman & Rendall, 2005). Analysis of the acoustic structure of these vocalizations and subsequent playbacks have helped shed further light on how exactly these calls have their effects. For example, the loud roars of red deer, *Cervus elaphus*, have long been known to represent sexually selected signals, being produced more frequently during the rutting or mating season (Clutton-Brock & Albon, 1979). Through applying a source-filter framework to the analysis of their roars it has additionally been shown that honest, accurate information on body size is cued through filter-related acoustic parameters, or formants, with larger males having more dispersed formant frequencies in their roars (Fitch & Reby, 2001). What is more, both males and females attend to this information and use it to modify their mating/fighting-based decisions with males avoiding and females approaching larger-sounding roars (Reby et al., 2003, Charlton et al., 2007).

A number of studies have now shown that long-distance vocalizations have the potential to cue an array of sender-specific, phenotypic information including the caller’s identity (Barbary macaques, *Macaca sylvanus*: Fischer, Hammerschmidt & Todt, 1998;

chacma baboons, *Papio ursinus*: Fischer, Hammerschmidt, Cheney & Seyfarth, 2001; Dolphins: Sayigh et al., 2007; chickens, *Gallus gallus domesticus*: Kent, 1987; meerkats, *Suricata suricatta*: Townsend & Manser, 2011), sex (Rendall et al., 2004; Charlton et al., 2009b), age (Charlton et al., 2009b) and group membership (Vehrencamp et al., 2003; Crockford et al., 2004). In addition to this, some species also encode more dynamic motivational, behavioural and/or contextual information in their loud calls. Encoding of contextual information in animal vocalizations, such as black-capped chickadees, *Poecile atricapillus*, expressing information about the size of a predator in their alarm calls (Templeton, Greene & Davis, 2005), has received considerable research attention over the years, partly due to its ostensible similarity to the highly context-specific nature of human language and the potential implications for understanding its evolutionary origins (Townsend & Manser, 2013; cf. Wheeler & Fischer, 2012). Furthermore, the capacity for both sender- and context-specific information to be encoded in a single call type has additionally been demonstrated (Briefer, Vannoni & McElligott, 2010; Cornec et al., 2015; Lemasson et al., 2009; Theis et al., 2007; Volodin et al., 2016). For example, male giant pandas, *Ailuropoda melanoleuca*, dynamically modulate the fundamental frequency (rate of vocal-fold vibration in the larynx) of their bleats to reflect their motivational state, increasing it when alone in order to broadcast their quality to potential mates (Charlton et al., 2015), whereas other acoustic features signal the size and sex of the individual (Charlton et al., 2009a). Indeed, the multi-encoding of static and dynamic features in a single call may, alongside sequentially combining vocalizations (e.g. Outtarra, Lemasson & Zuberbuhler, 2009), represent an additional mechanism by which animals can maximize the expressive power of a limited vocal repertoire (Manser, Seyfarth & Cheney, 2002). Here, we follow up existing work investigating whether this capacity is present in howls, the stereotypical loud call of wolves.

As with other social mammal loud calls, wolf howls are thought to function to mediate spacing within their groups (Mech & Boitani, 2010; Mazinni et al., 2013). This is likely to facilitate contact not only between separated group members but also between groups (Mech & Boitani, 2010, Zaccaroni et al., 2012; Nowak et al., 2007). Recent research has begun to shed light on the proximate mechanisms by which these effects come about, demonstrating, for example, that the acoustic structure of howls can be used to accurately predict individuality (Palacios et al., 2007; Root-Gutteridge et al., 2014) and group membership (Zaccaroni et al., 2012). Interestingly, previous work has also suggested that wolves produce howls in subtly different contexts: howls occur at increased rates spontaneously after sunrise (Gazzola et al., 2002; Harrington & Mech, 1982) and when faced with the temporary absence of group members (hereafter 'elicited' howls), both in the wild (Mech & Boitani, 2010; Nowak et al., 2007) and in captivity (Mazzini et al., 2013). Furthermore, individuals have also been shown to howl more often when separated from closely affiliated individuals (Mazzini et al., 2013). However, until now it was unknown whether calls produced in these different contexts also systematically differ in their acoustic structure. We therefore extended this body of work using a substantial data set to investigate whether, in addition to more static, individual-specific information types, wolf howls can also encode external, context-specific information.

Specifically, we examined the influence of various phenotypic attributes of callers and accompanying behavioural contexts on the acoustic structure of timber wolf howls. In line with the findings discussed above, we investigated the extent to which howls vary between individuals (Palacios et al., 2007; Root-Gutteridge et al., 2014) and packs (Mech & Boitani, 2010; Zaccaroni et al., 2012). Furthermore, in light of the consistent differences in size between the sexes (females are on average a third smaller than males, MacNulty et al, 2009) and the impact this has on vocal tract anatomy (Taylor & Reby, 2010), we expected to find sex-specific influences on overall acoustic structure of howls. Similarly, we also predicted that the howls of adult (24+ months) individuals would differ from those of juveniles (5–24 months) due to differences in size resulting from maturation. Regarding context, we determined whether howls produced in a spontaneous (just after sunrise) or elicited (by the temporary absence of a pack mate taken for a walk by care staff) context were acoustically distinct from one another.

Methods

Study Site and Subjects

All howls were recorded at the Wolf Science Center (WSC) in Ernstbrunn, Austria. Thirteen wolves, kept in three different packs, were subject to behavioural observations and acoustic recording (Table 1). All individuals were born in captivity from lineages originating in North America but came from different locations in North America and Europe. All were hand-raised in peer groups at the Wolf Science Center after being separated from their mothers in the first 10 days after birth (for details see Range & Viranyi, 2014). Puppies were bottle-fed and, after 3–4 weeks, hand-fed with solid food. All individuals had continuous access to humans for the first 5 months of their life. After 5 months, the wolves were integrated into established packs of the previous generations. We broadly defined two age categories in line with accepted definitions from the literature (Mech & Boitani, 2010). Adults were classified as individuals that were at least 2 years of age. Juveniles were classified as individuals that were between 5 months and 2 years of age. The wolves participated in training and/or cognitive and behavioural experiments at least once a day and, hence, still had frequent social contact with humans (Range & Viranyi, 2011). The enclosures of each of the three packs range over 4000–8000 m^2. They are equipped with trees, bushes, logs and shelters and water for drinking is permanently available. The wolves receive a diet of meat and dry food. All raising and keeping procedures of wolves at the Wolf Science Center are in line with the animal protection law in Austria (Tierversuchsgesetz 2012–TVG 2012). No special permission for use of animals (wolves) in such sociocognitive studies is required in Austria. The relevant committee that allows research on animals without special permission is Tierversuchskommission am Bundesministerium für Wissenschaft und Forschung (Austria).

Data Collection

Wolf howls were recorded with a directional microphone (ME66/K6 and a MZW66 pro windscreen, frequency response 40–20 000 Hz ± 2.5 dB; Sennheiser, Old Lyme, CT, U.S.A.) attached to a solid-state recorder (Marantz PMD 661), sampled at a frequency of 44.1 kHz. All howls were recorded at a distance of 1–10 m. Comments by the observer documenting the howling individual or the context were simultaneously recorded with a second speaker microphone (Sony FV100). Given that wild and captive observations both suggest that howling is most intense between mid-summer and mid-spring (Joslin, 1967; Harrington & Mech, 1982; Gazzola et al., 2002; Nowak et al., 2007) all recordings were performed over this period (June 2012–March 2013). Specifically, howl recordings were conducted during two different observational contexts to assess whether there were acoustic differences between the calls: (1) morning sessions which started at dawn and ended 2 h later (hereafter 'spontaneous' calls) and (2) leash walk sessions which took place each week and involved several individuals from the different packs being leash walked by an animal trainer at the WSC (hereafter 'elicited' calls). Morning recording sessions were performed on at least 5 days of the week. Leash walk recording sessions were performed as and when they were scheduled at the WSC, resulting in approximately three to four sessions per week. The remaining individuals in the enclosure were observed and all howls recorded, beginning when the individual on the walk was out of visual contact with the pack and ending when it returned.

Acoustic analysis

Recorded howls were uploaded from a solid-state recorder (Marantz PMD 661) to a PC notebook (IBM T41–Intel Centrino). All sound files were visually and audibly assessed to identify and select single howls with a high signal to noise ratio for analysis (see Fig. 1). Only howls that did not occur as part of a chorus were used for analysis since it was not possible to extract acoustic measures from overlapping calls. Ongoing work is investigating how the acoustic features of chorus howls and single howls vary. From these selected howls, a number of spectral and temporal acoustic parameters were extracted (see Table 2) using a custom-built script in Praat (version 5.5.53, praat.org, Reby & McComb, 2003). Source-related vocal parameters were measured by extracting the fundamental frequency (F0) contour of each call using a cross-correlation method ([Sound: To Pitch (cc) command] time step = 0.005 s, pitch floor = 20 Hz, pitch ceiling = 1200 Hz). To check whether the F0 contour was accurately tracked by Praat, the extracted F0 contour was visually compared to the F0 contour visualized in the spectrogram (e.g. Fig. 1). To filter out background noise, frequencies from 0 to 150 Hz were filtered from each howl, as all howls in a random sample of 30 howls had a minimum F0 of over 200 Hz. When inspecting analysis outputs, we noted sporadically high F0 measures. Detailed visual inspection of these calls indicated this was also due to miscellaneous background noise (e.g. birdsong) in higher frequency ranges. The exception to this was individual 'YU' who genuinely produced high-pitched howls. Consequently, for all other individuals we applied a filter that constrained F0 measures to a maximum of 1200 Hz. For 147 of the 913 calls analysed, the automated script was unable to extract a measure for peak frequency. Visual inspection of a subset of these howls suggested no obvious signal to noise ratio issues with the recordings. Hence, to avoid having to exclude these from the final analyses, we manually extracted peak frequency from the calls (by examining a spectral slice of the whole howl). For four individuals (SH, KA, TA, W), fewer calls were collected in at least one of the behavioural contexts than acoustic measurements were used for analysis (N = 15). Consequently, these individuals were excluded from the analysis reported here (Mundry & Sommer, 2007). Interobserver reliability was carried out by running identical acoustic analyses on a random selection of calls (N = 20). We found strong interobserver reliability, with an agreement of over 90%.

Statistical analysis

All acoustic parameters were initially assessed for multicollinearity to obtain a set of uncorrelated acoustic parameters. Multicollinearity is known to misleadingly inflate the standard errors of tested coefficients (Graham, 2003; Farrar & Robert, 1967). Q–Q plots were used to assess whether the data were normally distributed. For variables that were not normally distributed and could be improved by a log transformation, this was carried out. Other variables were not transformed. Variables with a variance inflation factor (VIF) greater than 10 were excluded from all analyses (Table 2).

To test whether the acoustic structure of single howls predictably differed between different classes of phenotypic factors (age class, ID, sex, pack membership) and between call production contexts, we used permuted discriminant function analyses (pDFA) with 1000 permutations (Mundry & Sommer, 2007). This was a necessary alternative to conventional DFA allowing us to control for the statistical conflict of using multiple data points per individual and estimate the significance of the number of correctly cross-validated single howls. A further advantage of the pDFA method is that it can handle unbalanced data sets, as is the case here, where there are different numbers of data points per factor level. When the individuals included in a specific data set contributed to only one class of the tested phenotypic factor (pack membership) a nested pDFA was performed. For data sets where all individuals contributed to more than one class of the test factor (e.g. call production context), a crossed pDFA was performed (Mundry & Sommer, 2007). In nested pDFAs where one of the levels of a test factor is nested within levels of another factor, it is possible to classify this as a restriction factor, causing permutations to only take place within that factor. Table 3 shows how each model was specified and the type of test used. Since preliminary work suggested an influence of context on howl structure (Hegland, 2014), where applicable, context of call production (spontaneous versus elicited) was used as a restriction factor. Where pDFAs reported statistically significant levels of call discrimination, we explored which acoustic factors contribute towards this by using principal components analyses (PCA). We retained principal components with eigenvalues greater than one (Kaiser's criterion) and factors were interpreted as loading highly if they had a correlation coefficient greater than 0.4 with the corresponding principal component (Budaev, 2010). For examining context-based differences, we then fitted a generalized linear mixed-effects model (GLMM) with each of the principal components as fixed effects, individual as a random effect and production context as the outcome variable. The purpose of this GLMM was to determine which principal components varied significantly between production contexts, and accordingly which corresponding factors were likely to contribute towards context-based discrimination of howls.

All statistical analyses were conducted in R version 3.42 (R Development Core Team, 2011) with RStudio v. 1.1.383, using the software package 'MASS' (Ripley et al. 2013). Scripts for carrying out pDFAs were provided by R. Mundry. All R scripts and data used to run this analysis are located at www.osf.io/5ptxf/

Results

We analysed 913 single howls from nine different individuals over a period of 10 months. Of these howls, 448 were recorded during morning observation sessions (spontaneous calls) and 465 during leash walk observation sessions. The individual contributions from each wolf as well as their sex, age class and the pack membership are listed in Table 1. The pDFAs found that calls could be correctly categorized at significantly above chance level by individual identity (correct: 38.75%; expected: 26.46; $P = 0.003$) and context of call production (correct: 62.38%; expected: 53.33; $P = 0.009$), but not by age class, pack or sex (see Table 3). Because one individual ('YU') was well known at the study site for producing atypical howls at a very high frequency, we wanted to be sure that this individual was not driving our pDFAs' ability to discriminate between individuals. Consequently, we reran the identity and context pDFAs without including howls from this individual but found that it was still able to correctly classify howls at above chance level (identity: correct: 30.5%; expected: 21.2%; $P = 0.011$; context: correct: 59.6%; expected: 52.5%; $P = 0.006$). Because there was no significant effect of age class upon howl acoustic structure, adult and juvenile calls were pooled for all other analyses.

To determine which variables contributed most towards individual differences in howls, we took the median of each acoustic measure for each individual and conducted a PCA on these data. The PCA produced nine principal components, the first of which had an eigenvalue greater than one, accounting for 47.7% of the variance (Table 4).

To examine which acoustic variables contributed towards discrimination between spontaneous and elicited contexts, we ran a PCA on the data used by the corresponding pDFA. This resulted in 15 principal components, the first six of which had eigenvalues greater than one and which cumulatively explained 70% of the variance. A GLMM determined that, of these six principal components, PC2, PC3, PC4 and PC6 varied significantly between elicited and spontaneous contexts ($P < 0.05$). Factor loadings greater than 0.4 were not clustered around frequency-, intensity- or energy-based variables (Table 5, Fig. 2).

Discussion

We analysed the acoustic structure of a large number of howls (913) from nine captive wolves to determine the types of phenotypic and contextual (whether the call was spontaneous or elicited) information that are encoded. We found that calls could be classified statistically according to the identity of the caller (but not their age class, sex or pack) and the context in which the call was produced.

Our findings confirm recent studies suggesting that wolf howls are individually distinctive (Palacios et al., 2007; Root-Gutteridge et al., 2014) with variance in acoustic structure between individuals probably attributable to interindividual anatomical differences (Yin and McCowen, 2004; Townsend et al., 2014; Charlton et al., 2009a). Interestingly, despite being statistically significant, the percentage with which the pDFA was able to correctly classify howls according to identity was lower (ca. 38%) than previous work reporting individual differences in wolf howls (e.g. 72% in Iberian wolves, Palacios et al., 2007). This may, to an extent, be due to differences in statistical approach: a conventional DFA, as used by Palacios et al. (2007), correctly classified our howl sample at 45% (expected: ca. 10%). However, because our howls were produced in different contexts it was necessary to simultaneously control for this, something traditional DFAs cannot do. While recognition of individual differences in vocalizations is taxonomically widespread (birds: Godard, 1991; primates: Keenan et al., 2016; elephants: McComb et al., 2000; cetaceans: Bruck, 2013), it is yet to be demonstrated in wolves and this is crucial to understanding the relevance of detected individual signatures in howls. Habituation/discrimination playback experiments whereby subjects are habituated to the howls of one individual and then exposed to the howls of a different individual (discrimination phase) could be one viable approach to test this.

It was somewhat surprising that howls did not differ according to age class or sex, since differences in size typically impact vocal anatomy (Taylor & Reby, 2010). Male and female wolves are known to differ in terms of gross anatomy (MacNulty et al., 2009) and probably also in underlying physiology (Dabbs & Mallinger, 1999, Deaux et al., 2016). The apparent absence of a sex effect may therefore be due to our relatively small and unbalanced sample size in this respect, consisting of seven males and just two females from whom we recorded enough howls to be used for analysis. Similarly, we had six adults and just three juveniles in the final sample. It may therefore be that we lacked the statistical power to identify the effects of these factors. Alternatively, in the case of age class, given that juveniles were towards the younger end of their age category (and therefore probably smaller; see Table 1) it may be that the acoustic structure of calls crystallizes during early adolescence with little appreciable further change into adulthood (despite further physical changes). However, to confirm this, it would be necessary to carry out a fully longitudinal design in which calls were collected from the same individuals during both adolescence and adulthood.

In contrast to previous work (Zaccaroni et al., 2012), we found no evidence for group-specific differences in howl structure. However, this finding should be interpreted cautiously as, although we had access to a large number of howls, these were derived from only a small number of individuals per group (black pack N = 3, red pack N = 3 and green pack N = 3). Hence the absence of evidence for group signatures may well be a by-product of insufficient statistical power to detect group differences. It is also worth noting that so called 'dialects' in animal vocalizations are often, although not exclusively (Elowson & Snowdon, 1994; Crockford et al., 2004; Watson et al., 2015), a consequence of genetic relatedness leading to greater within-group vocal tract similarities than between groups and, as such, more similar calls (Gouzoules & Gouzoules, 1990; Townsend et al., 2010; Kershenbaum et al., 2016). The packs at the WSC, on the other hand, are artificially composed. While in some packs a few animals are related to one another, in other packs none of the animals are related and some individuals are related to animals from other packs. This means that genetically driven acoustic variation is likely to be as great within packs as between them. Furthermore, if wild populations of wolves typically deploy kin-based social learning of call structures (e.g. matriline-based vocal learning in killer whales, *Orcinus orca*: Miller & Bain, 2000), or directed social learning dependent on a critical period (e.g. song learning in zebra finches, *Taeniopygia guttata*: George et al., 1995), this would not be expressed in our sample as they were hand-raised in peer groups (with animals from unrelated litters) that were later split to form different packs.

Lastly, our data indicate that there is a degree of flexibility in howl acoustic structure: howls systematically varied according to the motivational or behavioural context in which they were produced, namely a difference between spontaneous howls given shortly after sunrise and those elicited by the absence of a group member. Interestingly, previous work has also shown that the production of elicited calls is under flexible control of the caller, given that they are produced more frequently when the absent group mate is closely affiliated with the caller (Mazzini et al., 2013). Data from a range of species have demonstrated that both long- and short-distance calls can and do convey rich information sets associated with the ongoing behavioural context (see Townsend & Manser, 2013 for a review). For example, the screams of chimpanzees differ systematically based on the severity of aggression experienced (Slocombe & Zuberbuhler, 2007) and playback experiments have demonstrated that these differences are salient to receivers (Slocombe et al., 2009). Furthermore, dog growls or the groans of fallow deer, *Dama dama*, have also been shown to be influenced by either the valence of the context (play versus aggression) or the presence of specific individuals, respectively (Farago et al., 2010; Yin & McCowan, 2004; Charlton & Reby, 2011). While context-specific howls have been previously posited (Harrington, 1987; Palacios et al., 2007), to our knowledge this is the first systematic observational evidence that wolves utilize distinctive howl variants in different behavioural contexts.

According to our PCA, individual differences in call structure were associated with End F0, a fundamental frequency-based measure. This is in line with the vocalizations of other species, such as pandas, where fundamental frequency conveys information about individual level attributes such as age and size (Charlton et al., 2009b). With regard to contextual differences in howl acoustics, a number of frequency-, energy- and amplitude-based parameters loaded highly in PCs that differed significantly between elicited and spontaneous contexts (Tables 4, 5). However, caution should be taken when interpreting energy-based measures such as Fpeak and EfPeak, which loaded highly for context differences, as these are known to be sensitive to changes in recording distance to subject (Zollinger et al., 2012). In this study, spontaneous and elicited howls were always recorded at 1–10 m, but, owing to the long-term nature of the data, we do not have sufficient information to determine whether there were systematic differences in recording distance between contexts. Nevertheless, the lack of clustering around a category (e.g. energy) of variable in our results suggests that there is no single acoustic feature differentiating calls produced in different contexts, but rather that the 'holistic' structure of wolf howls has the capacity to encode, through a variety of acoustic features, information regarding the individual's motivational or behavioural states.

From a proximate perspective, differences in howl structure resulting from production context are likely to be a product of differing arousal levels experienced by the signaller (Charlton & Reby, 2011) driving concomitant changes in spectral and temporal parameters (Owren, Amoss & Rendall, 2011; but see Mazinni et al., 2013). However, these data can also help shed more general light on exactly how wolf howls can serve multiple recruitment and territorial functions. Specifically, our findings suggest that subtle differences in acoustic structure could potentially help receivers differentiate between howls directed at recruiting individuals back to the pack (elicited) versus those signalling territory and mediating intergroup spacing (spontaneous). However, systematic playback experiments are still necessary to determine whether these acoustic differences are indeed meaningful to receivers by examining whether they elicit differential behavioural responses.

Our results indicate that wolf howls encode information on both the identity of the caller and the behavioural context of production. They support recent work demonstrating that social carnivore vocal systems display an intriguing degree of complexity and hence represent a relevant model group for understanding the evolution and emergence of vocal complexity (Holekamp et al., 1999; Manser et al., 2014; Kershenbaum et al., 2016). Naturally, for each of the information sets detected, rigorous experimental verification is central to test whether these information sets are not just anatomical artefacts but are meaningful and relevant to receivers (see Townsend et al., 2010), reducing their uncertainty regarding the identity of the signaller and the behavioural context in which the call was produced (Seyfarth & Cheney, 2010).

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgments

The Wolf Science Center was established by Zsofia Virányi, Kurt Kotrschal and Friederike Range and we thank all the helpers who made this possible hence indirectly supporting this research. We thank the WSC staff for logistical support during the study and Roger Mundry for providing the R pDFA script. Particular thanks to Mauro Hegland for his help with data collection and logistical support. The project was supported by funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement to FR. [311870]. SWT was supported by the University of Zurich and SNF grant PP00P3-163850. We further thank many private sponsors including Royal Canin for financial support and the Game Park Ernstbrunn for hosting the Wolf Science Center. We are grateful to two anonymous referees for their helpful feedback on previous versions of the manuscript.

References

Bradbury, J. W., & Vehrencamp, S. L. 1998. *Principles of animal communication.* Sunderland, MA: Sinauer.

Briefer, E., & McElligott, A. G. (2011). Indicators of age, body size and sex in goat kid calls revealed using the source–filter theory. *Applied Animal Behaviour Science*, *133*(3), 175-185.

Bruck, J. N. (2013). Decades-long social memory in bottlenose dolphins. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*(1768), 20131726.

Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: caveats and guidelines. *Ethology*, *116*(5), 472-480.

Catchpole, C. K., & Slater, P. J. (2003). *Bird song: biological themes and variations*. Cambridge, U.K.: Cambridge University Press.

Charlton, B. D., & Reby, D. (2011). Context-related acoustic variation in male fallow deer (Dama dama) groans. PLoS One, 6(6), e21066.

Charlton, B. D., Keating, J. L., Rengui, L., Huang, Y., & Swaisgood, R. R. (2015). The acoustic structure of male giant panda bleats varies according to intersexual context. *The Journal of the Acoustical Society of America*, *138*(3), 1305-1312.

Charlton, B. D., Reby, D., & McComb, K. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*(4), 382-385.

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009a). Vocal cues to identity and relatedness in giant pandas (*Ailuropoda melanoleuca*). *The Journal of the Acoustical Society of America*, *126*(5), 2721-2732.

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009b). The information content of giant panda, *Ailuropoda melanoleuca*, bleats: acoustic cues to sex, age and size. *Animal Behaviour*, *78*(4), 893-898.

Clutton-Brock, T. H., & Albon, S. D. (1979). The roaring of red deer and the evolution of honest advertisement. *Behaviour*, *69*(3), 145-170.

Cornec, C., Hingrat, Y., Robert, A., & Rybak, F. (2015). The meaning of boom calls in a lekking bird: identity or quality information? *Animal behaviour*, *109*, 249-264.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild Chimpanzees Produce Group- Specific Calls: a Case for Vocal Learning? *Ethology*, *110*(3), 221-243.

Dabbs Jr, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and individual differences*, *27*(4), 801-804.

Déaux, E. C., Clarke, J. A., & Charrier, I. (2016). Dingo Howls: The Content and Efficacy of a Long- Range Vocal Signal. *Ethology*, *122*(8), 649-659.

Elowson, A. M., & Snowdon, C. T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour*, *47*(6), 1267-1277.

Faragó, T., Pongrácz, P., Range, F., Virányi, Z., & Miklósi, Á. (2010). 'The bone is mine': affective and referential aspects of dog growls. *Animal Behaviour*, *79*(4), 917-925.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. The Review of Economic and Statistics, 92-107.

Fischer, J., Hammerschmidt, K., & Todt, D. (1998). Local variation in Barbary macaque shrill barks. *Animal Behaviour*, *56*(3), 623-629.

Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2001). Acoustic features of female chacma baboon barks. *Ethology*, *107*(1), 33-54.

Fischer, J., Metz, M., Cheney, D. L., & Seyfarth, R. M. (2001). Baboon responses to graded bark variants. *Animal Behaviour*, *61*(5), 925-931.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*(1477), 1669-1675.

Gazzola, A., Avanzinelli, E., Mauri, L., Scandura, M., & Apollonio, M. (2002). Temporal changes of howling in south European wolf packs. *Italian Journal of Zoology*, *69*(2), 157-161.

Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological Reviews*, *77*(1), 57-76.

George, J. M., Jin, H., Woods, W. S., & Clayton, D. F. (1995). Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron*, *15*(2), 361-372.

Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird. *Nature*, *350*(6315), 228.

Gouzoules, H., & Gouzoules, S. (1990). Matrilineal signatures in the recruitment screams of pigtail macaques, *Macaca nemestrina*. *Behaviour*, *115*(3), 327-347.

Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. Ecology, 84(11), 2809-2815.

Grinnell, J., & McComb, K. (2001). Roaring and social communication in African lions: the limitations imposed by listeners. *Animal Behaviour*, *62*(1), 93-98.

Gustison, M. L., & Townsend, S. W. (2015). A survey of the context and structure of high-and low-amplitude calls in mammals. *Animal Behaviour*, *105*, 281-288.

Harrington, F. H. (1987). Aggressive howling in wolves. Animal Behaviour, 35(1), 7-12.

Harrington, F. H., & Mech, L. D. (1982). An analysis of howling response parameters useful for wolf pack censusing. *The Journal of Wildlife Management*, 686-693.

Harrington, F. H., Asa, C. S., Mech, L., & Boitani, L. (2003). Wolf communication. In L. D. Mech & L. Boitani (Eds). *Wolves: Behavior, ecology, and conservation*, (pp. 66-103). Chicago, IL: University of Chicago Press.

Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Hegland, M. (2014). The social information content and vocal flexibility of howls in timber wolves. (MSc thesis). Zurich, Switzerland: University of Zurich.

Holekamp, K. E., Boydston, E. E., Szykman, M., Graham, I., Nutt, K. J., Birch, S., ... & Singh, M. (1999). Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. *Animal Behaviour*, *58*(2), 383-395.

Joslin, P. W. (1967). Movements and Home Sites of Timber Wolves in Alǵonquin Park. *American Zoologist*, *7*(2), 279-288.

Keenan, S., Mathevon, N., Stevens, J. M., Guéry, J. P., Zuberbühler, K., & Levréro, F. (2016). Enduring voice recognition in bonobos. *Scientific Reports*, *6*, 22046.

Kent, J. P. (1987). Experiments on the relationship between the hen and chick (*Gallus gallus*): the role of the auditory mode in recognition and the effects of maternal separation. *Behaviour*, *102*(1), 1-13.

Kershenbaum, A., Root-Gutteridge, H., Habib, B., Koler-Matznick, J., Mitchell, B., Palacios, V., & Waller, S. (2016). Disentangling canid howls across multiple species and subspecies: Structure in a complex communication channel. *Behavioural Processes*, *124*, 149-157.

Lemasson, A., Boutin, A., Boivin, S., Blois-Heulin, C., & Hausberger, M. (2009). Horse (*Equus caballus*) whinnies: a source of social information. *Animal Cognition*, *12*(5), 693-704.

MacNulty, D. R., Smith, D. W., Mech, L. D., & Eberly, L. E. (2009). Body size and predatory performance in wolves: is bigger better?. *Journal of Animal Ecology*, *78*(3), 532-539.

Manser, M. B., Jansen, D. A., Graw, B., Hollén, L. I., Bousquet, C. A., Furrer, R. D., & le Roux, A. (2014). Vocal complexity in meerkats and other mongoose species. *Advances in the Study of Behavior*, (46), 281-310.

Manser, M. B., Seyfarth, R. M., & Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences*, *6*(2), 55-57.

Mazzini, F., Townsend, S. W., Virányi, Z., & Range, F. (2013). Wolf howling is mediated by relationship quality rather than underlying emotional stress. *Current Biology*, *23*(17), 1677-1680.

McComb, K., Moss, C., Sayialel, S., & Baker, L. (2000). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, *59*(6), 1103-1109.

Mech, L. D., & Boitani, L. (Eds.). (2010). *Wolves: behavior, ecology, and conservation*. Chicago, IL: University of Chicago Press.

Miller, P. J., & Bain, D. E. (2000). Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca*. *Animal Behaviour*, *60*(5), 617-628.

Mundry, R., & Sommer, C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. Animal Behaviour, 74(4), 965-976.

Notman, H., & Rendall, D. (2005). Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Animal Behaviour*, *70*(1), 177-190.

Nowak, S., Jędrzejewski, W., Schmidt, K., Theuerkauf, J., Mysłajek, R. W., & Jędrzejewska, B. (2007). Howling activity of free-ranging wolves (*Canis lupus*) in the Białowieża Primeval Forest and the Western Beskidy Mountains (Poland). *Journal of Ethology*, *25*(3), 231-237.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, *106*(51), 22026-22031.

Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, *73*(6), 530-544.

Palacios, V., Font, E., & Márquez, R. (2007). Iberian wolf howls: acoustic structure, individual variation, and a comparison with North American populations. *Journal of Mammalogy*, *88*(3), 606-613.

Range, F., & Virányi, Z. (2011). Development of gaze following abilities in wolves (*Canis lupus*). *PLoS One*, *6*(2), e16888.

Range, F., & Virányi, Z. (2014). Wolves are better imitators of conspecifics than dogs. *PLoS One*, *9*(1), e86559.

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, *65*(3), 519-530.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society of London B: Biological Sciences*, *272*(1566), 941-947.

Rendall, D., Owren, M. J., Weerts, E., & Hienz, R. D. (2004). Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. *The Journal of the Acoustical Society of America*, *115*(1), 411-421.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'mass'. http://cran.r-projectorg/web/packages/MASS/.

Root-Gutteridge, H., Bencsik, M., Chebli, M., Gentle, L. K., Terrell-Nield, C., Bourit, A., & Yarnell, R. W. (2014). Identifying individual wild Eastern grey wolves (*Canis lupus lycaon*) using fundamental frequency and amplitude of howls. *Bioacoustics*, *23*(1), 55-66.

Sayigh, L. S., Esch, H. C., Wells, R. S., & Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour*, *74*(6), 1631-1642.

Seyfarth, Robert M, and Dorothy L Cheney. 2010. Primate vocal communication. In M. Platt & A. A. Ghazanfar (Eds). *Primate neuroethology* (pp. 84-97). Oxford, U.K. Oxford University Press.

Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, *104*(43), 17228-17233.

Slocombe, K. E., Townsend, S. W., & Zuberbühler, K. (2009). Wild chimpanzees (*Pan troglodytes schweinfurthii*) distinguish between different scream types: evidence from a playback study. *Animal cognition*, *12*(3), 441-449.

Taylor, A. M., & Reby, D. (2010). The contribution of source–filter theory to mammal vocal communication research. *Journal of Zoology*, *280*(3), 221-236.

Team, R. (2015). *RStudio: integrated development for R.* Boston, MA: RStudio, Inc. http://www. rstudio. com.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science*, *308*(5730), 1934-1937.

Theis, K. R., Greene, K. M., Benson-Amram, S. R., & Holekamp, K. E. (2007). Sources of variation in the long-distance vocalizations of spotted hyenas. *Behaviour*, *144*(5), 557-584.

Townsend, S. W., & Manser, M. B. (2011). The function of nonlinear phenomena in meerkat alarm calls. *Biology letters*, *7*(1), 47-49.

Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: the past, present and the future. *Ethology*, *119*(1), 1-11.

Townsend, S. W., Charlton, B. D., & Manser, M. B. (2014). Acoustic cues to identity and predator context in meerkat barks. Animal behaviour, 94, 143-149.

Townsend, S. W., Hollén, L. I., & Manser, M. B. (2010). Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. *Animal Behaviour*, *80*(1), 133-138.

Vehrencamp, S. L., Ritter, A. F., Keever, M., & Bradbury, J. W. (2003). Responses to playback of local vs. distant contact calls in the orange- fronted conure, *Aratinga canicularis*. *Ethology*, *109*(1), 37-54.

Volodin, I. A., Sibiryakova, O. V., & Volodina, E. V. (2016). Sex and age-class differences in calls of Siberian wapiti *Cervus elaphus sibiricus*. *Mammalian Biology-Zeitschrift für Säugetierkunde*, *81*(1), 10-20.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., ... & Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, *25*(4), 495-499.

Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: a promising paradigm whose time has passed. *Evolutionary Anthropology: Issues, News, and Reviews*, *21*(5), 195-205.

Yin, S., & McCowan, B. (2004). Barking in domestic dogs: context specificity and individual identification. *Animal Behaviour*, *68*(2), 343-355.

Zaccaroni, M., Passilongo, D., Buccianti, A., Dessi-Fulgheri, F., Facchini, C., Gazzola, A., ... & Apollonio, M. (2012). Group specific vocal signature in free-ranging wolf packs. *Ethology Ecology & Evolution*, *24*(4), 322-331.

Zollinger, S. A., Podos, J., Nemeth, E., Goller, F., & Brumm, H. (2012). On the relationship between, and measurement of, amplitude and frequency in birdsong. *Animal Behaviour*, *84*(4), e1-e9.

Tables

Table 1. Study subjects (N = 13) with details on their age class, sex, pack and the number of howls collected in each context

Individual	Age class	Birth date	Sex	Pack	Total howls	Spontaneous howls	Elicited howls
AM	J	12 Apr	M	2	193	88	105
AR	A	8 May	M	1	59	26	33
CH	J	12 Apr	M	1	33	17	16
GE	A	9 May	M	2	149	93	56
KA	A	8 May	M	1	72	45	27
KAY*	J	12 Apr	F	2	15	9	6
KE	A	10 Apr	M	2	88	60	28
NA	A	9 Apr	M	3	74	21	53
SH*	A	8 May	F	1	92	6	86
TA*	J	12 Apr	F	1	46	12	34
UN	J	12 Apr	F	3	109	65	44
WA*	J	12 Apr	M	3	74	13	61
YU	A	9 May	F	3	136	33	103

A = Adult, J = Juvenile. M = Male, F = Female. Asterisks indicate individuals with fewer calls than number of acoustic parameters taken (<15); we excluded these from the analysis.

Table 2. List of acoustic measures extracted and used in analysis

Vocal parameter	Type	Definition
Duration	F0 (fundamental frequency)	Duration of the howl
Mean F0	F0	The mean of F0 values across the howl
F0 start	F0	The value of F0 at the start of the howl
F0 end	F0	The value of F0 at the end of the howl
Max F0	F0	The maximum value of F0 across the howl
Min F0	F0	The minimum value of F0 across the howl
% Time max F0	F0	The percentage of the total duration for which F0 was at maximum
F0 absolute slope*	F0	The mean absolute slope of F0
F0 var*	F0 variation	The mean F0 variation/s, calculated as the cumulative variation in the F0 contour in Hz divided by howl duration
FM extent*	F0 variation	The mean peak-to-trough variation of each F0 modulation (change in sign of the frequency gradient, see Charlton et al., 2009a)
FM rate*	F0 variation	The number of complete cycles (peak-to-trough-to-peak) of F0 modulation/s (Charlton et al., 2009a)
Jitter	F0 variation	The mean absolute difference between frequencies of consecutive F0 periods divided by mean F0 (Titze et al., 1987)
Shimmer	F0 variation	The mean absolute difference between the amplitudes of consecutive F0 periods divided by mean amplitude of F0
Q25%	Frequency	The frequency values at the upper limit of the first quartiles of energy, measured on a linear amplitude spectrum applied to the entire howl
Q50%	Frequency	The frequency values at the upper limit of the second quartiles of energy, measured on a linear amplitude spectrum applied to the entire howl
Q75%	Frequency	The frequency values at the upper limit of the third quartiles of energy, measured on a linear amplitude spectrum applied to the entire howl
Fpeak	Energy	The frequency with the highest power/energy of the howl
EfPeak	Energy	The maximum energy value of the frequency with highest power/energy of the howl
% EfPeak *	Energy	The percentage of the total howl duration where energy value of the frequency with the highest power/energy of the howl was maximum
% Time of max intensity	Intensity	The percentage of the total howl duration when the intensity was maximum
AM var*	Intensity	The mean variation/s of the intensity contour of the howl, calculated as the cumulative variation in amplitude divided by the howl duration
AM rate*	Intensity	The number of complete cycles of amplitude modulation/s of intensity contour of the howl
AM extent*	Intensity	The mean peak-to-peak variation of each amplitude modulation of the intensity contour of the howl (see Charlton et al., 2009a)

*Variable was removed from further analysis due to having a VIF greater than 10.

Table 3. Summary of pDFA details and outputs

pDFA type	Test factor	Control factor	Restriction factor	No. of individuals	No. of calls	Correctly cross-classified	Expected correctly cross-classified	*P*
Crossed	Context	Individual	None	9	913	62.38	53.33	0.009
Crossed	Individual	Context	None	9	913	38.75	26.46	0.003
Nested	Age class	Individual	Context	9	913	58.91	62.13	0.796
Nested	Pack	Individual	Context	9	913	49.56	53.10	0.870
Nested	Sex	Individual	Context	9	913	56.77	56.80	0.480

Table 4. Summary of outputs for individual identity PCA

	PC1
Eigenvalue	1.100
Proportion of variance	0.477
Factor loadings	
Duration	0.001
Mean F0	0.223
F0 start	0.321
F0 end	**0.418**
Max F0	-0.241
% Time Max F0	-0.107
Min F0	0.177
Q25%	**0.481**
Q50%	0.221
Q75%	0.003
Fpeak	0.056
EfPeak	0.093
% Time of max intensity	-0.097
Jitter	-0.352
Shimmer	-0.383

Bold indicates a factor loading of over 0.4.

Table 5. Summary of output for principal components that varied significantly between contexts

	PC2	PC3	PC4	PC6
Eigenvalue	1.705	1.639	1.208	1.012
Proportion of variance	0.114	0.109	0.081	0.067
Cumulative proportion	0.114	0.223	0.304	0.371
P	0.024	<0.001	0.019	<0.001
Factor loadings				
Duration	**-0.406**	0.094	0.067	0.391
Mean F0	0.203	0.358	0.137	-0.007
F0 start	0.002	**0.418**	-0.125	0.027
F0 end	0.047	0.201	**-0.615**	-0.181
Max F0	0.373	**0.434**	0.217	0.174
% Time Max F0	**0.402**	-0.303	-0.204	-0.317
Min F0	-0.095	0.149	**-0.502**	-0.198
Q25%	0.044	-0.321	0.067	-0.011
Q50%	0.135	-0.306	0.057	0.117
Q75%	0.196	-0.331	-0.154	0.173
Fpeak	**0.439**	0.012	-0.004	0.307
EfPeak	-0.186	0.117	0.241	**-0.516**
% Time of max intensity	0.179	0.037	0.357	**-0.474**
Jitter	0.373	0.114	0.025	-0.006
Shimmer	0.159	0.109	-0.157	0.121

Bold indicates a factor loading of over 0.4. *P* refers to the outcome of GLMM described above.

Figures

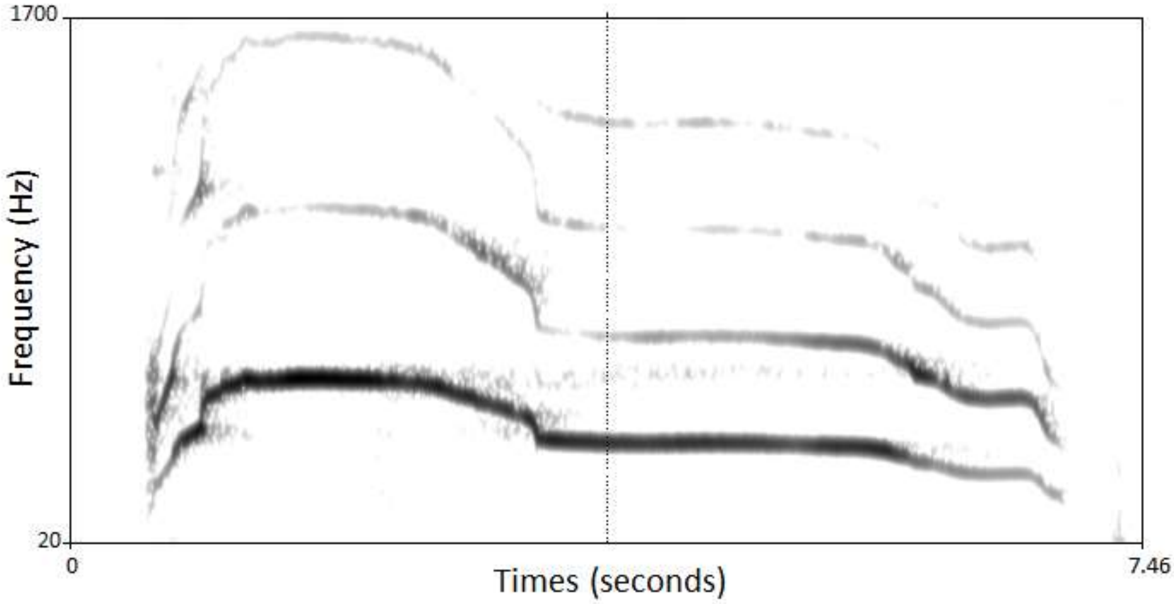

Figure 1. Example of a single howl spectral visualization. The fundamental frequency is the lowest thick band. Other measures extracted can be found in Table 2.

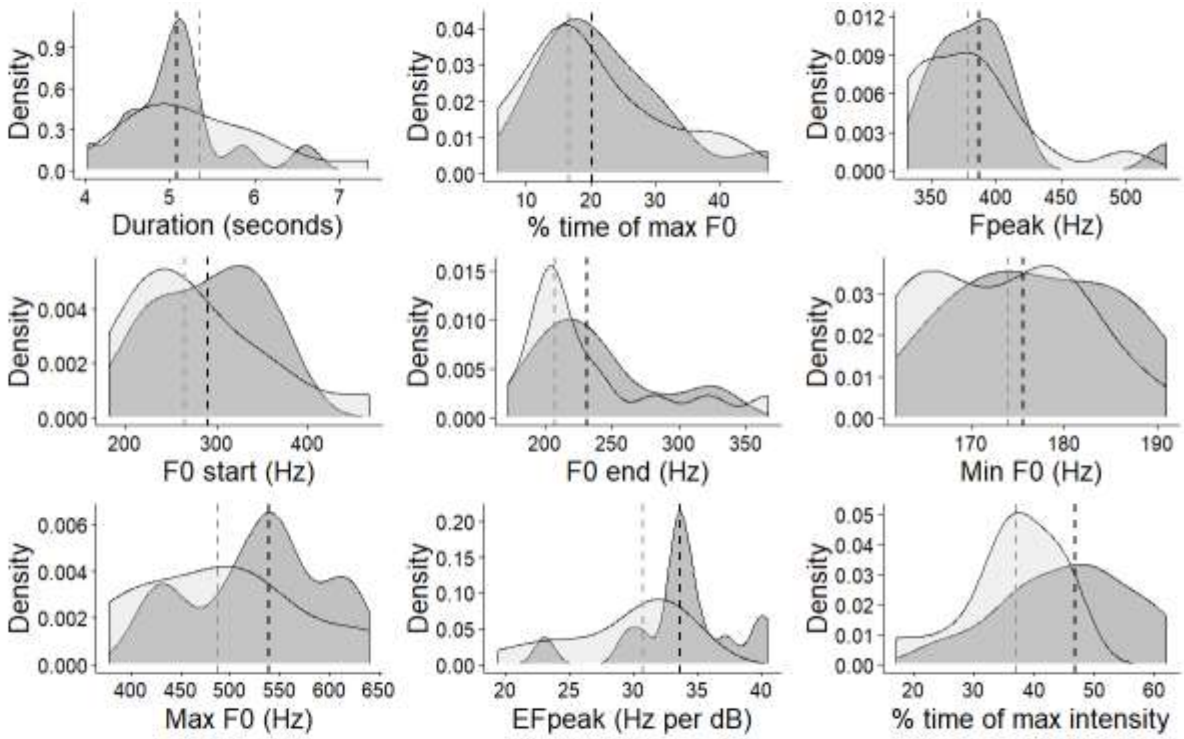

Figure 2. Density distributions for variables that had a loading greater than 0.4 in the context PCA. Light grey: spontaneous context; dark grey: elicited contexts. Dashed lines indicate the median value.