

1 TITLE:

2 Measuring acoustic complexity in continuously varying signals: how complex is a wolf howl?

3

4 AUTHORS

5 Arik Kershenbaum<sup>1</sup>, Éloïse C. Déaux<sup>2</sup>, Bilal Habib<sup>3</sup>, Brian Mitchell<sup>4</sup>, Vicente Palacios<sup>5</sup>, Holly Root-  
6 Gutteridge<sup>6</sup>, Sara Waller<sup>7</sup>

7 <sup>1</sup>Department of Zoology, University of Cambridge, UK

8 <sup>2</sup>Department of Biological Sciences, Macquarie University, Sydney, 2109, Australia

9 <sup>3</sup>Department of Animal Ecology and Conservation Biology, Wildlife Institute of India, Dehradun,  
10 India

11 <sup>4</sup>The Rubenstein School of Environment and Natural Resources, University of Vermont, USA

12 <sup>5</sup>Instituto Cavanilles de Biodiversidad y Biología Evolutiva. University of Valencia, Spain

13 <sup>6</sup>Department of Biology, Syracuse University, USA

14 <sup>7</sup>Montana State University, Bozeman, MT, USA

15

16 CORRESPONDING AUTHOR: Arik Kershenbaum, [arik.kershenbaum@gmail.com](mailto:arik.kershenbaum@gmail.com)

17

18 ABSTRACT

19

20 Communicative complexity is a key behavioural and ecological indicator in the study of animal  
21 cognition. Much attention has been given to measures such as repertoire size and syntactic structure in  
22 both bird and mammal vocalisations, as large repertoires and complex call combinations may give an  
23 indication of the cognitive abilities both of the sender and receiver. However, many animals  
24 communicate using a continuous vocal signal that does not easily lend itself to be described by  
25 concepts such as "repertoire". For example, dolphin whistles and wolf howls both have complex  
26 patterns of frequency modulation, so that no two howls or whistles are quite the same. Is there a sense  
27 in which some of these vocalisations may be more "complex" than others? Can we arrive at a  
28 quantitative metric for complexity in a continuously varying signal? Such a metric would allow us to  
29 extend familiar analyses of communicative complexity to those species where vocal behaviour is not  
30 restricted to sequences of stereotyped syllables. We present four measures of complexity in  
31 continuous signals (Wiener Entropy, Autocorrelation, Inflection Point Count, and Parsons Entropy),  
32 and examine their relevance using example data from members of the genus *Canis*. We show that  
33 each metric can lead to different conclusions regarding which howls could be considered complex or  
34 not. Ultimately, complexity is poorly defined and researchers must compare metrics to ensure that  
35 they reflect the properties for which the hypothesis is being tested.

36

37 KEYWORDS: Autocorrelation, Canids, Communication, Complexity, Entropy

38

39 INTRODUCTION

40

41 Vocal complexity is considered an important property of animal communication (Freeberg and Krams  
42 2015; Larson 2004; McCowan et al. 2002; Pollard and Blumstein 2012), despite being poorly defined,  
43 with little agreement how complexity should be quantified (Edmonds 1999). Despite this, complexity  
44 has been used to explain different aspects of animal behaviour. For instance, evidence exists in several  
45 bird species that females choose mates at least partially on the basis of the complexity of male song  
46 (Darolová et al. 2012; Hiebert et al. 1989), and other males may use complexity cues to make conflict  
47 escalation decisions (Leitão et al. 2006). It has been postulated that birdsong complexity acts as an  
48 index signal; being positively correlated with nutritional competence and cognitive abilities, and  
49 negatively correlated with early life stress (Nowicki et al. 2002). Similarly, recent studies have  
50 suggested that a correlation exists between communicative complexity and social complexity, such  
51 that species with more complex social systems also have more complex communicative interactions  
52 (Freeberg and Krams 2015; Krams et al. 2012; Pollard and Blumstein 2012). This in turn could shed  
53 light on possible evolutionary pathways to the development of language as an adaptation of highly  
54 complex social groups in early hominins (Seyfarth and Cheney 2014).

55 Communicative complexity can also have practical implications for the identification of and  
56 discrimination between similar sub-populations where complexity varies geographically (Briefer et al.  
57 2010; Kershenbaum et al. 2012). While many species can distinguish between the vocalisations of in-  
58 group and out-group individuals, e.g. wolves *Canis lupus* (Palacios et al. 2015; Zaccaroni et al. 2012),  
59 elephants *Loxodonta africana* (O'Connell-Rodwell et al. 2007), and multiple bird species (Nakagawa  
60 et al. 2001; Radford 2005), it is often not clear what vocal cues are being used to make this  
61 discrimination, and complexity characteristics may play a role (Briefer et al. 2008). Geographic  
62 differences in vocal complexity may be particularly noticeable where ecological conditions lead to  
63 differences in food availability, cognitive development, and hence vocal repertoire size (Byers and  
64 Kroodsma 2009; Kipper et al. 2006; Pfaff et al. 2007). In parallel, researchers can make use of  
65 differences in repertoire size, for example, to distinguish between sub-populations of birds and

66 mammals (Gwilliam et al. 2008; Pitcher et al. 2012). Clearly, vocal complexity is an important  
67 phenomenon with far ranging implications for the study of animal communication.

68 Previous studies of vocal complexity have focussed largely on birdsong, because of three essential  
69 properties that make this communication modality particularly tractable: (a) most birdsong can be  
70 divided into discrete syllables or notes (Marler and Slabbekoorn 2004); (b) there exists a simple  
71 metric - repertoire size - for measuring purported complexity (Byers and Kroodsma 2009); and (c) the  
72 well-established role of birdsong in mate choice provides the opportunity for manipulative as well as  
73 correlative experiments to be carried out, quite clearly demonstrating the role of song complexity in  
74 enhancing fitness (Searcy 1992). Even when birdsong is open-ended so that repertoire cannot  
75 adequately be defined, for example in the northern mockingbird *Mimus polyglottos* (Gammon 2014),  
76 the discrete nature of the song syllables means that other measurements of communicative complexity  
77 can be used, most notably Shannon entropy (Briefer et al. 2010; Da Silva et al. 2000; Kershenbaum  
78 2013). Although some animals from other taxa have vocal communication systems that are similarly  
79 discrete, e.g. rock hyrax *Procapra capensis* (Kershenbaum et al. 2012), or are closed-ended, e.g.  
80 several primates (Cäsar et al. 2012; Zuberbühler 2002), outside of passerine birds they are the  
81 exception rather than the rule. Indeed, the vocal communication system of some species consists  
82 entirely of signals whose properties are continuously varying, and with such signals the existing  
83 concepts of complexity (e.g. based on repertoire and entropy) cannot be applied. For example,  
84 considerable empirical evidence supports the existence of semantic information in the whistles of  
85 bottlenose dolphins *Tursiops truncatus*; in particular, the use of signature whistles to signal individual  
86 identity (Fripp et al. 2005; Kershenbaum et al. 2013; Quick and Janik 2012; Sayigh et al. 1999).

87 However, dolphin whistles are relatively long and unbroken tonal signals that are continuously  
88 frequency modulated. Similarly, wolf howls are continuous frequency modulated signals that have  
89 been shown to contain individual identity information (Palacios et al. 2007; Root-Gutteridge et al.  
90 2014; Tooze et al. 1990), but cannot be classified into stereotyped categories. One possible  
91 interpretation of complexity, which may or may not be intuitive to the reader, is that the frequency of  
92 a *simple* signal varies little, or predictably, with time, whereas a *complex* signal varies greatly, and  
93 unpredictably. The reader may examine Figure 1, which shows several examples of wolf howls and

94 decide which howls are simple and which complex. However, although some howls may appear  
95 intuitively more complex than others, the lack of an objective definition of complexity for such  
96 signals renders the judgement unhelpful. As such, previous studies of information content in  
97 continuously varying animal vocal signals have had to make use of alternative techniques, example  
98 for measuring similarity between pairs of vocalisations, rather than quantifying characteristics of the  
99 vocalisations themselves (Kershenbaum et al. 2013).

100 Complexity in itself is poorly defined (Edmonds 1999), and as a result any particular use of the term  
101 in examining animal behaviour is liable to be criticised. In particular, information theoretical  
102 definitions of complexity based on concepts such as entropy are often regarded with suspicion by  
103 biologists, because the most complex (highest entropy) signals are in fact random signals - something  
104 that most ethologists would consider to be non-complex (Suzuki et al. 2005). We agree that defining  
105 complexity is difficult, however we hope to mitigate this difficulty somewhat by insisting that  
106 researchers should examine and understand what kind of signal a particular definition of complexity  
107 would deem either complex or non-complex. It would then be possible to determine whether a  
108 particular definition of complexity meets the demands of discriminating between signals in relation to  
109 the study's hypotheses.

110 In this paper, we aim to show how concepts of continuous complexity can be measured using different  
111 approaches, and we illustrate how use of a particular complexity metric can lead to different  
112 conclusions from using other metrics. We present four candidate complexity metrics and compare  
113 their performance against each other, identifying which kinds of signals each particular metric would  
114 indicate to be "complex" or "not complex". We recognise that in drawing up a proposed metric of  
115 complexity for continuous signals, it is inevitable that subjective interpretations of the complexity or  
116 simplicity of a signal must necessarily influence the decisions a researcher makes in designing an  
117 experiment. Rather than attempt to avoid this subjective tendency, we hope to formalise it somewhat,  
118 by presenting a range of complexity metrics, along with illustration of their significance for signals of  
119 different types. That way, researchers can choose a metric that captures the "kind of" complexity for  
120 which they are searching; providing a quantitative measure for an essentially subjective property.

121

122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147

## METHODS

We describe below four quantitative metrics that have been previously used for quantifying complexity in continuous signals: Wiener Entropy, Autocorrelation, Inflection Point Count, and Parsons Entropy. For each of these metrics, we define how they are computed and explain in what way they could be considered to be measures of complexity in a continuous signal, giving detailed examples using simulated waveforms as shown in Figure 2. As each of our proposed metrics measures quite different features of acoustic signals, it is constructive to illustrate the behaviour of each metric by showing examples of signals with both high and low metric values, as this provides an indication of which signal features are being emphasised by each metric. We do this using example acoustic signals drawn from an empirical data set consisting of 2,014 coded canid howls from 16 different species and subspecies, as described in a previous work (Kershenbaum et al. 2016). We calculated each complexity metric for each howl and then plot on a time-frequency graph the three howls with the lowest metric values, and the three howls with the highest values.

### *Wiener Entropy*

Wiener Entropy, or spectral flatness, has been proposed as a measure of the complexity of birdsong elements, and has been applied to the analysis of increasing complexity during the process of song learning in juvenile birds (Baker and Logue 2003; Tchernichovski et al. 2000). Wiener Entropy is a measure of the extent to which a signal contains a mixture of frequencies, as opposed to a single frequency or tone. As such, a value of 0 represents a single sine wave, and a value of 1 indicates white noise, in which all frequencies are equally represented. To calculate the Wiener Entropy, we first perform a Fourier transform to calculate the spectral power  $P$  present at each of  $N$  distinct frequency bins. The formal definition is given as the ratio of the geometric means of the spectral powers to the arithmetic mean:

148

$$WE = \frac{\sqrt{\prod_{f=1}^N P(f)}}{\frac{1}{N} \sum_{f=1}^N P(f)}$$

149 Wiener Entropy can be applied to the signal spectrum, where  $P(f)$  corresponds to the FFT of the input  
150 waveform (column C in Figure 2), or to the spectrographic representation of the signal, where  $P(f)$   
151 corresponds to  $F(t)$  (column B in Figure 2). The former definition also measures the entropic  
152 contribution of background noise and discards any information on temporal variation in frequency,  
153 therefore, we calculate Wiener Entropy only on the dominant signal frequency, i.e. column B in  
154 Figure 2. In addition, we square-root transform this metric for normality.

155

### 156 *Autocorrelation*

157 Autocorrelation (Figure 2, column D) measures the self-similarity of a signal, and so quantifies the  
158 extent to which the signal contains repetitions of the same pattern, or is varied without similarity  
159 (Stoica and Moses 2005). The autocorrelation sequence of a signal consisting of  $N$  discrete samples,  
160  $F(1...N)$  is measured by shifting the signal by time lag  $l$  and calculating the correlation between the  
161 shifted and the unshifted signals:

162

$$ac(l) = \sum_{n=0}^N F(n) \cdot F(n - l)$$

163 The Autocorrelation metric is then calculated as the sum of  $ac(l)$  for all  $l$ . As with Weiner entropy, we  
164 measure the repetitiveness of the howl frequency modulation, rather than flatness *per se*, by  
165 calculating autocorrelation on the dominant signal frequency in the spectrogram. In addition, we log-  
166 transform this metric for normality.

167

### 168 *Inflection points*

169 A number of studies, particularly with cetaceans, have measured vocal complexity by counting the  
170 number of inflection points in a vocalisation (Janik et al. 1994; May-Collado and Wartzok 2008). A  
171 more complex signal, in this context, is a signal in which the frequency is changing direction  
172 (rising/falling) often (e.g. Figure 2, panel 4B). In keeping with these studies, we define the number of

173 inflection points as the count of changes in gradient direction of the vocalisation. To ensure that this  
174 complexity metric takes continuous values, we divide the number of inflection points by the length of  
175 the signal, which also standardises the metric, to remove the correlation between vocalisation length  
176 and complexity that would otherwise be present. In addition, we square-root transformed this metric  
177 for normality. Simplified Matlab code for counting the number of inflection points in a vector **X** is  
178 shown below:

```
179         1. g=gradient(X) ;  
180         2. s=sign(g) ;  
181         3. d=g(1:end-1)~=g(2:end) ;  
182         4. C=sum(d) ;
```

183 The above algorithm (1) measures the gradient at each point, (2) determines the sign of the gradient  
184 (positive, increasing, or negative, decreasing), (3) tests whether the sign of the gradient at this point is  
185 different from the sign of the gradient at the next point, which would indicate a change in direction,  
186 and (4) counts the number of such changes in direction.

187

### 188 *Parsons Entropy*

189 The Parsons code is a reduced representation of a varying signal, used primarily for music retrieval  
190 systems (Downie 2003; Parsons and Levin 1975). However, a recent study showed that dolphin  
191 signature whistles can be represented as Parsons codes, while maintaining much of the individual  
192 identity information (Kershenbaum et al. 2013). To convert a continuous signal to a Parsons code, we  
193 divide the signal into a fixed number of slices, and record whether the frequency from one slice to the  
194 next is rising, falling, or remaining constant. To increase the descriptive power of the Parsons code,  
195 we can distinguish between rises and falls of different magnitude, classifying the 10% of the largest  
196 magnitude changes as “big rise” or “big drop” (Müllensiefen and Frieler 2004; Pauws 2002), and thus  
197 dividing changes in pitch into seven categories: big rise, medium rise, small rise, no change, small  
198 drop, medium drop, big drop (Figure 3). The implementation of this coding is described more fully in  
199 (Kershenbaum et al. 2013). The continuous signal has now been converted into a series of discrete  
200 characters from a finite alphabet of seven step categories, and so we can then calculate the simple



201 Shannon entropy (Cover and Thomas 1991) as is often done with stereotyped signals, where  $P(n)$  is  
202 the probability of occurrence of step category  $n$ . We refer to this entropy metric that measures the  
203 unpredictability of the Parsons code as the “Parsons Entropy”.

$$204 \quad PE = - \sum_{n=1}^7 P(n) \log_7 P(n)$$

205  
206 As each of the four metrics Wiener Entropy, Autocorrelation, Inflection Point Count, and Parsons  
207 Entropy are all purported to measure the same property - signal complexity - the metrics may  
208 potentially measure similar features of the acoustic signals. Therefore, we test directly for correlation  
209 between the different metrics by calculating the Pearson's correlation coefficient between each pair of  
210 metrics. All analyses were carried out in Matlab R2014b (The Mathworks, Inc).

211

212

## 213 RESULTS

214

### 215 *Complexity of simulated waveforms*

216 The values of the four metrics for each of the five waveforms in Figure 2 are shown in Figure 4. All  
217 metrics gave the constant frequency (1) the lowest score, representing the simplest waveform. The  
218 random waveform (5) received the highest scores from three of the metrics, Wiener Entropy,  
219 Inflection Point Count, and Parsons Entropy, indicating (as with traditional entropy measures) that  
220 randomness is interpreted as high complexity. Autocorrelation was low for the random waveform.  
221 The oscillating frequency (4) scored highly for complexity with all metrics (especially  
222 Autocorrelation and Parsons Entropy), consistent with the intuitive interpretation of this as a complex  
223 signal. However, the two-frequency waveform (2) received similar Wiener Entropy and  
224 Autocorrelation scores to constantly increasing frequency (3), but notably higher scores than (3) for  
225 Inflection Point Count and Parsons Entropy. Indeed, both Inflection Point Count and Parsons Entropy  
226 considered the constantly increasing frequency (3) to be approximately as complex as the constant  
227 frequency (1).

228

229 *Complexity of wolf howls*

230 Examples of howls with the lowest, and highest values of each of the complexity metrics are shown in  
231 Figure 5. Wiener Entropy and Autocorrelation illustrate the difficulty of traditional metrics as  
232 indicators of complexity. It is not clear that the howls with the lowest Wiener Entropy or  
233 Autocorrelation are necessarily less complex by subjective interpretation than those with high Wiener  
234 Entropy or Autocorrelation. A low Wiener Entropy score is achieved by a signal possessing a single  
235 frequency (i.e. approximating a part of a sine wave), whereas high Wiener Entropy score is achieved  
236 by flat signals, which are transformed by FFT to a mixture of a large number of frequencies, and  
237 hence high entropy. Neither appear to be particularly complex by intuitive definition. Low  
238 Autocorrelation scores are achieved by irregular but not repetitive signals, and such irregularity is a  
239 promising trait of complexity; however high Autocorrelation scores are achieved by signals with a  
240 single frequency, which does not appear to be either complex or simple.  
241 In contrast, the number of inflection points seems an intuitive measure of complexity, as high  
242 Inflection Point Count howls are very varied, whereas low Inflection Point Count howls appear  
243 simpler. However, some howls with low Inflection Point Count still have considerable variation in  
244 frequency. Parsons Entropy also detects subjectively complex howls, and those with low Parsons  
245 Entropy appear subjectively simple; specifically, the frequency of these simple howls varies  
246 monotonically.

247

248 *Colinearity among complexity metrics*

249 Weak correlations existed between all metrics (Figure 6), however, a stronger negative correlation  
250 was found between Wiener Entropy and Parsons Entropy ( $R=-0.38$ ). Taken with the other differences  
251 found between metrics, there does not appear to be grounds for describing any pair of metrics as co-  
252 varying, and the metrics appear to be measuring different aspects of complexity.

253

254

255 DISCUSSION

256

257 We have illustrated the similarities and differences between four different metrics, each of which  
258 could be considered a quantitative measure of complexity in a continuously varying signal. Despite  
259 the poorly defined nature of signal complexity, we have provided the reader with both quantitative  
260 comparisons, and qualitative illustrations of the result of using each of these metrics in the evaluation  
261 of the complexity of simulated signals and natural of canid howls.

262 All metrics distinguished clearly between a constant flat frequency, and a randomly varying signal,  
263 with all metrics except Autocorrelation placing the random signal at the most complex end of the  
264 quantitative scale. Researchers will need to consider whether or not the characterisation of a random  
265 signal as "complex" (a definition taken from the field of entropy and information theory) is consistent  
266 with the hypotheses that they are testing. In contrast to Wiener Entropy, Inflection Point Count, and  
267 Parsons Entropy, Autocorrelation gave a higher score to a regularly varying signal than to the random  
268 one. Parsons Entropy also gave a much higher value for a regularly varying signal than for the flat,  
269 rising, and random frequencies, indicating that Parsons Entropy, which measures changes in slope, is  
270 a good metric for measuring the extent to which a signal changes with time - either regularly or  
271 irregularly. When examining actual howls qualitatively, both Inflection Point Count and Parsons  
272 Entropy appeared to distinguish between howls that the authors felt looked "simple" (i.e. varying in  
273 frequency in a constant way) and those that looked "complex" (i.e. varying in an inconsistent way  
274 with time), although clearly this subjective distinction may not be globally applicable. When  
275 examining the example howls for low and high Wiener Entropy and Autocorrelation, there did not  
276 appear to be as much of a subjective difference in complexity. However, the essence of these results is  
277 to provide the comparison, rather than to impose subjective conclusions, and Figure 5 makes this  
278 comparison clearly.

279 Despite the fact that complexity in vocal signals of any kind remains poorly defined (Kershenbaum  
280 2013), the concept of complexity is still widely used for investigating questions of proximal behaviour  
281 (Darolová et al. 2012; Demartsev et al. 2014; Gustison and Bergman 2016), ultimate fitness (Freeberg  
282 et al. 2012; Ord et al. 2012; Pollard and Blumstein 2012), and drivers of the evolution of social  
283 systems (Bergman and Beehner 2015; Freeberg and Krams 2015; Krams et al. 2012). Even if we

284 accept the definition of entropy as a proxy for complexity (Doyle et al. 2008), it is not clear how such  
285 a metric can be applied to continuously varying signals. We have shown that multiple approaches are  
286 possible, each with its benefits and disadvantages. For example, counting the number of inflection  
287 points is a useful method for identifying highly frequency modulated signals, but can become  
288 overwhelmed in the presence of a highly random signal. Parsons Entropy may suffer less from this  
289 constraint, as the signal is divided into discrete segments. Wiener Entropy measures the noisiness of a  
290 signal, but can misinterpret a simple upsweep as complex as it contains many frequencies, albeit  
291 spread through time. Autocorrelation is a powerful tool for detecting repetition, but returns a low  
292 value for asymmetric changes in frequency.

293 We have no objective gold-standard of complexity to compare to our metrics and to indicate whether  
294 a particular metric truly captures the property of complexity or not. Yet we believe that our study has  
295 merit precisely because it allows quantification of the subjective measure for which researchers may  
296 be searching. If an oscillating signal is the nature of complexity being tested, then Inflection Point  
297 Count or Parsons Entropy may be the best metric to use. However, if for a particular hypothesis,  
298 randomness is best rejected as not complex, then perhaps Autocorrelation is a better-suited metric. It  
299 is vital for researchers to understand the implications of their choice of a particular complexity metric,  
300 rather than to make use of a metric whose properties may be unknown, and perhaps surprising. Our  
301 approach of defining a quantitative metric also has the advantage of enabling clearer comparisons  
302 between different study systems that may use similar but non-quantitative assessments of vocalisation  
303 type. Qualitative descriptions of frequency modulations in continuous signals, e.g. "flat", "rising",  
304 "step up" (Hallberg 2007; Palacios et al. 2007) are useful, but difficult to compare between studies.  
305 Vocal signal complexity is likely an important property in the communication of many species,  
306 including birds (Briefer et al. 2010; Darolová et al. 2012; Freeberg 2008; Kipper et al. 2006; Leitão et  
307 al. 2006), amphibians (Larson 2004; Narins and Capranica 1978), terrestrial mammals (Demartsev et  
308 al. 2014; Gustison and Bergman 2016; Schlenker et al. in press) and cetaceans (Doyle et al. 2008;  
309 Ferrer-i-Cancho and McCowan 2009; Garland et al. 2013; Nash and Bowles 2011), and also human  
310 language (Ferrer-i-Cancho 2005; Ferrer-i-Cancho and Solé 2003; Kershenbaum et al. 2014;  
311 Montemurro and Zarette 2015). Complexity is a difficult property to measure, particularly when the

312 signals are continuously varying, rather than a sequence of discrete notes. To investigate questions  
313 such as the connection between social complexity and vocal complexity, appropriate measures of  
314 complexity must be found (Freeberg and Krams 2015; Kershenbaum 2013). Species with  
315 continuously varying vocalisations, such as wolves and dolphins, share many of the properties of  
316 particular interest in the investigation of the evolution of communication. Wolves and dolphins in  
317 particular are highly social and cooperative, as well as having intelligent problem-solving abilities  
318 (wolves: Mech and Boitani 2010; Mech et al. 2015, dolphins: Gazda et al. 2005; King and Janik 2015;  
319 Wells 2003), and so it would be unfortunate if a quantitative assessment of their vocal complexity  
320 were neglected.

321 We have shown that multiple metrics do exist for capturing the complexity of these vocalisations, and  
322 we have provided quantitative tools to assess the suitability of the different types of metrics. We  
323 encourage researchers to make use of such quantitative measures when testing hypotheses in these  
324 and similar species, thereby extending the investigation of complex communication in animals to  
325 those species not previously amenable to quantitative analysis.

326

327

## 328 ACKNOWLEDGEMENTS

329

330 AK thanks Sheela Hira and Knoxville Zoo, Laura Pearson and Asheville Zoo, and Chris Lasher and  
331 North Carolina Zoo for access to record their red wolves, and Arthur Povey for assistance coding the  
332 data. Recording work was approved by the Institutional Animal Care and Use Committee of the  
333 University of Tennessee. AK is supported by a Herchel Smith postdoctoral fellowship at the  
334 University of Cambridge. Part of this work was carried out while AK was a Postdoctoral Fellow at the  
335 National Institute for Mathematical and Biological Synthesis, an Institute sponsored by the National  
336 Science Foundation through NSF Award #DBI-1300426, with additional support from The University  
337 of Tennessee, Knoxville. BH is thankful to the State Forest Departments of Himachal Pradesh, J&K,  
338 and Maharashtra, and to various zoos in India for permitting us to record howls. HRG is grateful to all  
339 who helped with the project: the staff at Colchester Zoo; the Wildwood Trust, the Borror Laboratory

340 of Bioacoustics; the British Library; Lupus Laetus; Polish Mammal Research Institute; Tigress  
341 Productions; the BBC Natural History Unit; Longleat Safari Park; Tierstimmen Archiv; Wild Sweden;  
342 Wolf Park; the Macaulay Sound Library and the UK Wolf Conservation Trust; and Mike Collins,  
343 Teresa Palmer, Monty Sloan, Karl-Heinz Frommolt, Yorgos Iliopoulos, Christine Anhalt, Louise  
344 Gentle, Richard Yarnell, Victoria Allison Hughes and Susan Parks. BRM thanks the  
345 USDA/APHIS/WS/National Wildlife Research Center for supporting his doctoral research and  
346 providing access to captive coyotes; recording work was approved by the NWRC IACUC. SW thanks  
347 Mariana Olsen for assistance with data collection, and Yellowstone National Park for permission to  
348 record. ECD thanks Amanda and Gavin McDowell from Colong Station, NSW, Australia for  
349 permission to record their animals.

350

351

352

353

354

355

## References

- 356 Baker MC and Logue DM. 2003. Population differentiation in a complex bird sound: A comparison of  
357 three bioacoustical analysis procedures. *Ethology* 109(3):223-42.
- 358 Bergman TJ and Beehner JC. 2015. Measuring social complexity. *Anim Behav* 103:203-9.
- 359 Briefer E, Osiejuk TS, Rybak F, Aubin T. 2010. Are bird song complexity and song sharing shaped  
360 by habitat structure? An information theory and statistical approach. *J Theor Biol* 262(1):151-64.
- 361 Briefer E, Aubin T, Lehongre K, Rybak F. 2008. How to identify dear enemies: The group signature  
362 in the complex song of the skylark *Alauda arvensis*. *J Exp Biol* 211(3):317-26.
- 363 Byers BE and Kroodsma DE. 2009. Female mate choice and songbird song repertoires. *Anim Behav*  
364 77(1):13-22.
- 365 Cäsar C, Byrne RW, Hoppitt W, Young RJ, Zuberbühler K. 2012. Evidence for semantic  
366 communication in titi monkey alarm calls. *Anim Behav* 84(2):405-11.
- 367 Cover TM and Thomas JA. 1991. *Elements of information theory*. New York, NY: John Wiley &  
368 Sons, Inc.
- 369 Da Silva ML, Piqueira JRC, Vielliard JME. 2000. Using Shannon entropy on measuring the  
370 individual variability in the rufous-bellied thrush *Turdus rufiventris* vocal communication. *J*  
371 *Theor Biol* 207(1):57-64.
- 372 Darolová A, Krištofík J, Hoi H, Wink M. 2012. Song complexity in male marsh warblers: Does it  
373 reflect male quality? *Journal of Ornithology* 153(2):431-9.
- 374 Demartsev V, Kershenbaum A, Ilany A, Barocas A, Ziv EB, Koren L, Geffen E. 2014. Male hyraxes  
375 increase song complexity and duration in the presence of alert individuals. *Behav Ecol* 25:1451-  
376 8.
- 377 Downie JS. 2003. Music information retrieval. *Annual Review of Information Science and*  
378 *Technology* 37(1):295-340.
- 379 Doyle LR, McCowan B, Hanser SF, Chyba C, Bucci T, Blue JE. 2008. Applicability of information  
380 theory to the quantification of responses to anthropogenic noise by southeast Alaskan humpback  
381 whales. *Entropy* 10(2):33-46.
- 382 Edmonds B. 1999. What is complexity? The philosophy of complexity per se with application to some  
383 examples in evolution. In: *The evolution of complexity*. Heylighen F and Aerts D, editors.  
384 Dordrecht: Kluwer.
- 385 Ferrer-i-Cancho R. 2005. The variation of Zipf's law in human language. *The European Physical*  
386 *Journal B-Condensed Matter and Complex Systems* 44(2):249-57.
- 387 Ferrer-i-Cancho R and McCowan B. 2009. A law of word meaning in dolphin whistle types. *Entropy*  
388 11(4):688-701.
- 389 Ferrer-i-Cancho R and Solé RV. 2003. Least effort and the origins of scaling in human language. *Proc*  
390 *Natl Acad Sci USA* 100(3):788-91.

- 391 Freeberg TM, Dunbar RIM, Ord TJ. 2012. Social complexity as a proximate and ultimate factor in  
392 communicative complexity. *Philos Trans R Soc Lond B Biol Sci* 367(1597):1785-801.
- 393 Freeberg TM. 2008. Complexity in the chick-a-dee call of carolina chickadees (*Poecile carolinensis*):  
394 Associations of context and signaler behavior to call structure. *Auk* 125(4):896-907.
- 395 Freeberg TM and Krams I. 2015. Does social complexity link vocal complexity and cooperation?  
396 *Journal of Ornithology* 156:1-8.
- 397 Fripp D, Owen C, Quintana-Rizzo E, Shapiro A, Buckstaff K, Jankowski K, Wells R, Tyack P. 2005.  
398 Bottlenose dolphin (*Tursiops truncatus*) calves appear to model their signature whistles on the  
399 signature whistles of community members. *Anim Cogn* 8(1):17-26.
- 400 Gammon DE. 2014. Seasonal patterns of vocal mimicry in northern mockingbirds *Mimus polyglottos*.  
401 *J Avian Biol* 45(6):545-50.
- 402 Garland EC, Noad MJ, Goldizen AW, Lilley MS, Rekdahl ML, Garrigue C, Constantine R, Hauser  
403 ND, Poole MM, Robbins J. 2013. Quantifying humpback whale song sequences to understand  
404 the dynamics of song exchange at the ocean basin scale. *J Acoust Soc Am* 133:560-9.
- 405 Gazda SK, Connor RC, Edgar RK, Cox F. 2005. A division of labour with role specialization in  
406 group-hunting bottlenose dolphins (*Tursiops truncatus*) off Cedar Key, Florida. *Proc Biol Sci*  
407 272(1559):135-40.
- 408 Gustison ML and Bergman TJ. 2016. Vocal complexity influences female responses to gelada male  
409 calls. *Scientific Reports* 6:19680.
- 410 Gwilliam J, Charrier I, Harcourt RG. 2008. Vocal identity and species recognition in male Australian  
411 sea lions, *Neophoca cinerea*. *J Exp Biol* 211(Pt 14):2288-95.
- 412 Hallberg KI. 2007. Information in a long-distance vocal signal: Chorus howling in the coyote (*Canis*  
413 *latrans*). PhD thesis. The Ohio State University.
- 414 Hiebert SM, Stoddard PK, Arcese P. 1989. Repertoire size, territory acquisition and reproductive  
415 success in the song sparrow. *Anim Behav* 37:266-73.
- 416 Janik VM, Todt D, Dehnhardt G. 1994. Signature whistle variations in a bottlenosed dolphin,  
417 *Tursiops truncatus*. *Behav Ecol Sociobiol* 35(4):243-8.
- 418 Kershenbaum A. 2013. Entropy rate as a measure of animal vocal complexity. *Bioacoustics* 23:195-  
419 208.
- 420 Kershenbaum A, Sayigh LS, Janik VM. 2013. The encoding of individual identity in dolphin  
421 signature whistles: How much information is needed? *PLoS One* 8(10):e77671.
- 422 Kershenbaum A, Ilany A, Blaustein L, Geffen E. 2012. Syntactic structure and geographical dialects  
423 in the songs of male rock hyraxes. *Proc R Soc Lond B Biol Sci* 279(1740):2974-81.
- 424 Kershenbaum A, Bowles AE, Freeberg TM, Jin DZ, Lameira AR, Bohn K. 2014. Animal vocal  
425 sequences: Not the Markov chains we thought they were. *Proceedings of the Royal Society B:*  
426 *Biological Sciences* 281(1792):20141370.



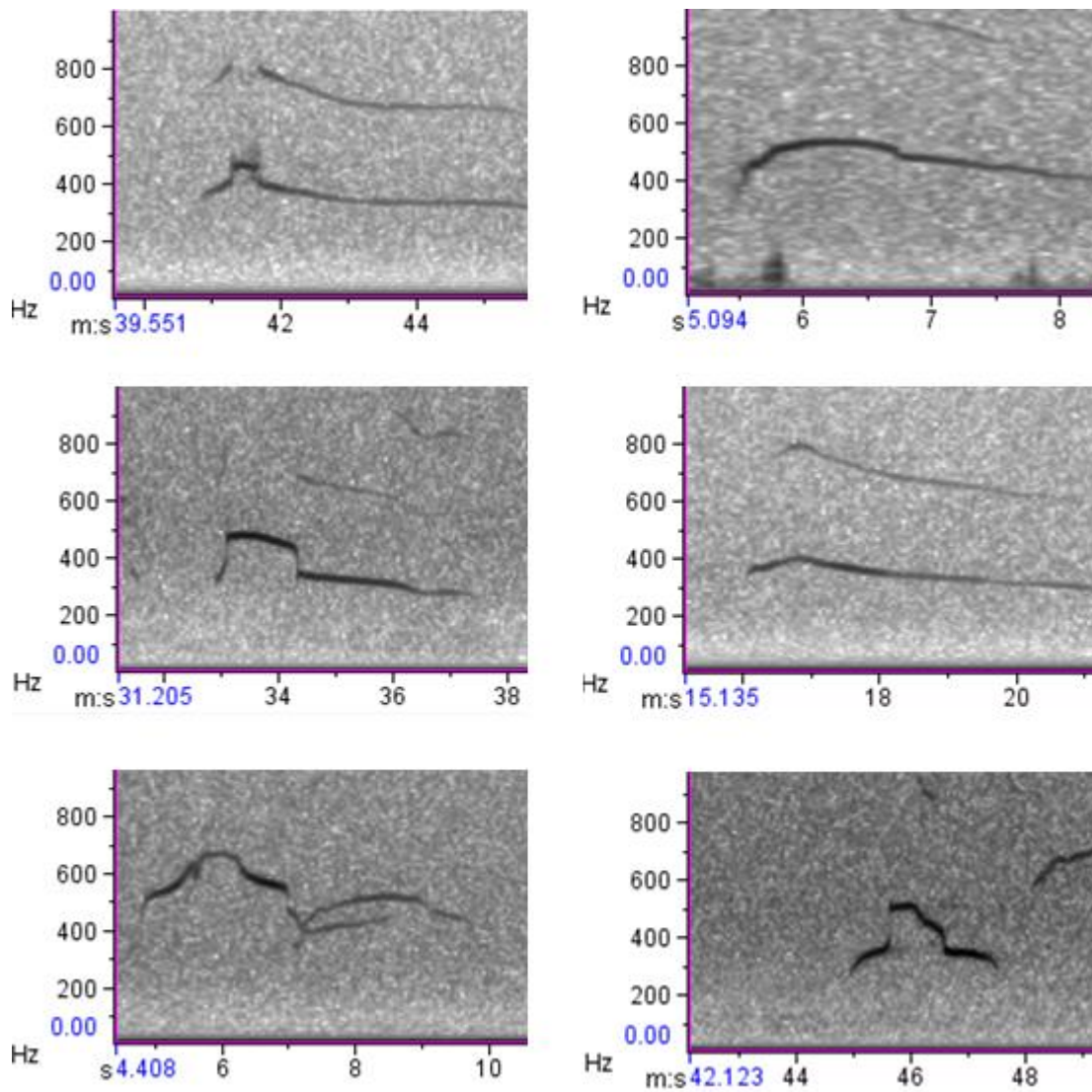
- 427 Kershenbaum A, Root-Gutteridge H, Habib B, Koler-Matznick J, Mitchell B, Palacios V, Waller S.  
428 2016. Disentangling canid howls across multiple species and subspecies: Structure in a complex  
429 communication channel. *Behav Processes* 124:149-57.
- 430 King SL and Janik VM. 2015. Come dine with me: Food-associated social signalling in wild  
431 bottlenose dolphins (*Tursiops truncatus*). *Animal Cognition* 18:1-6.
- 432 Kipper S, Mundry R, Sommer C, Hultsch H, Todt D. 2006. Song repertoire size is correlated with  
433 body measures and arrival date in common nightingales, *Luscinia megarhynchos*. *Anim Behav*  
434 71(1):211-7.
- 435 Krams I, Krama T, Freeberg TM, Kullberg C, Lucas JR. 2012. Linking social complexity and vocal  
436 complexity: A parid perspective. *Philosophical Transactions of the Royal Society B: Biological*  
437 *Sciences* 367(1597):1879-91.
- 438 Larson KA. 2004. Advertisement call complexity in northern leopard frogs, *Rana pipiens*. *Copeia*  
439 2004(3):676-82.
- 440 Leitão A, Ten Cate C, Riebel K. 2006. Within-song complexity in a songbird is meaningful to both  
441 male and female receivers. *Anim Behav* 71(6):1289-96.
- 442 Marler P and Slabbekoorn HW. 2004. *Nature's music: The science of birdsong*. Academic Press.
- 443 May-Collado LJ and Wartzok D. 2008. A comparison of bottlenose dolphin whistles in the Atlantic  
444 Ocean: Factors promoting whistle variation. *J Mammal* 89(5):1229-40.
- 445 McCowan B, Doyle LR, Hanser SF. 2002. Using information theory to assess the diversity,  
446 complexity, and development of communicative repertoires. *J Comp Psychol* 116(2):166-72.
- 447 Mech LD and Boitani L. 2010. *Wolves: Behavior, ecology, and conservation*. University of Chicago  
448 Press.
- 449 Mech LD, Smith DW, MacNulty DR. 2015. *Wolves on the hunt: The behavior of wolves hunting wild*  
450 *prey*. University of Chicago Press.
- 451 Montemurro MA and Zanette DH. 2015. Complexity and universality in the long-range order of  
452 words. *arXiv Preprint arXiv:1503.01129* .
- 453 Müllensiefen D and Frieler K. 2004. Cognitive adequacy in the measurement of melodic similarity:  
454 Algorithmic vs. human judgments. *Computing in Musicology* 13:147-76.
- 455 Nakagawa S, Waas JR, Miyazaki M. 2001. Heart rate changes reveal that little blue penguin chicks  
456 (*Eudyptula minor*) can use vocal signatures to discriminate familiar from unfamiliar chicks.  
457 *Behav Ecol Sociobiol* 50(2):180-8.
- 458 Narins PM and Capranica RR. 1978. Communicative significance of the two-note call of the treefrog  
459 *Eleutherodactylus coqui*. *Journal of Comparative Physiology* 127(1):1-9.
- 460 Nash JS and Bowles AE. 2011. N3 call types produced long-term by a killer whale of the northern  
461 resident community under controlled conditions: Characteristics, variation, and behavioral  
462 context. *J Acoust Soc Am* 130(4):2358.

- 463 Nowicki S, Searcy W, Peters S. 2002. Brain development, song learning and mate choice in birds: A  
 464 review and experimental test of the "nutritional stress hypothesis". *Journal of Comparative*  
 465 *Physiology A* 188(11-12):1003-14.
- 466 O'Connell-Rodwell CE, Wood JD, Kinzley C, Rodwell TC, Poole JH, Puria S. 2007. Wild African  
 467 elephants (*Loxodonta africana*) discriminate between familiar and unfamiliar conspecific seismic  
 468 alarm calls. *J Acoust Soc Am* 122:823.
- 469 Ord TJ, Garcia-Porta J, Ord TJ, Garcia-Porta J. 2012. Is sociality required for the evolution of  
 470 communicative complexity? Evidence weighed against alternative hypotheses in diverse  
 471 taxonomic groups. *Philos Trans R Soc Lond B Biol Sci* 367(1597):1811-28.
- 472 Palacios V, Font E, Márquez R, Carazo P. 2015. Recognition of familiarity on the basis of howls: A  
 473 playback experiment in a captive group of wolves. *Behaviour* 152(5):593-614.
- 474 Palacios V, Font E, Márquez R. 2007. Iberian wolf howls: Acoustic structure, individual variation,  
 475 and a comparison with north American populations. *J Mammal* 88(3):606-13.
- 476 Parsons D and Levin B. 1975. *The directory of tunes and musical themes*. Cambridge: S. Brown.
- 477 Pauws S. 2002. Cuby hum: A fully operational query-by-humming system. ISMIR 2002 conference  
 478 proceedings, IRCAM, 2002.
- 479 Pfaff JA, Zanette L, MacDougall-Shackleton SA, MacDougall-Shackleton EA. 2007. Song repertoire  
 480 size varies with HVC volume and is indicative of male quality in song sparrows (*Melospiza*  
 481 *melodia*). *Proceedings of the Royal Society B: Biological Sciences* 274(1621):2035-40.
- 482 Pitcher BJ, Harcourt RG, Charrier I. 2012. Individual identity encoding and environmental constraints  
 483 in vocal recognition of pups by Australian sea lion mothers. *Anim Behav* 83(3):681-90.
- 484 Pollard KA and Blumstein DT. 2012. Evolving communicative complexity: Insights from rodents and  
 485 beyond. *Philos Trans R Soc Lond B Biol Sci* 367(1597):1869-78.
- 486 Quick NJ and Janik VM. 2012. Bottlenose dolphins exchange signature whistles when meeting at sea.  
 487 *Proc R Soc Lond B Biol Sci* 279(1738):2539-45.
- 488 Radford A. 2005. Group-specific vocal signatures and neighbour-stranger discrimination in the  
 489 cooperatively breeding green woodhoopoe. *Anim Behav* 70(5):1227-34.
- 490 Root-Gutteridge H, Bencsik M, Chebli M, Gentle LK, Terrell-Nield C, Bourit A, Yarnell RW. 2014.  
 491 Identifying individual wild eastern grey wolves (*Canis lupus lycaon*) using fundamental  
 492 frequency and amplitude of howls. *Bioacoustics* 23:55-66.
- 493 Sayigh LS, Tyack PL, Wells RS, Solow AR, Scott MD, Irvine AB. 1999. Individual recognition in  
 494 wild bottlenose dolphins: A field test using playback experiments. *Anim Behav* 57(1):41-50.
- 495 Schlenker P, Chemla E, Cäsar C, Ryder R, Zuberbühler K. in press. Titi semantics: Context and  
 496 meaning in titi monkey call sequences. *Natural Language and Linguistic Theory* 35(1):271-298.
- 497 Searcy WA. 1992. Song repertoire and mate choice in birds. *Am Zool* 32(1):71-80.
- 498 Seyfarth RM and Cheney DL. 2014. The evolution of language from social cognition. *Curr Opin*  
 499 *Neurobiol* 28:5-9.

- 500 Stoica P and Moses RL. 2005. Spectral analysis of signals. Upper Saddle River, NJ: Pearson/Prentice  
501 Hall.
- 502 Suzuki R, Buck JR, Tyack PL. 2005. The use of Zipf's law in animal communication analysis. Anim  
503 Behav 69(1):9-17.
- 504 Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP. 2000. A procedure for an automated  
505 measurement of song similarity. Anim Behav 59(6):1167-76.
- 506 Tooze Z, Harrington F, Fentress J. 1990. Individually distinct vocalizations in timber wolves, *Canis*  
507 *lupus*. Anim Behav 40(4):723-30.
- 508 Wells RS. 2003. Dolphin social complexity: Lessons from long-term study and life history. In:  
509 Animal social complexity: Intelligence, culture, and individualized societies. de Waal FBM and  
510 Tyack PL, editors. Harvard University Press.
- 511 Zaccaroni M, Passilongo D, Buccianti A, Dessi-Fulgheri F, Facchini C, Gazzola A, Maggini I,  
512 Apollonio M. 2012. Group specific vocal signature in free-ranging wolf packs. Ethol Ecol Evol  
513 24(4):322-31.
- 514 Zuberbühler K. 2002. A syntactic rule in forest monkey communication. Anim Behav 63(2):293-9.
- 515
- 516

517 FIGURES

518



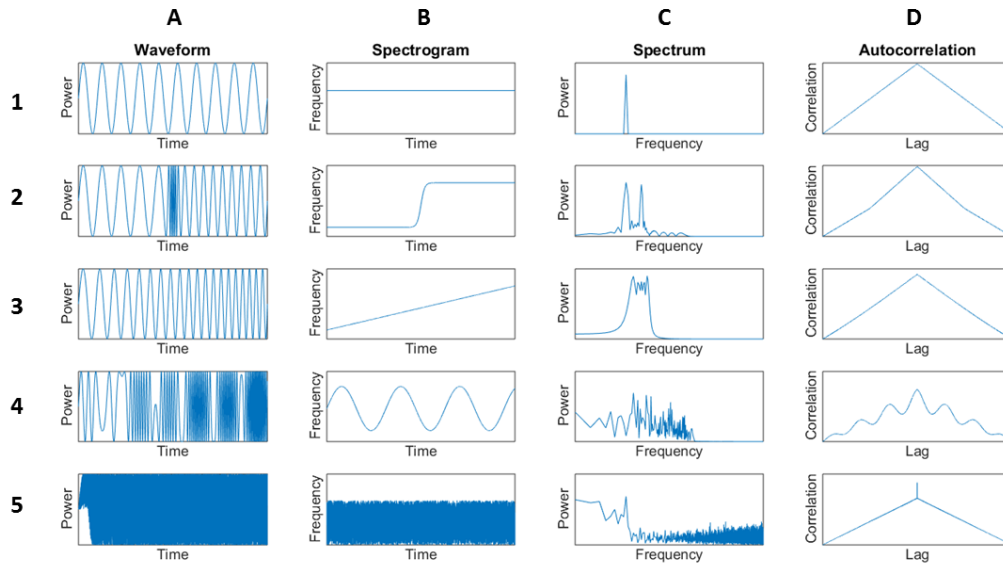
519

520 Figure 1. Six examples of wolf howls, with time on the x-axis and frequency on the y-axis. All howls

521 show some pattern of rising and falling frequencies, but the determination of which howl is most

522 complex appears wholly subjective.

523

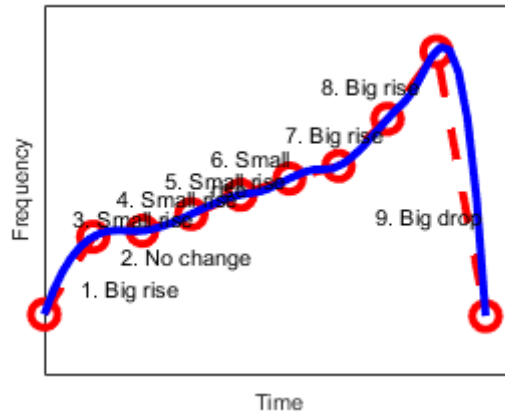


524

525 Figure 2. Simulated sounds demonstrating the relationship between waveform, spectrogram,  
 526 spectrogram, and autocorrelation. The first column (A) shows five different waveforms: (1) constant  
 527 frequency sine wave; (2) rapid doubling in frequency; (3) constantly increasing in frequency; (4)  
 528 oscillating frequency; and (5) random waveform. The second column (B) shows the frequency of the  
 529 waveform with time. Column (C) shows overall spectra for these waveforms: a single peak where one  
 530 frequency is present (1), two peaks where two frequencies exist (2), a range of frequencies in (3) and  
 531 (4), and all frequencies present in the case of white noise (5). Column (D) shows the autocorrelation  
 532 of the spectrogram.

533

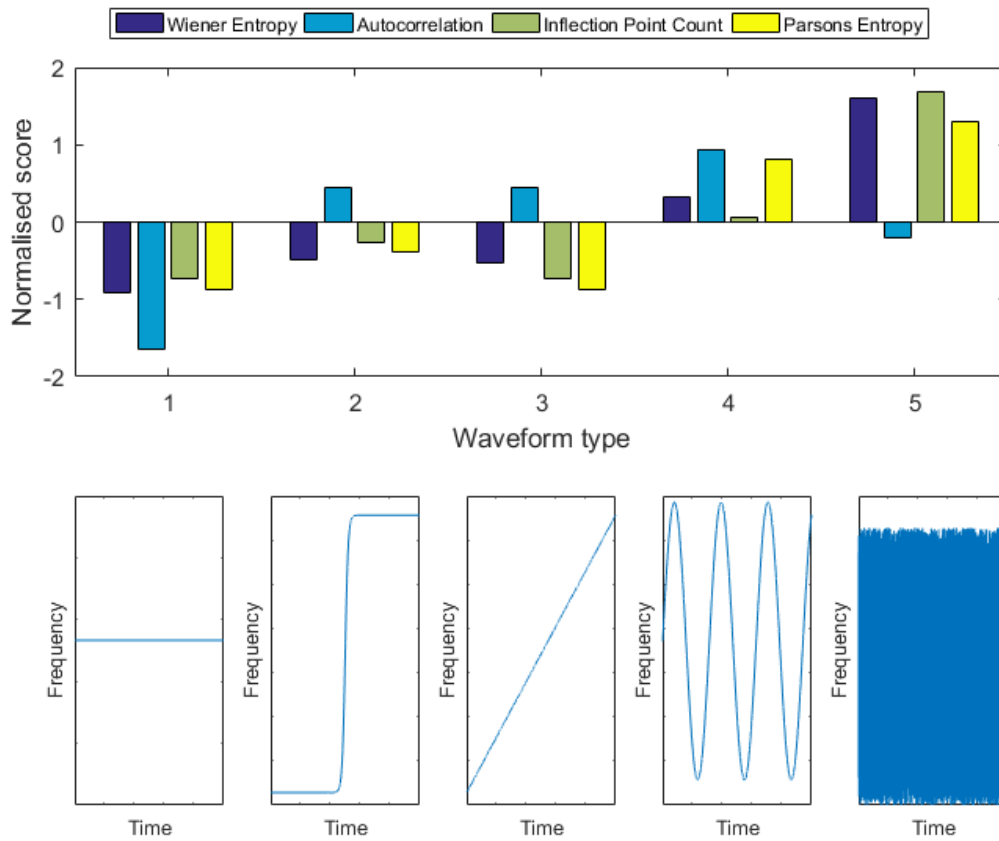
534



535

536 Figure 3: Example of a Parsons code representation of a vocal signal (blue). The time-course is  
 537 divided into ten equal sections (red), and the frequency change for each section is recorded only as  
 538 "big rise", "small rise", "no change", "small drop", or "big drop".

539

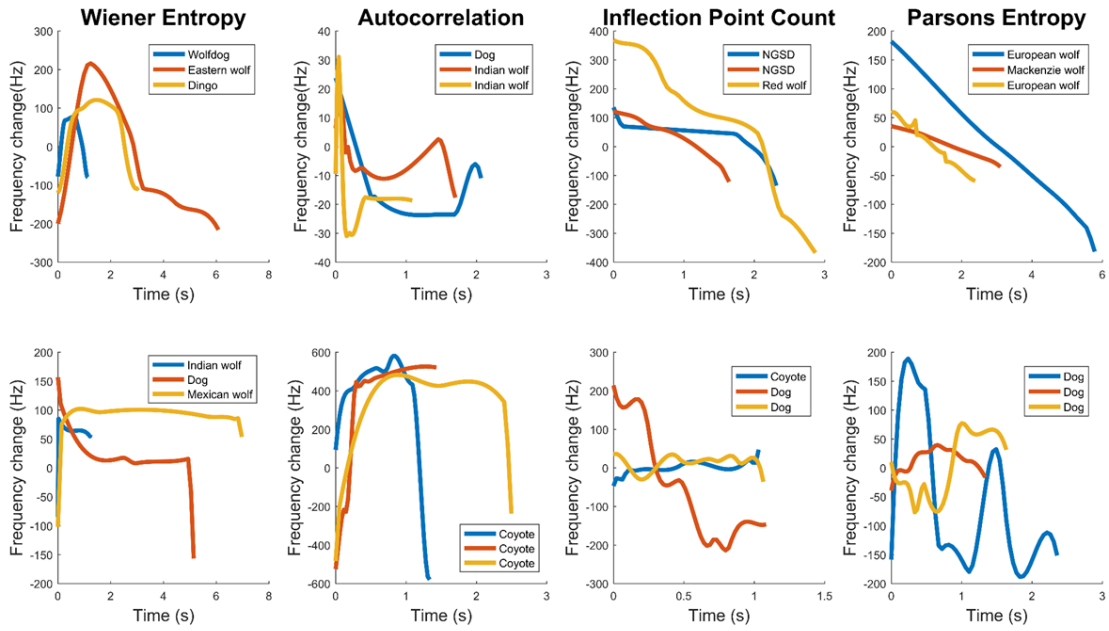


540

541 Figure 4. Complexity scores for the four metrics, for each of the five waveforms shown in Figure 2.

542 The bar charts are normalised for comparison by subtracting the mean and dividing by the standard  
 543 deviation.

544

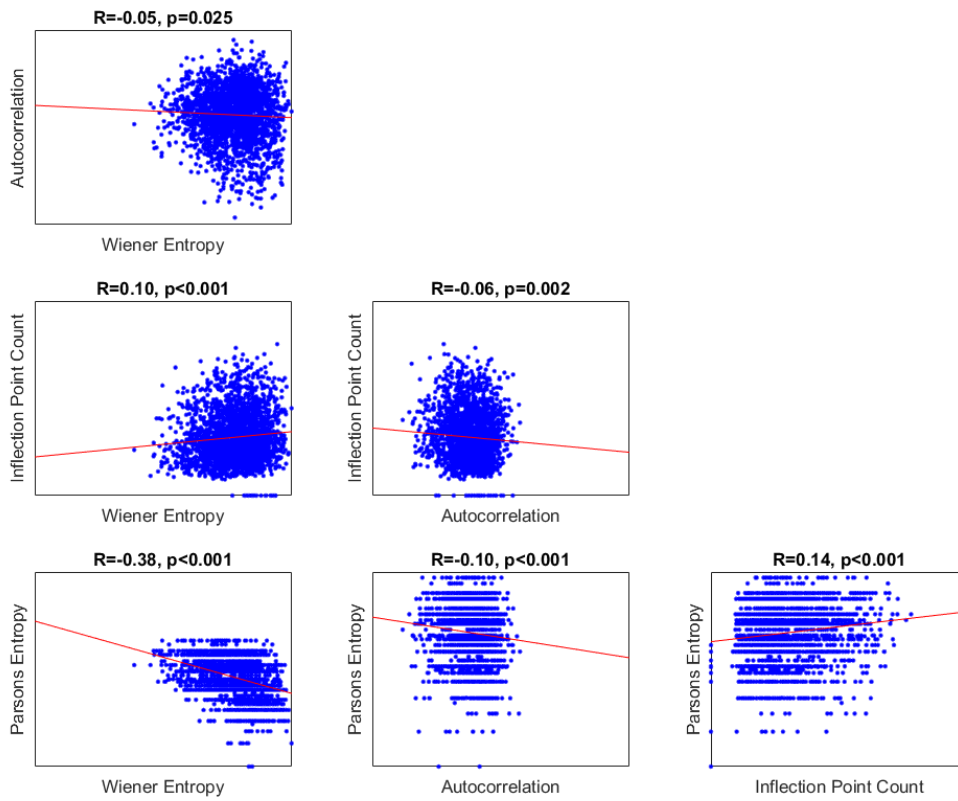


545

546 Figure 5. Examples of howls with low metric values (top row), and high metric values (bottom row),  
 547 for each of the four metric types. Each plot shows those three sample howls with the highest or lowest  
 548 values for each particular metric (not necessarily the same sample howls for each metric). The y-axis  
 549 indicates frequency deviation from the median value of each howl, to allow a clear comparison  
 550 between howls of differing frequencies.

551





552

553 Figure 6. Pairwise correlation between each of the metrics. The red line indicates line of best fit, and

554  $R$  and  $p$  values given in the title of each plot are for Pearson correlation.

555

556

557