



## Acoustic sequences in non-human animals: A tutorial review and prospectus

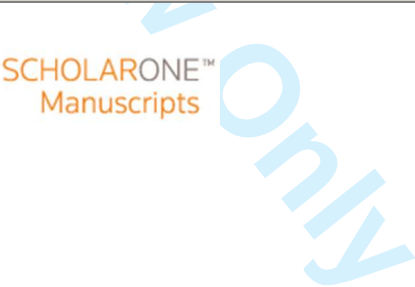
Journal:	<i>Biological Reviews</i>
Manuscript ID:	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Kershenbaum, Arik; National Institute for Mathematical and Biological Synthesis,  Blumstein, Dan; University of California, Ecol Evol Biol  Roch, Marie; San Diego State University, Department of Computer Science  Akçay, Caglar; Cornell University, Lab of Ornithology  Backus, Gregory; North Carolina State University, Department of Biomathematics  Bee, Mark; University of Minnesota, Department of Ecology, Evolution, &amp; Behavior  Bohn, Kirsten; Florida International University, Integrated Science  Cao, Yan; University of Texas at Dallas, Department of Mathematical Sciences  Carter, Gerald; University of Maryland, Department of Ecology, Evolution, &amp; Systematics  Căsar, Cristiane; University of St. Andrews, Department of Psychology &amp; Neuroscience  Coen, Michael; University of Wisconsin, Department of Biostatistics and Medical Informatics  DeRuiter, Stacy; University of St. Andrews, School of Mathematics and Statistics  Doyle, Laurance; SETI Institute,  Edelman, Shimon; Cornell University, Department of Psychology  Ferrer-i-Cancho, Ramon; Universitat Politècnica de Catalunya, Departament de Llenguatges i Sistemes Informàtics  Freeberg, Todd; University of Tennessee, Department of Psychology  Garland, Ellen; AFSC/NOAA, National Marine Mammal Laboratory  Gustison, Morgan; University of Michigan, Department of Psychology  Harley, Heidi; New College of Florida, Department of Psychology  Huetz, Chloe; Université Paris Sud, Department of Bioacoustics  Hughes, Melissa; College of Charleston, Department of Biology  Hyland Bruno, Julia; Hunter College, The City University of New York, Department of Psychology  Ilany, Amiyaal; National Institute for Mathematical and Biological Synthesis,  Jin, Dezhe; Pennsylvania State University, Department of Physics  Johnson, Michael; Marquette University, Department of Electrical and Computer Engineering  Ju, Chenghui; Queen College, The City Univ. of New York, Department of</p>

This is the peer reviewed version of the following article: Kershenbaum, A., Blumstein, D., Roch, M., Ferrer-i-Cancho, R... [et al.]. *Acoustic sequences in non-human animals: a tutorial review and prospectus*. "Biological reviews", Febrer 2016, vol. 91, núm. 1, p. 13-52, which has been published in final form at DOI 10.1111/brv.12160 .

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

	<p>Biology Karnowski, Jeremy; University of California San Diego, Department of Cognitive Science Lohr, Bernard; University of Maryland Baltimore County, Department of Biological Sciences Manser, Marta; University of Zurich, Department of Evolutionary Biology and Environmental Studies McCowan, Brenda; University of California Davis, Department of Veterinary Medicine Mercado, Eduardo; State University of New York, Department of Psychology Narins, Peter; University of California Los Angeles, Department of Integrative Biology &amp; Physiology Piel, Alex; University of Cambridge, Department of Biological Anthropology Rice, Megan; California State University San Marcos, Department of Psychology Salmi, Roberta; University of Georgia at Athens, Department of Anthropology Sasahara, Kazutoshi; Nagoya University, Graduate School of Information Science Sayigh, Laela; Woods Hole Oceanographic Institute, Department of Biology Shiu, Yu; Cornell University, Lab of Ornithology Taylor, Charles; University of California, Ecol Evol Biol Vallejo, Edgar; Monterrey Technical University, Department of Computer Science Waller, Sara; Montana State University, Department of Philosophy Zamora-Gutierrez, Veronica; University of Cambridge, Department of Zoology</p>
Keywords:	<p>acoustic communication, information, information theory, machine learning, Markov model, meaning, network analysis, sequence analysis, vocalisation</p>



# Acoustic sequences in non-human animals: A tutorial review and prospectus

Arik Kershenbaum<sup>1</sup>, Daniel T. Blumstein<sup>2</sup>, Marie A. Roch<sup>3</sup>, Çağlar Akçay<sup>4</sup>, Gregory Backus<sup>5</sup>, Mark Bee<sup>6</sup>, Kirsten Bohn<sup>7</sup>, Yan Cao<sup>8</sup>, Gerald Carter<sup>9</sup>, Cristiane Căsar<sup>10</sup>, Michael Coen<sup>11</sup>, Stacy L. DeRuiter<sup>12</sup>, Laurance Doyle<sup>13</sup>, Shimon Edelman<sup>14</sup>, Ramon Ferrer-i-Cancho<sup>15</sup>, Todd M. Freeberg<sup>16</sup>, Ellen C. Garland<sup>17</sup>, Morgan Gustison<sup>18</sup>, Heidi E. Harley<sup>19</sup>, Chloé Huetz<sup>20</sup>, Melissa Hughes<sup>21</sup>, Julia Hyland Bruno<sup>22</sup>, Amiyaal Ilany<sup>23</sup>, Dezhe Z. Jin<sup>24</sup>, Michael Johnson<sup>25</sup>, Chenghui Ju<sup>26</sup>, Jeremy Karnowski<sup>27</sup>, Bernard Lohr<sup>28</sup>, Marta B. Manser<sup>29</sup>, Brenda McCowan<sup>30</sup>, Eduardo Mercado III<sup>31</sup>, Peter M. Narins<sup>32</sup>, Alex Piel<sup>33</sup>, Megan Rice<sup>34</sup>, Roberta Salmi<sup>35</sup>, Kazutoshi Sasahara<sup>36</sup>, Laela Sayigh<sup>37</sup>, Yu Shiu<sup>38</sup>, Charles Taylor<sup>39</sup>, Edgar E. Vallejo<sup>40</sup>, Sara Waller<sup>41</sup>, Veronica Zamora-Gutierrez<sup>42</sup>

<sup>1</sup> National Institute for Mathematical and Biological Synthesis, Knoxville, USA

<sup>2</sup> Department of Ecology and Evolutionary Biology, University of California Los Angeles, USA

<sup>3</sup> Department of Computer Science, San Diego State University, USA

<sup>4</sup> Lab of Ornithology, Cornell University, USA

<sup>5</sup> Department of Biomathematics, North Carolina State University, USA

<sup>6</sup> Department of Ecology, Evolution and Behavior, University of Minnesota, USA

<sup>7</sup> Integrated Science, Florida International University, USA

<sup>8</sup> Department of Mathematical Sciences, University of Texas at Dallas, USA

<sup>9</sup> Biological Sciences Graduate Program, University of Maryland, USA

<sup>10</sup> Department of Psychology & Neuroscience, University of St. Andrews, UK

<sup>11</sup> Department of Biostatistics and Medical Informatics, University of Wisconsin, USA

<sup>12</sup> School of Mathematics and Statistics, University of St. Andrews, UK

<sup>13</sup> Carl Sagan Center for the Study of Life in the Universe, SETI Institute, USA

<sup>14</sup> Department of Psychology, Cornell University, USA

<sup>15</sup> Department of Computer Science, Universitat Politècnica de Catalunya, (Catalonia) Spain

<sup>16</sup> Department of Psychology, University of Tennessee, USA

<sup>17</sup> National Marine Mammal Laboratory, AFSC/NOAA, USA

<sup>18</sup> Department of Psychology, University of Michigan, USA

<sup>19</sup> Division of Social Sciences, New College of Florida, USA

<sup>20</sup> CNRS, CNRS UMR 8195, Université Paris-Sud, Orsay, France

<sup>21</sup> Department of Biology, College of Charleston, USA

<sup>22</sup> Department of Psychology, Hunter College and the Graduate Center, The City University of New York, USA

<sup>23</sup> National Institute for Mathematical and Biological Synthesis, Knoxville, USA

<sup>24</sup> Department of Physics, Pennsylvania State University, USA

<sup>25</sup> Department of Electrical and Computer Engineering, Marquette University, USA

<sup>26</sup> Department of Biology, Queen College, The City Univ. of New York, USA

<sup>27</sup> Department of Cognitive Science, University of California San Diego, USA

<sup>28</sup> Department of Biological Sciences, University of Maryland Baltimore County, USA

<sup>29</sup> Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland

<sup>30</sup> Department of Veterinary Medicine, University of California Davis, USA

<sup>31</sup> Department of Psychology; Evolution, Ecology, & Behavior, University at Buffalo, The State University of New York, USA

<sup>32</sup> Department of Integrative Biology & Physiology, University of California Los Angeles, USA

<sup>33</sup> Division of Biological Anthropology, University of Cambridge, UK

<sup>34</sup> Department of Psychology, California State University San Marcos, USA

<sup>35</sup> Department of Anthropology, University of Georgia at Athens, USA

<sup>36</sup> Graduate School of Information Science, Nagoya University, Japan

<sup>37</sup> Biology Department, Woods Hole Oceanographic Institution, USA

<sup>38</sup> Lab of Ornithology, Cornell University, USA

<sup>39</sup> Department of Ecology and Evolutionary Biology, University of California Los Angeles, USA

<sup>40</sup> Department of Computer Science, Monterrey Institute of Technology, Mexico

<sup>41</sup> Department of Philosophy, Montana State University, USA

<sup>42</sup> Department of Zoology, University of Cambridge, UK, Centre for Biodiversity and Environmental Research, University College London, UK

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

ABSTRACT

Animal acoustic communication often takes the form of complex sequences, made up of multiple distinct acoustic units. Apart from the well-known example of birdsong, other animals such as insects, amphibians, and mammals (including bats, rodents, primates, and cetaceans) also generate complex acoustic sequences. Occasionally, such as with birdsong, the adaptive role of these sequences seems clear (e.g., mate attraction and territorial defence). More often however, researchers have only begun to characterise – let alone understand – the significance and meaning of acoustic sequences. Hypotheses abound, but there is little agreement as to how sequences should be defined and analysed. Our work here aims to forge such an agreement on key hypotheses, to outline suitable methods for testing these hypotheses, and to describe the major limitations to our current and near-future knowledge on questions of acoustic sequences.

This review and prospectus is the result of a collaborative effort between 43 scientists from the fields of animal behaviour, ecology and evolution, signal processing, machine learning, quantitative linguistics, and information theory, who gathered for a 2013 workshop entitled, “Analysing vocal sequences in animals”. Our goal is to present not just a review of the state of the art, but to propose a methodological framework that summarises what we suggest are the best practices for research in this field, across taxa and across disciplines. We also provide a tutorial-style introduction to some of the most promising algorithmic approaches for analysing sequences.

We divide our review into three sections: describing the different ways that information can be contained within a sequence, identifying the distinct units of an acoustic sequence, and analysing the structure of that sequence. Each of these sections is further subdivided to address the key questions and approaches in that area.

We propose a uniform, systematic, and comprehensive approach to studying sequences, with the goal of clarifying research terms used in different fields, and facilitating collaboration and comparative

37 studies. Allowing greater interdisciplinary collaboration will facilitate the investigation of many important  
38 questions in the evolution of communication and sociality.

39  
40 Keywords: acoustic communication, information, information theory, machine learning, Markov model,  
41 meaning, network analysis, sequence analysis, vocalisation

42

43

For Review Only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

44	CONTENTS
45	I. INTRODUCTION
46	II. THE CONCEPTS OF INFORMATION AND MEANING
47	(1) <i>Information embedding paradigms</i>
48	(2) <i>Six information embedding paradigms</i>
49	III. ACOUSTIC UNITS
50	(1) <i>Identifying potential units</i>
51	(2) <i>Identifying production units</i>
52	(3) <i>Identifying perceptual units</i>
53	(4) <i>Identifying analytical units</i>
54	IV. THE STRUCTURE OF SEQUENCES
55	(1) <i>Markov chains</i>
56	(2) <i>Hidden Markov models</i>
57	(3) <i>Network models</i>
58	(4) <i>Formal grammars</i>
59	(5) <i>Temporal structure</i>
60	V. FUTURE DIRECTIONS
61	(1) <i>Two case studies</i>
62	(2) <i>Outstanding questions in animal vocal sequences</i>
63	VI. CONCLUSIONS
64	VII. ACKNOWLEDGEMENTS
65	VIII. REFERENCES

## 69 I. INTRODUCTION

70

71 Sequences are everywhere, from the genetic code, to behavioural patterns such as foraging, as well as the  
72 sequences that comprise music and language. Often, but not always, sequences convey meaning (and can  
73 do so more effectively than other types of signals; Shannon et al., 1949), and individuals can take  
74 advantage of the information contained in a sequence to increase their own fitness (Bradbury &  
75 Vehrencamp, 2011). Acoustic communication is widespread in the animal world, and very often  
76 individuals communicate using a sequence of distinct acoustic elements, the order of which may contain  
77 information of potential benefit to the receiver. In some cases, acoustic sequences appear to be ritualised  
78 signals where the signaller benefits if the signal is detected and acted upon by a receiver. The most  
79 studied examples include birdsong, where males may use sequences to advertise their potential quality to  
80 rival males and to receptive females (Catchpole & Slater, 2003). Acoustic sequences can contain  
81 information on species identity (e.g., in many frogs and insects; Gerhardt & Huber, 2002), on individual  
82 identity and traits (e.g., in starlings *Sturnus vulgaris*, Gentner & Hulse, 1998; wolves *Canis lupus*, Root-  
83 Gutteridge et al., 2014; dolphins *Tursiops truncatus*, Sayigh et al., 2007; and hyraxes *Procavia capensis*,  
84 Koren & Geffen, 2011), and in some cases, on contextual information such as resource availability (e.g.,  
85 food calls in chimpanzees *Pan troglodytes*, Slocombe & Zuberbühler, 2006), or predator threats (e.g., in  
86 marmots *Marmota spp.*, Blumstein, 2007; primates, Schel, Tranquilli & Zuberbühler, 2009; Cäsar et al.,  
87 2012b; and parids, Baker & Becker, 2002). In many cases, however, the ultimate function of  
88 communicating in sequences is unclear. Understanding the proximate and ultimate forces driving and  
89 constraining the evolution of acoustic sequences, as well as decoding the information contained within  
90 them, is a growing field in animal behaviour (Freeberg et al., 2012). New analytical techniques are  
91 uncovering characteristics shared between diverse taxa, and offer the potential of describing and  
92 interpreting the information within animal communication signals. The field is ripe for a review and a  
93 prospectus to guide future empirical research.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

94 Progress in this field has been somewhat hampered in the past partly by inconsistent terminology,  
95 conflicting assumptions, and different research goals, both between disciplines (e.g., between biologists  
96 and mathematicians), and also between researchers concentrating on different taxa (e.g., ornithologists  
97 and primatologists). Therefore, we aim to do more than provide a glossary of terms. Rather, we build a  
98 framework that identifies the key conceptual issues common to the study of acoustic sequences of all  
99 types, while providing specific definitions useful for clarifying questions and approaches in more narrow  
100 fields. Our approach identifies three central questions: What are the units of which the sequence is  
101 composed? How do we assess the structure with which these units are combined? How is information  
102 contained within the sequence? Figure 1 illustrates a conceptual flow diagram linking these questions, and  
103 their sub-components, and should be broadly applicable to any study involving animal acoustic  
104 sequences.

105 Our aims in this review are as follows: (1) to identify the key issues and concepts necessary for  
106 the successful analysis of animal acoustic sequences; (2) to describe the commonly used analytical  
107 techniques, and importantly, also those underused methods deserving of more attention; (3) to encourage  
108 a cross-disciplinary approach to the study of animal acoustic sequences that takes advantage of tools and  
109 examples from other fields to create a broader synthesis; and (4) to facilitate the investigation of new  
110 questions through the articulation of a solid conceptual framework.

111 In Section II we ask why sequences are important, and how information may be embedded within  
112 them. We present this section first, rather than in the order shown in Figure 1, because it is necessary  
113 early on to define and establish the significance of the terminology that will follow in the review. In  
114 Section III, we examine the questions of what units make up a sequence and how to identify them. In  
115 some applications the choice seems trivial, however in many study species, sequences can be represented  
116 at different hierarchical levels of abstraction, and the choice of sequence “unit” may depend on the  
117 hypotheses being tested. In Section IV, we examine the structure of the sequence, the mathematical and  
118 statistical models that quantify how units are combined, and how these models can be analysed,  
119 compared, and assessed. In Section V, we provide some case studies that illustrate our approach, describe



1  
2  
3 120 some of the evolutionary and ecological questions that can be addressed by analysing animal acoustic  
4  
5 121 sequences, and look at some promising future directions and new approaches.  
6  
7  
8 122  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Review Only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

123

124 II. THE CONCEPTS OF INFORMATION AND MEANING

125

126 The complementary terms, “meaning” and “information” in communication, have been variously defined,  
127 and have long been the subject of some controversy (Dawkins & Krebs, 1978; Stegmann, 2013). In this  
128 section we explore some of the different definitions from different fields, and their significance for  
129 research on animal behaviour. The distinction between meaning and information is sometimes portrayed  
130 with meaning, on the one hand, as activity, and information, on the other hand, as form, or structure  
131 (Bohm, 1989).

132 Philosophical understanding of meaning is rooted in studies of human language and has a variety  
133 of schools of thought. Philosophers consider intension (a meaning or sense, e.g., a chair is something that  
134 one sits on) and extension (objects that are instances of an intension, e.g., a particular lounge chair),  
135 prototype theory, whether meaning is innate or learned, mental representations, and cognitive content.  
136 Philosophers also view meaning as computational/functional, as atomic or holistic, as bound to both  
137 speaker and audience, as speech act and performance, as rule bound or as referentially based, as  
138 description, as conventional, and as a game dependent on a form of life, among other examples  
139 (Christiansen & Chater, 2001; Martinich & Sosa, 2013).

140 Biologists (particularly behavioural ecologists), and cognitive neuroscientists have different  
141 understandings of meaning. For most biologists, meaning relates to the function of signalling. The  
142 function of signals is examined in agonistic and affiliative interactions, in courtship and mating decisions,  
143 and in communicating about environmental stimuli, such as the detection of predators (Bradbury &  
144 Vehrencamp, 2011). Behavioural ecologists study meaning by determining the degree of production  
145 specificity, the degree of response specificity, and contextual independence (e.g., Evans, 1997). Cognitive  
146 neuroscientists generally understand meaning through mapping behaviour onto structure-function  
147 relationships in the brain (Chatterjee, 2005).

Mathematicians understand meaning by developing theories and models to interpret the observed signals. This includes defining and quantifying the variables (observable and unobservable), and the formalism for combining various variables into a coherent framework (e.g., pattern theory; Mumford & Desolneux, 2010). One approach to examining a signal mathematically is to determine the entropy, or amount of structure (or lack thereof) present in a sequence.

The amount of structure measured does not indicate meaning, nor does measuring the amount of structure quantify the complexity of the meaning, if it is present. As an example, the structure may be indicative of meaning, but it could also be related to a strategy to ensure acoustic propagation in an adverse environment. A distinction is often made between a signal's "content", or *broadcast information*, and its "efficacy", or *transmitted information* – the characteristics or features of signals that actually reach receivers (Wiley, 1983; Hebets & Papaj, 2005). This is basically the distinction between *bearing functional information* and *getting that information across* to receivers in conditions that can be adverse to clear signal propagation. A sequence may also contain elements that do not in themselves contain meaning, but are intended to get the listeners' attention, in anticipation of future meaningful elements (e.g., Richards, 1981; Call & Tomasello, 2007; Arnold & Zuberbühler, 2013).

Context has a profound influence on signal meaning, and this should apply to the meaning of sequences as well. Context includes internal and external factors that may influence both the production and perception of acoustic sequences; the effects of context can partially be understood by considering how it specifically influences the costs and benefits of producing a particular signal or responding to it. For instance, an individual's motivational, behavioural, or physiological state may influence response (Lynch et al., 2005; Goldbogen et al., 2013); hungry animals respond differently to signals than satiated ones, and an individual in oestrus or musth may respond differently than ones not in those altered physiological states (Poole, 1999). Gender may influence response as well (Tyack, 1983; Darling, Jones & Nicklin, 2006; Smith et al., 2008; van Schaik, Damerius & Isler, 2013). The social environment may influence the costs and benefits of responding to a particular signal (Bergman et al., 2003; Wheeler, 2010a; Ilany et al., 2011; Wheeler & Hammerschmidt, 2012) as might environmental attributes, such as

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

174 temperature or precipitation. Knowledge from other social interactions or environmental experiences can  
175 also play a role in context, e.g., habituation (Krebs, 1976). Context can also alter a behavioural response  
176 when hearing the same signal originate from different spatial locations. For instance in neighbour-stranger  
177 discrimination in songbirds, territorial males typically respond less aggressively toward neighbours  
178 compared with strangers, so long as the two signals are heard coming from the direction of the  
179 neighbour's territory. If both signals are played back from the centre of the subject's territory, or from a  
180 neutral location, subjects typically respond equally aggressively to both neighbours and strangers (Falls,  
181 1982; Stoddard, 1996). Identifying and testing for important contextual factors appears to be an essential  
182 step in decoding the meaning of sequences.

183         Qualitatively, we infer meaning in a sequence if it modifies the receiver's response in some  
184 predictable way. Quantitatively, information theory measures the amount of information (usually in units  
185 of bits) transmitted and received within a communication system (Shannon et al., 1949). Therefore,  
186 information theory approaches can describe the complexity of the communication system. Information  
187 theory additionally can characterise transmission errors and reception errors, and has been  
188 comprehensively reviewed in the context of animal communication in (Bradbury & Vehrencamp, 2011).

189         Considerable debate exists over the nature of animal communication and the terminology used in  
190 animal communication research (Owren, Rendall & Ryan, 2010; Seyfarth et al., 2010; Ruxton &  
191 Schaefer, 2011; Stegmann, 2013), and in particular the origin of and relationship between meaning and  
192 information, and their evolutionary significance. For our purposes, we will use the term "meaning" when  
193 discussing behavioural and evolutionary processes, and the term "information" when discussing the  
194 mathematical and statistical properties of sequences. This parallels (but is distinct from) the definitions  
195 given by Ruxton & Schaefer (2011), in particular because we wish to have a single term ("information")  
196 that describes inherent properties of sequences, without reference to the putative behavioural effects on  
197 receivers, or the ultimate evolutionary processes that caused the sequence to take the form that it does. In  
198 the remainder of this section on information and meaning, we address the question of how information  
199 can be embedded into signal sequences.

200

201 (1) *Information embedding paradigms*

202 A “sequence” can be defined as an ordered list of units. Animals produce sequences of sounds through a  
203 wide range of mechanisms (e.g., vocalisation, stridulation, percussion), and different uses of the sound-  
204 producing apparatus can produce different sound “units” with distinct and distinguishable properties. The  
205 resulting order of these varied sound units may or may not contain information that can be interpreted by  
206 a receiver, irrespective of whether or not the signaller intended to convey meaning. Given that a sequence  
207 must consist of more than one “unit” of one or more different types, the delineation and definition of the  
208 unit types is clearly of vital importance. We discuss this question at length in Section III. However,  
209 assuming that units have been successfully assigned short-hand labels (e.g., *A*, *B*, *C*, etc.), what different  
210 methods can be used to arrange these units in a sequence, in such a way that the sequence can contain  
211 information?

212 Although it seems intuitively obvious that a sequence of such labels may contain information, this  
213 intuition arises from our own natural human dispensation to language and writing, and may not be  
214 particularly useful in identifying information in animal sequences. We appreciate birdsong, for instance,  
215 as a complex combination of notes, and may be tempted to compare this animal vocalisation to human  
216 music (Baptista & Keister, 2005; Araya-Salas, 2012; Rothenberg et al., 2013). An anthropocentric  
217 approach, however, is not likely in all cases to identify structure relevant to animal communication.  
218 Furthermore, wide variation can be expected between the structure of sequences generated by different  
219 taxa, from the pulse-based stridulation of insects (Gerhardt & Huber, 2002) to song in whales (reviewed  
220 in Cholewiak, Sousa-Lima & Cerchio, 2012), and a single analytical paradigm derived from a narrow  
221 taxonomic view is also likely to be inadequate. A more rigorous analysis is needed, one that indicates the  
222 fundamental structural properties of acoustic sequences, in all their diversity. Looking for information  
223 only, say, in the order of units can lead researchers to miss information encoded in unit timing, or pulse  
224 rate.

We suggest a classification scheme based on six distinct paradigms for encoding information in sequences (Figure 2): (a) *Repetition*, where a single unit is repeated more than once; (b) *Diversity*, where information is represented by the number of distinct units present; (c) *Combination*, where sets of units have different information from each unit individually; (d) *Ordering*, where the relative position of units to each other is important; (e) *Overlapping*, where information is conveyed in the relationship between sequences of two or more individuals; and (f) *Timing*, where the time gap between units conveys information. This framework can form the basis of much research into sequences, and provides a useful and comprehensive approach for classifying information-bearing sequences. We recommend that in any research into animal acoustic communication with a sequential component, researchers first identify the place(s) of their focal system in this framework, and use this structure to guide the formulation of useful, testable hypotheses. Considering the formal structures of possible information embedding systems may provide supportive insights into the cognitive and evolutionary processes taking place (Chatterjee, 2005; Seyfarth, Cheney & Bergman, 2005). Of course, any particular system might have properties of more than one of the six paradigms in this framework, and the boundaries between them may not always be clearly distinguished. Sperm whale *Physeter macrocephalus* coda exchanges (Watkins & Schevill, 1977) provide an example of this. A coda is a sequence of clicks (*Repetition* of the acoustic unit) where the *Timing* between echolocation clicks moderates response. In duet behaviour, *Overlap* also exists, with one animal producing and another responding with another coda (Schulz et al., 2008). Each of these paradigms is now described in more detail below.

(2) *Six information embedding paradigms*

1. Repetition: Sequences are made of repetitions of discrete units, and repetitions of the same unit affect receiver responses. For instance, the information contained in a unit *A* given in isolation may convey a different meaning to a receiver than an iterated sequence of unit *A* (e.g., *AAAA*, etc.). For example, greater numbers of D notes in the *chick-a-dee* calls of chickadee species *Poecile spp.* can be related to the immediacy of threat posed by a detected predator (Krams et al., 2012). Repetition in alarm

calls are related to situation urgency (meerkats *Suricata suricatta*: Manser, 2001, marmots *Marmota spp.*: Blumstein, 2007, colobus monkeys *Colobus spp.*: Schel, Candiotti & Zuberbühler, 2010, Campbell's monkeys *Cercopithecus campbelli*: Lemasson et al., 2010, lemurs *Lemur catta* and *Varecia variegata*: Macedonia, 1990). As an additional example, many frog species produce pulsatile advertisement calls consisting of the same repeated element. If it is the case that both number of pulses and pulse rate affect receiver responses, as shown in some hyliid treefrogs (Gerhardt, 2001), then information is being embedded using both the *Repetition* (1) and the *Timing* (6) paradigms simultaneously. Such use of multiple embedding techniques may be quite common, for instance in intrasexual competitive and intersexual reproductive contexts (Gerhardt & Huber, 2002).

2. Diversity: Sequences of different units (e.g., A, B, C) are produced, but those units are functionally interchangeable, and therefore ordering is unimportant. For instance, many songbirds produce songs with multiple different syllables. In many species, however, the particular syllables are substitutable (e.g., Eens, Pinxten & Verheyen, 1991; Farabaugh & Dooling, 1996, but see Lipkind et al., 2013) and receivers attend to the overall diversity of sounds in the songs or repertoires of signallers (Catchpole & Slater, 2003). Large acoustic repertoires have been proposed to be sexually selected in species such as great reed warblers *Acrocephalus arundinaceus* and common starlings *Sturnus vulgaris* (Eens, Pinxten & Verheyen, 1993; Hasselquist, Bensch & von Schantz, 1996; Eens, 1997), in which case diversity embeds information (that carries meaning) on signaller quality (e.g., Kipper et al., 2006). Acoustic "diversity" has additionally been proposed as a means of preventing habituation on the part of the receiver (Hartshorne, 1956; Hartshorne, 1973; Kroodsma, 1990) as well as a means of avoiding (neuromuscular) "exhaustion" on the part of the sender (Lambrechts & Dhondt, 1987; Lambrechts & Dhondt, 1988). We do note that these explanations remain somewhat controversial, especially if the transitions between acoustic units are, indeed, biologically constrained (Weary & Lemon, 1988; Weary et al., 1988; Weary & Lemon, 1990; Weary, Lambrechts & Krebs, 1991; Riebel & Slater, 2003; Brumm & Slater, 2006).

1  
2  
3 277  
4  
5 278 3. Combination: Sequences may consist of different discrete acoustic units (e.g., *A*, *B*, *C*) each of  
6  
7  
8 279 which is itself meaningful, and the combining of the different units conveys distinct information. Here,  
9  
10 280 order does not matter (in contrast to the *Ordering* paradigm below) – the sequence of unit *A* followed by  
11  
12 281 unit *B* has the same information as the sequence of unit *B* followed by unit *A*. For example, titi monkeys  
13  
14 282 *Callicebus nigrifrons* (Cäsar et al., 2013) use semantic alarm combinations, in which interspersing avian  
15  
16 283 predator alarms calls (A-type) with terrestrial predator alarm calls (B-type) indicates the presence of a  
17  
18 284 raptor on the ground. In this case, the number of calls (i.e. *Repetition*) also appears to influence the  
19  
20 285 information present in each call sequence (Cäsar et al., 2013).  
21  
22 286  
23  
24 287 4. Ordering: Sequences of different discrete acoustic units (e.g., *A*, *B*, *C*) each of which is itself  
25  
26 288 meaningful and the specific order of which is meaningful. Here, order matters – and the ordered  
27  
28 289 combination of discrete units may result in emergent responses. For instance, *A* followed by *B* may elicit  
29  
30 290 a different response than either *A* or *B* alone, or *B* followed by *A*. Examples include primate alarm calls  
31  
32 291 which, when combined, elicit different responses related to the context of the predatory threat (Arnold &  
33  
34 292 Zuberbühler, 2006a; Arnold & Zuberbühler, 2008). Human languages are a sophisticated example of  
35  
36 293 ordered information encoding (Hauser, Chomsky & Fitch, 2002).  
37  
38 294  
39  
40 295 5. Overlapping: Sequences are combined from two or more individuals into exchanges for which the  
41  
42 296 order of these overlapping sequences has information distinct from each signaller's signals in isolation.  
43  
44 297 Overlapping can be in the time dimension (i.e., two signals emitted at the same time) or in acoustic space  
45  
46 298 (e.g., song type matching, Krebs, Ashcroft & Orsdol, 1981, and frequency matching, Mennill & Ratcliffe,  
47  
48 299 2004) For example, in different parid species (Paridae: chickadees, tits, and titmice), females seem to  
49  
50 300 attend to the degree to which their males' songs are overlapped (in time) by neighbouring males' songs,  
51  
52 301 and seek extra-pair copulations when their mate is overlapped (Otter et al., 1999; Mennill, Ratcliffe &  
53  
54 302 Boag, 2002). Overlapping is also used for social bonding, spatial perception, and reunion, such as chorus  
55  
56  
57  
58  
59  
60



1  
2  
3 303 howls in wolves (Harrington et al., 2003) and sperm whale codas (Schulz et al., 2008). Overlapping as  
4  
5 304 song type matching (overlapping in acoustic space) is also an aggressive signal in some songbirds (Akçay  
6  
7 305 et al., 2013), though this may depend on whether it is the sequence or the individual unit that is  
8  
9  
10 306 overlapped (Searcy & Beecher, 2011). Coordination between the calling of individuals can also give  
11  
12 307 identity cues (Carter et al., 2008).  
13  
14 308

15  
16 309 6. Timing: The temporal spacing between units in a sequence can contain information. In the  
17  
18 310 simplest case, pulse rate and interpulse interval can distinguish between different species, for example in  
19  
20 311 insects and anurans (Gerhardt & Huber, 2002; Nityananda & Bee, 2011), rodents (Randall, 1997), and  
21  
22 312 primates (Hauser, Agnetta & Perez, 1998). Call timing can indicate fitness and aggressive intent, e.g.,  
23  
24 313 male howler monkeys *Alouatta pigra* attend to howling delay as an indicator of aggressive escalation  
25  
26 314 (Kitchen, 2004). Additionally, when sequences are produced by different individuals, a receiver may  
27  
28 315 interpret the timing differences between the producing individuals to obtain contextual information. For  
29  
30 316 instance, ground squirrels *Spermophilus richarsonii* use the spatial pattern and temporal sequence of  
31  
32 317 conspecific alarm calls to provide information on a predator's movement trajectory (Thompson & Hare,  
33  
34 318 2010). This information only emerges from the sequence of different callers initiating calls (Blumstein,  
35  
36 319 Verneyre & Daniel, 2004). Such risk tracking could also emerge from animals responding to sequences of  
37  
38 320 heterospecific alarm signals produced over time.  
39  
40  
41  
42 321

43  
44 322 We conclude this section with a discussion of two examples of how sequences of acoustic signals  
45  
46 323 produced by signallers can influence meaning to receivers. These two examples come from primates and  
47  
48 324 exemplify the *Diversity* and *Ordering* types of sequences illustrated in Figure 2. The example of the  
49  
50 325 *Diversity* type is the system of serial calls of titi monkeys, *Callicebus molloch*, used in a wide range of  
51  
52 326 social interactions. Here, the calls comprise several distinct units, many of which are produced in  
53  
54 327 sequences. Importantly, the units of this call system seem to have meaning primarily in the context of the  
55  
56 328 sequence – this call system therefore seems to represent the notion of phonological syntax (Marler, 1977).  
57  
58  
59  
60

1  
2  
3 329 One sequence has been tested via playback studies – the ‘honks-bellows-pumps’ sequence is used  
4  
5 330 frequently by males that are isolated from and not closely associated with females and may recruit non-  
6  
7 331 paired females (Robinson, 1979). Robinson (1979) played back typical sequences of honks-bellows-  
8  
9 332 pumps sequences and atypical sequences of honks-pumps-bellows and found little evidence that groups of  
10  
11 333 titi monkeys responded differently to the two playbacks (though they gave one call type – a ‘moan’,  
12  
13 334 produced often during disturbances caused by other conspecific or heterospecific monkey groups – more  
14  
15 335 often to the atypical sequences). Unfortunately, the playbacks were not done to groups at a distance to  
16  
17 336 determine whether sequence order mattered to the question of recruitment of females.  
18  
19

20  
21 337 The second example relates to the *Ordering* type of sequence (Figure 2), and stems from two  
22  
23 338 common calls of putty-nosed monkeys, *Cercopithecus nictitans martini*. ‘Pyow’ calls can be produced  
24  
25 339 individually or in strings of pyows, and seem to be used by putty-nosed monkeys frequently when  
26  
27 340 leopards are detected in the environment (Arnold & Zuberbühler, 2006a), and more generally as an  
28  
29 341 attention-getting signal related to recruitment of receivers and low level alarm (Arnold & Zuberbühler,  
30  
31 342 2013). ‘Hack’ calls can also be produced individually or in strings of hacks, and seem to be used  
32  
33 343 frequently when eagles are detected in the environment, and more generally as a higher-level alarm call  
34  
35 344 (Arnold & Zuberbühler, 2013). Importantly, pyow and hack calls are frequently combined into pyow-  
36  
37 345 hack sequences. Both naturalistic observational data as well as experimental call playback results indicate  
38  
39 346 that pyow-hack sequences influence receiver behaviour differently than do pyow or hack sequences alone  
40  
41 347 – pyow-hack sequences seem to mean “let’s go!”, and produce greater movement distances in receivers  
42  
43 348 (Arnold & Zuberbühler, 2006b). The case of the pyow-hack sequence therefore seems to represent  
44  
45 349 something closer to the notion of lexical syntax – individual units and ordered combinations of those units  
46  
47 350 have distinct meanings from one another (Marler, 1977).  
48  
49  
50  
51

52  
53 351 These two examples of primate calls illustrate the simple but important point that sequences  
54  
55 352 matter in acoustic signals – combinations or different linear orderings of units (whether those units have  
56  
57 353 meaning individually or not) can have different meanings to receivers. In the case of titi monkeys, the call  
58  
59  
60

sequences seem to serve the function of female attraction for male signallers, whereas in the case of putty-nosed monkeys, the call sequences serve anti-predatory and group cohesion functions.

We have so far been somewhat cavalier in how we have described the structures of call sequences, using terms like notes, units, and, indeed, calls. In the next section of our review, we describe in depth the notion of signalling ‘units’ in the acoustic modality.

### III. ACOUSTIC UNITS

Sequences are made of constituent units. Thus the accurate analysis of potential information in animal acoustic sequences depends on appropriately characterising their constituent acoustic units. We recognise, however, that there is no single definition of a unit. Indeed definitions of units, how they are identified, and the semantic labels we assign them vary widely across researchers working with different taxonomic groups (Gerhardt & Huber, 2002) or even within taxonomic groups, as illustrated by the enormous number of names for different units in the songs of songbird species. Our purpose in this section is to discuss issues surrounding the various ways the acoustic units composing a sequence may be characterised.

Units may be identified based on either production mechanisms, which focus on how the sounds are generated by signallers, or by perceptual mechanisms, which focus on how the sounds are interpreted by receivers. How we define a unit will therefore be different if the biological question pertains to production mechanisms or perceptual mechanisms. For example, in birdsong even a fairly simple note may be the result of two physical production pathways, each made on a different side of the syrinx (Catchpole & Slater, 2003). In practice, however, the details of acoustic production and perception are often hidden from the researcher, and so the definition of acoustic units is often carried out on the basis of observed acoustic properties (see Catchpole & Slater, 2003). It is not always clear to what extent these observed acoustic properties accurately represent the production/perceptual constraints on

1  
2  
3 379 communication, and the communicative role of the sequence. Identifying units is made all the more  
4  
5 380 challenging because acoustic units produced by animals often exhibit graded variation in their features  
6  
7 381 (e.g., absolute frequency, duration, rhythm or tempo, or frequency modulation), but most analytical  
8  
9 382 methods for unit classification assume that units can be divided into discrete, distinct categories (e.g.,  
10  
11 383 Clark, Marler & Beeman, 1987). Thus, how we identify units may differ depending on whether the  
12  
13 384 biological question pertains to production mechanisms, perceptual mechanisms, or acoustical analyses of  
14  
15 385 information content in the sequences. If the unit classification scheme must reflect animal sound  
16  
17 386 production or perception, care must be taken to base unit identification on the appropriate features of a  
18  
19 387 signal, and features that are biologically relevant (e.g., Clemins & Johnson, 2006). In cases where  
20  
21 388 sequences carry meaning, it is likely that they can be correlated with observational behaviours (possibly  
22  
23 389 context-dependent) observed over a large number of trials. There is still no guarantee that the sequence  
24  
25 390 assigned by the researcher is representative of the animal's perception of the same sequence. To some  
26  
27 391 degree, this can be tested with playback trials where the signals are manipulated with respect to the  
28  
29 392 hypothesised unit sequence (Kroodsma, 1989; Fischer, Noser & Hammerschmidt, 2013).

30  
31  
32  
33  
34 393       Whatever technique for identifying potential acoustic units is used, we emphasise here that there  
35  
36 394 are four acoustic properties that are commonly used to delineate potential units (Figure 3). First, the  
37  
38 395 spectrogram may show a silent gap between two acoustic elements (Figure 3a). When classifying units  
39  
40 396 “by eye”, separating units by silent gaps is probably the most commonly used criterion. Second,  
41  
42 397 examination of a spectrogram may show that an acoustic signal changes its properties at a certain time,  
43  
44 398 without the presence of a silent “gap” (Figure 3b). For example, a pure tone may become harmonic or  
45  
46 399 noisy, as the result of the animal altering its articulators (e.g., lips), without ceasing sound production in  
47  
48 400 the source (e.g., larynx). Third, a series of similar sounds may be grouped together as a single unit,  
49  
50 401 regardless of silent gaps between them, and separated from dissimilar units (Figure 3c). This is  
51  
52 402 characteristic of pulse trains and “trills”. Finally, there may be a complex hierarchical structure to the  
53  
54 403 sequence, in which combinations of sounds, which might otherwise be considered fundamental units,  
55  
56 404 always appear together, giving the impression of a coherent, larger unit of communication (Figure 3d). A  
57  
58  
59  
60

consideration of these four properties together can provide valuable insights into defining units of production, units of perception, and units for sequence analyses.

In Table 1, we give examples of the wide range of studies that have used these different criteria for dividing acoustic sequences into units. Although not intended to be comprehensive, the table shows how all of the four criteria listed above have been used for multiple species and with multiple aims – whether simply characterising the vocalisations, defining units of production/perception, or identifying the functional purpose of the sequences.

#### (1) *Identifying potential units*

Before we discuss in more detail how acoustic units may be identified in terms of production, perception, and analysis methods, we point out here that practically all such efforts require scientists to identify *potential* units at some early stage of their planned investigation or analysis. Two practical considerations are noteworthy.

First, a potential unit can be considered that part of a sequence that can be replaced with a label for analysis purposes (e.g., unit *A* or unit *B*), without adversely affecting the results of a planned investigation or analysis. Because animal acoustic sequences are sometimes hierarchical in nature (e.g., humpback whale *Megaptera novaengliae* song, reviewed in Cholewiak, Sousa-Lima & Cerchio, 2012), distinct sequences of units may themselves be organised into longer, distinctive sequences (i.e., “sequences of sequences”, Berwick et al., 2011). Thus, an important consideration in identifying potential acoustic units for sequence analyses is that they can be hierarchically nested, such that a sequence of units can itself be considered as a unit and replaced with a label.

Second, potential acoustic units are almost always identified based on acoustic features present in a spectrographic representation of the acoustic waveform. Associating combinations of these features with a potential unit can be performed either manually (i.e., examining the spectrograms “by eye”), or automatically by using algorithms for either supervised classification (where sounds are placed in categories according to pre-defined exemplars) or unsupervised clustering (where labelling units is

performed without prior knowledge of the types of units that occur). We return to these analytical methods in a subsequent section, and elaborate here on spectrographic representations.

Spectrograms consisting of discrete Fourier transforms of short, frequently overlapped, segments of the signal are ubiquitous and characterise well those acoustic features related to spectral profile and frequency modulation, many of which are relevant in animal acoustic communication. Examples of such features include minimum and maximum fundamental frequency, slope of the fundamental frequency, number of inflection points, and the presence of harmonics (Oswald et al., 2007) that vary, for example, between individuals (Buck & Tyack, 1993; Blumstein & Munos, 2005; Koren & Geffen, 2011; Ji et al., 2013; Kershenbaum, Sayigh & Janik, 2013; Root-Gutteridge et al., 2014), and in different environmental and behavioural contexts (Matthews et al., 1999; Taylor, Reby & McComb, 2008; Henderson, Hildebrand & Smith, 2011).

Other less used analytical techniques, such as cepstral analysis, may provide additional detail on the nature of acoustic units, and are worth considering for additional analytical depth. Cepstra are the Fourier (or inverse Fourier) transform of the log of the power spectrum (Oppenheim & Schaffer, 2004), and can be thought of as producing a spectrum of the power spectrum. Discarding coefficients can yield a compact representation of the spectrum (Figure 4). Further, while Fourier transforms have uniform temporal and frequency resolution, other techniques vary this resolution by using different basis sets, and this provides improved frequency resolution at low frequencies and better temporal resolution at higher frequencies. Examples of these other techniques include multi-taper spectra (Thomson, 1982; Tchernichovski et al., 2000; Baker & Logue, 2003), Wigner-Ville spectra (Martin & Flandrin, 1985; Cohn, 1995), and wavelet analysis (Mallat, 1999). While spectrograms and cepstra are useful for examining frequency-related features of signals, they are less useful when analysing temporal patterns of amplitude modulation. This is an important issue worth bearing in mind, because amplitude modulations are probably critical in signal perception by many animals (Henry et al., 2011), including speech perception by humans (Remez et al., 1994).

(2) *Identifying production units*

One important approach to identifying acoustic units stems from considering the mechanisms for sound production. In stridulating insects, for example, relatively simple, repeated sounds are typically generated by musculature action that causes hard physical structures to be engaged, such as the file and scraper located on the wings of crickets or the tymbal organs of cicadas (Gerhardt & Huber, 2002). The resulting units, variously termed “chirps,” or, “pulses,” can be organised into longer temporal sequences often termed “trills” or “echemes” (Ragge & Reynolds, 1988). Frogs can produce sounds with temporally structured units in a variety of ways (Martin & Gans, 1972; Martin, 1972; Gerhardt & Huber, 2002). In some species, a single acoustic unit (sometimes called a “pulse,” “note,” or a “call”) is produced by a single contraction of the trunk and laryngeal musculature that induces vibrations in the vocal folds (e.g., Girgenrath & Marsh, 1997). In other instances, frogs can generate short sequences of distinct sound units (also often called “pulses”) produced by the passive expulsion of air forced through the larynx that induces vibrations in structures called arytenoid cartilages, which impose temporal structure on sound (Martin & Gans, 1972; Martin, 1972). Many frogs organise these units into trills (e.g., Gerhardt, 2001), while other species combine acoustically distinct units (e.g., Narins, Lewis & McClelland, 2000; Larson, 2004). In songbirds, coordinated control of the two sides of the syrinx can be used to produce different units of sound, or “notes” (Suthers, 2004). These units can be organised into longer sequences, of “notes,” “trills,” “syllables,” “phrases,” “motifs,” and “songs” (Catchpole & Slater, 2003). In most mammals, sounds are produced as an air source (pressure squeezed from the lungs) causes vibrations in the vocal membranes, which are then filtered by a vocal tract (Peterson & Barney, 1952; Titze, 1994). When resonances occur in the vocal tract, certain frequencies known as formants are reinforced. Formants and formant transitions have been strongly implicated in human perception of vowels and voiced consonants, and may also be used by other species to perceive information (Raemaekers, Raemaekers & Haimoff, 1984; Fitch, 2000).

As the variety in these examples illustrates, there is incredible diversity in the mechanisms animals use to produce the acoustic units that are subsequently organised into sequences. Moreover, there



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

are additional mechanisms that constrain the production of some of the units. For example, in zebra finches *Taeniopygia guttata*, songs can be interrupted between some of its constitutive units but not others (Cynx, 1990). This suggests that at a neuronal level, certain units share a common, integrated neural production mechanism. Such examples indicate that identifying units based on metrics of audition or visual inspection of spectrograms (e.g., based on silent gaps) may not always be justified, and that there may be essential utility that emerges from a fundamental understanding of unit production. Thus, a key consideration in identifying functional units of production is that doing so may often require knowledge about production mechanisms that can only come about through rigorous experimental studies.

(3) *Identifying perceptual units*

While there may be fundamental insights gained from identifying units based on a detailed understanding of sound production, there may not always be a one-to-one mapping of the units of production or the units identified in acoustics analyses, onto units of perception (e.g., Blumstein, 1995). Three key considerations should be borne in mind when thinking about units of perception and the analysis of animal acoustic sequences (Figure 5).

First, it is possible that units of production or the units a scientist might identify on a spectrogram are perceptually bound together by receivers into a single unit of perception (Figure 5a). In this sense, a unit of perception is considered a perceptual auditory object in terms familiar to cognitive psychologists and auditory scientists. There are compelling reasons for researchers to consider vocalisations and other sounds as auditory objects (Miller & Cohen, 2010). While the rules governing auditory object formation in humans have been well studied (Griffiths & Warren, 2004; Bizley & Cohen, 2013), the question of precisely how, and to what extent, non-humans group acoustic information into coherent perceptual representations remains a largely open empirical question (Hulse, 2002; Bee & Micheyl, 2008; Miller & Bee, 2012).



Second, studies of categorical perception in humans and other animals (Harnad, 1990) show that continuous variation can nevertheless be perceived as forming discrete categories. In the context of units of perception, this means that the graded variation often seen in spectrograms may nevertheless be perceived categorically by receivers (Figure 5b). Thus, in instances where there are few discrete differences in production mechanisms or in spectrograms, receivers might still perceive distinct units (Nelson & Marler, 1989; Baugh, Akre & Ryan, 2008).

Third, well-known perceptual constraints related to the limits of spectrotemporal resolution may identify units of perception in ways that differ from analytical units and the units of production (Figure 5c). For example, due to temporal integration by the auditory system (Recanzone & Sutter, 2008), some short units of production might be produced so rapidly that they are not perceived as separate units. Instead, they might be integrated into a single percept having a pitch proportional to the repetition rate. For example, in both bottlenose dolphins *Tursiops truncatus* and Atlantic spotted dolphins *Stenella frontalis*, the “squawking” sound that humans perceive as having some tonal qualities is actually a set of rapid echolocation clicks known as a burst pulse (Herzing, 1996). The perceived pitch is related to the repetition rate, the faster the repetition, the higher the pitch. Given the perceptual limits of gap detection (Recanzone & Sutter, 2008), some silent gaps between units of production may be too short to be perceived by the receiver. Clearly, while it may sometimes be desirable or convenient to use “silence” as a way to create analysis boundaries between units, a receiver may not always perceive the silent gaps that we see in our spectrograms. Likewise, some transitions in frequency may reflect units of production that are not perceived because the changes remain unresolved by auditory filters (Moore & Moore, 2003; Recanzone & Sutter, 2008). Indeed, some species may be forced to trade off temporal and spectral resolution to optimise signalling efficiency in different environmental conditions. Frequency modulated signals are more reliable than amplitude modulation in reverberant habitats, such as forests, so woodland birds are adapted to greater frequency resolution and poorer temporal resolution, while the reverse is true of grassland species (Henry & Lucas, 2010; Henry et al., 2011).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The question of what constitutes a unit that is perceptually meaningful to the animal demands rigorous experimental approaches that put this question to the animal itself. There simply is no convenient shortcut to identifying perceptual units. Experimental approaches ranging from operant conditioning (e.g., Dooling et al., 1987; Brown, Dooling & O'Grady, 1988; Dent et al., 1997; Tu, Smith & Dooling, 2011; Ohms et al., 2012; Tu & Dooling, 2012), to field playback experiments, often involving the habituation-discrimination paradigm (e.g., Nelson & Marler, 1989; Wyttenbach, May & Hoy, 1996; Evans, 1997; Searcy, Nowicki & Peters, 1999; Ghazanfar et al., 2001; Weiss & Hauser, 2002) have the potential to identify the boundaries of perceptual units. Playbacks additionally can determine whether units can be discriminated (as in 'go no-go' tasks stemming from operant conditioning), or whether they can be recognised and are functionally meaningful to receivers.

Obviously some animals and systems are more tractable than others when it comes to experimentally assessing units of perception, but those not easy to manipulate experimentally (e.g., baleen whales, Balaenopteridae) should not necessarily be excluded from communication sequence research, although the inevitable constraints must be recognised.

(4) *Identifying analytical units*

In many instances, it is desirable to analyse sequences of identified units in acoustic recordings without having *a priori* knowledge about how those units may be produced or perceived by the animals themselves. Such analyses are often a fundamental first step toward investigating the potential meaning of acoustic sequences. Before turning to our main discussion of algorithms for analysing sequences of identified units, we briefly discuss methods by which scientists can identify and validate units for sequence analyses from acoustic recordings.

Sounds are typically assigned classifications to units based on the consistency of acoustic characteristics. When feasible, external validation of categories (i.e., comparing animal behavioural

1  
2  
3 557 responses to playback experiments) should be performed. Even without directly testing hypotheses of  
4  
5 558 biological significance by playback experiment, there may be other indicators of the validity of a  
6  
7  
8 559 classification scheme based purely on acoustic similarity. For example, naive observers correctly divide  
9  
10 560 dolphin signature whistles into groups corresponding closely to the individuals that produced them  
11  
12 561 (Sayigh et al., 2007), and similar (but poorer) results are achieved using quantitative measures of  
13  
14 562 spectrogram features (Kershenbaum, Sayigh & Janik, 2013).  
15  
16

17 563       When classifying units on the basis of their acoustic properties, errors can occur both as the result  
18  
19 564 of perceptual bias, and as the result of poor repeatability. Perceptual bias occurs either when the  
20  
21 565 characteristics of the sound that are used to make the unit assignment are inappropriate for the  
22  
23 566 communication system being studied, or when the classification scheme relies too heavily on those  
24  
25 567 acoustic features that appear important to human observers. For example, analysing spectrograms with a  
26  
27 568 50 Hz spectral resolution would be appropriate for human speech, but not for Asian elephants (*Elephas*  
28  
29 569 *maximus*), which produce infrasonic calls that are typically between 14-24 Hz (Payne, Langbauer Jr &  
30  
31 570 Thomas, 1986), as details of the elephant calls would be unobservable. Features that appear important to  
32  
33 571 human observers may include tonal modulation shapes, often posed in terms of geometric descriptors,  
34  
35 572 such as "upsweep", "concave", and "sine" (e.g., Bazúa-Durán & Au, 2002), which are prominent to the  
36  
37 573 human eye, but may or may not be of biological relevance. Poor repeatability, or variance, can occur both  
38  
39 574 in human classification, as inter-observer variability, and in machine learning, where computer  
40  
41 575 classification algorithms can make markedly different decisions after training with data that are very  
42  
43 576 similar (overtraining). Poor repeatability can be a particular problem when the classification scheme  
44  
45 577 ignores, or fails to give sufficient weight to, the features that are of biological significance, or the  
46  
47 578 algorithm (human or machine) places too much emphasis on particular classification cues that are specific  
48  
49 579 to the examples used to learn the categories. Repeatability suffers particularly when analysing signals in  
50  
51 580 the presence of noise, which can mask fine acoustic details (Kershenbaum & Roch, 2013).  
52  
53  
54  
55  
56  
57  
58  
59  
60

Two approaches have been used to classify units by their acoustic properties: visual inspection of spectrograms, and application of automatic algorithms that assign classifications based on mathematical rules.

(a) Visual classification, “by eye”

Traditionally, units are “hand-scored” by humans searching for consistent patterns in spectrograms (or even listening to sound recordings without the aid of a spectrogram). Visual classification has been an effective technique that has led to many important advances in the study both of birdsong (e.g., Kroodsma, 1985; Podos et al., 1992, and reviewed in Catchpole & Slater, 2003), and acoustic sequences in other taxa (e.g., Narins, Lewis & McClelland, 2000; Larson, 2004). Humans are usually considered to be good at visual pattern recognition – and better than most computer algorithms (Ripley, 2007; Duda, Hart & Stork, 2012), which makes visual classification an attractive approach to identifying acoustic units. However, drawbacks to visual classification exist (Clark, Marler & Beeman, 1987). Visual classification is time consuming and prevents taking full advantage of large acoustic data sets generated by automated recorders. Similarly, the difficulty in scoring large data sets means that sample sizes used in research may be too small to draw firm conclusions (Kershenbaum, 2013). Furthermore, visual classification can be prone to subjective errors (Jones, ten Cate & Bijleveld, 2001), and inter-observer reliability should be used (and reported) as a measure of the robustness of the visual assessments (Burghardt et al., 2012).

(b) Automatic classification

As an alternative to visual classification, automated methods may extract specific metrics, or features, from the acoustic data for input to classification algorithms. Although the acoustic features used by automated systems also may not reflect the cues used by the focal species, automated systems have the advantage of being able to handle large data sets. In principle, automatic classification is attractive as it is

not susceptible to the inter-observer variability of visual classification (Tchernichovski et al., 2000). However, current implementations generally fall short of the performance desired (Janik, 1999), for instance by failing to recognise subtle features that can be detected both by humans, and by the focal animals, and visual classification has been shown to out-perform automated systems in cases where the meaning of acoustic signals is known *a priori* (e.g., Sayigh et al., 2007; Kershenbaum, Sayigh & Janik, 2013). However, once an automatic algorithm is defined, large datasets can be analysed.

A third possibility is to use a hybrid system. Automated techniques can be used to find regions of possible calls that are then verified and corrected by a human analyst (Helble et al., 2012). Machine assistance can allow analysts to process much larger data sets than before, but at the risk of possibly missing calls that they might have been able to detect.

Classification algorithms can accept acoustic data with varying degrees of pre-processing as inputs. For example, in addition to the commonly used spectrograms (Picone, 1993), cepstra (Oppenheim & Schaffer, 2004), multi-taper spectra (Thomson, 1982), wavelets (Mallat, 1999), and formants (Fitch, 1997) may be used, as they provide additional information on the acoustic characteristics of units, which may not be well represented by traditional spectrograms (Tchernichovski et al., 2000). Each of these methods provide analysis of the spectral content of a short segment of the acoustic production, and algorithms frequently examine how these parameters are distributed or change over time (e.g., Kogan & Margoliash, 1998).

Units may be classified automatically using supervised algorithms, in which the algorithm is taught to recognise unit types given some *a priori* known exemplars, or clustered using unsupervised algorithms, in which no *a priori* unit type assignment is known (Duda, Hart & Stork, 2012). In both cases, the biological relevance of units must be verified independently because mis-specification of units can obscure sequential patterns. Environmental noise or sounds from other species may be mistakenly classified as an acoustic unit, and genuine units may be incorrectly assigned to unit categories. When

1  
2  
3 628 using supervised algorithms, perceptual bias may lead to misinterpreting data when the critical bands,  
4  
5 629 temporal resolution, and hearing capabilities of a species are not taken into account, i.e., the exemplars  
6  
7  
8 630 themselves may be subject to similar subjective errors that can occur in visual classification. However,  
9  
10 631 validation of unsupervised clustering into units is also problematic, where clustering results cannot be  
11  
12 632 assessed against known unit categories. The interplay between unit identification and sequence model  
13  
14 633 validation is a non-trivial problem (e.g., Jin & Kozhevnikov, 2011). Similarly, estimating uncertainty in  
15  
16 634 unit classification and assessing how that uncertainty affects conclusions from a sequence analysis is a  
17  
18 635 key part of model assessment (Duda, Hart & Stork, 2012)  
19

20  
21  
22 636 When using supervised classification, one appropriate technique for measuring classification  
23  
24 637 uncertainty is cross-validation (Arlot & Celisse, 2010). For fully unsupervised clustering algorithms,  
25  
26 638 where the desired classification is unknown, techniques exist to quantify the stability of the clustering  
27  
28 639 result, as an indicator of clustering quality. Examples include “leave-k-out” (Manning, Raghavan &  
29  
30 640 Schütze, 2008), a generalisation of the “leave-one-out” cross-validation, and techniques based on  
31  
32 641 normalised mutual information (Zhong & Ghosh, 2005), which measure the similarity between two  
33  
34 642 clustering schemes (Fred & Jain, 2005). However, it must be clear that cluster stability (and  
35  
36 643 correspondingly, inter-observer reliability) is not evidence that the classification is appropriate (i.e.,  
37  
38 644 matches the true, unknown, biologically relevant categorisation), or will remain stable upon addition of  
39  
40  
41 645 new data (Ben-David, Von Luxburg & Pál, 2006). Other information theoretic tests provide an alternative  
42  
43 646 assessment of the validity of unsupervised clustering results, such as checking if units follow Zipf’s law of  
44  
45 647 abbreviation, which is predicted by a universal principle of compression (Zipf, 1949; Ferrer-i-Cancho et  
46  
47 648 al., 2013) or Zipf’s law for word frequencies, which is predicted by a compromise between maximizing  
48  
49 649 the distinctiveness of units and the cost of producing them (Zipf, 1949; Ferrer-i-Cancho, 2005).  
50

51  
52 650  
53  
54  
55  
56  
57  
58  
59  
60

#### IV. THE STRUCTURE OF SEQUENCES

Given that the researcher has successfully determined the units of an acoustic sequence that are appropriate for the hypothesis being tested, one must select and apply appropriate algorithms for analysing the sequence of units. Many algorithms exist for the analysis of sequences: both those produced by animals, and sequences in general (such as DNA, and stock market prices). Selection of an appropriate algorithm can sometimes be guided by the quantity and variability of the data, but there is no clear rule to be followed. In fact, in machine learning, the so-called *no free lunch* theorem (Wolpert & Macready, 1997) shows that there is no one pattern recognition algorithm that is best for every situation, and any improvement in performance for one class of problems is offset by lower performance in another problem class. In choosing an algorithm for analyses, one should be guided by the variability and quantity of the data for analysis, keeping in mind that models with more parameters require more data to estimate the parameters effectively.

The structure analysis algorithms discussed throughout this section can be used to model the different methods for combining units discussed earlier (Figure 2). *Repetition*, *Diversity*, and *Ordering* are reasonably well captured by models such as Markov chains, hidden Markov models, and grammars. Networks capture structure either with or without order, although much of the application of networks has been done on unordered associations (*Combination*). Temporal information can be modelled as an attribute of an acoustic unit requiring extensions to the techniques discussed below, or as a separate process.

Here we give a sample of some of the more important and more promising algorithms for animal acoustic sequence analysis, and discuss ways for selecting and evaluating analytical techniques. Selecting appropriate algorithms should involve the following steps. (i) Technique: understand the nature of the models and their mathematical basis. (ii) Suitability: assess the suitability of the models and their constraints with respect to the research questions being asked. (iii) Application: apply the models to the empirical data (training, parameter estimation). (iv) Assessment: extract metrics from the models that



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

summarise the nature of the sequences analysed. (v) Inference: compare metrics between data sets (or between empirical data and random null-models) to draw ecological, mechanistic, evolutionary, and behavioural inferences. (vi) Validate: determine the goodness of fit of the model to the data and uncertainty of parameter estimates. Bootstrapping techniques can allow validation with sets that were not used in model development.

We consider five models in this section: (1) Markov chains, (2) hidden Markov models, (3) network models, (4) formal grammars, and (5) temporal models.

(1) *Markov chains*

Markov chains, or N-grams models, capture structure in acoustic unit sequences based on the recent history of a finite number of discrete unit types. Thus, the occurrence of a unit (or the probability of occurrence of a unit) is determined by a finite number of previous units. The history length is referred to as the order, and the simplest such model is a 0<sup>th</sup> order Markov model, which assumes that each unit is independent of one another, and simply determines the probability of observing any unit with no prior knowledge. A 1<sup>st</sup> order Markov model is one in which the probability of each unit occurring is determined only by the preceding unit, together with the “transition probability” from one unit to the next. This transition probability is assumed to be constant (stationary). Higher order Markov models condition the unit probabilities based on more than one preceding units, as determined by the model order. An N-gram model conditions the probability on the N-1 previous units, and is equivalent to an N-1<sup>th</sup> order Markov model. A  $K^{th}$  order Markov model of a sequence with  $C$  distinct units is defined by a  $C^K \times C$  matrix of transition probabilities from each of the  $C^K$  possible preceding sequences, to each of the  $C$  possible subsequent units, or equivalently by a state transition diagram (Figure 6).

As the order of the model increases, more and more data are required for the accurate estimation of transition probabilities, i.e., sequences must be longer, and many transitions will have zero counts. This is particularly problematic when looking at new data, which may contain sequences that were not previously encountered, as they will appear to have zero probability. As a result, Markov models with



orders greater than two (trigram,  $N=3$ ) are rare. In principle, a  $K^{\text{th}}$  order Markov model requires sufficient data to provide accurate estimates of  $C^{K+1}$  transition probabilities.

Closed-form expressions for maximum likelihood estimates of the transition probabilities can be used with conditional counts (Anderson & Goodman, 1957). For example, assuming five acoustic units (A-E), maximum likelihood estimates of the transition probabilities for a first order Markov model (bigram,  $N=2$ ) can be found directly from the number of occurrences of each transition, e.g.

$$P(B|A) = \frac{\text{count}(AB)}{\sum_{i \in \{A,B,C,D,E\}} \text{count}(A,i)}$$

Although not widely used in the animal communication literature, research in human natural language processing has led to the development of methods known as back-off models (Katz, 1987), which account for the underestimated probability of rare sequences using Good-Turing counts (Gale & Sampson, 1995). When a particular state transition is never observed in empirical data, the back-off model offers the minimum probability for this state transition so as not to rule it out automatically during the testing. Standard freely available tools such as the SRI language modelling toolkit (Stolcke, 2002) implement back-off models and can reduce the effort of adopting these more advanced techniques.

Once Markovian transitions have been calculated and validated, the transition probabilities can be used to calculate a number of summary metrics using information theory (Shannon et al., 1949; Chatfield & Lemon, 1970; Hailman, 2008). For a review on the mathematics underlying information theories, we direct the readers to the overview in McCowan, Hanser & Doyle (1999) or Freeberg & Lucas (2012), which provides the equations as well as a comprehensive reference list to other previous work. Here we will define these quantitative measures with respect to their relevance in analysing of animal acoustic sequences. Zero-order entropy measures repertoire diversity:

$$H_0 = \log_2(C)$$

where  $C=|V|$  is the cardinality of the set of acoustic units  $V$ . First-order entropy begins to measure simple repertoire internal organisational structure by evaluating the relative frequency of use of different signal types in the repertoire:

$$H_1 = \sum_{v_i \in V} -P(v_i) \log_2 P(v_i)$$

where the probability of each acoustic unit is typically estimated based on frequencies of occurrence, as described earlier with N-grams. Higher-order entropies measure internal organisational structure, and thus one form of communication complexity, by examining how signals interact within a repertoire at the two-unit sequence level, the three-unit sequence level, and so forth.

One inferential approach is to calculate the entropic values from first-order and higher-order Markov models to summarise the extent to which sequential structure is present at each order. A random sequence would show no dependence of entropy on Markov order, whereas decreases in entropy as the order is increased would be an indication of sequential organisation, and thus higher communication complexity (Ferrer-i-Cancho & McCowan, 2012). These summary measures can then be further extended to compare the importance of sequential structure across different taxa, social and ecological contexts. These types of comparisons can provide novel insights into the ecological, environmental, social, and contextual properties that shape the structure, organisation, and function of signal repertoires (McCowan, Doyle & Hanser, 2002).

The most common application of the Markov model is to test whether or not units occur independently in a sequence. Model validation techniques include the sequential and  $\chi^2$  tests (Anderson & Goodman, 1957). For instance, Narins, Lewis & McClelland (2000) used a permutation test (Adams & Anthony, 1996) to evaluate the hypothesis that a frog with an exceptionally large vocal repertoire, *Bufo madagascariensis*, emitted any call pairs more often than would be expected by chance. Similar techniques were used to show non-random call production by Sayigh et al., (2012) with short-finned pilot whales *Globicephala macrorhynchus*, and by Bohn et al., (2009) with free-tailed bats *Tadarida*

748 *brasiliensis*. However, deviation from statistical independence does not in itself prove a sequence to have  
749 been generated by a Markov chain. Other tests, such as N-gram distribution (Jin & Kozhevnikov, 2011)  
750 may be more revealing.

751

## 752 (2) *Hidden Markov models*

753 HMMs are a generalisation of the Markov model. In Markov models, the acoustic unit history (of length  
754 N) can be considered the current “state” of the system. In hidden Markov models (HMMs) (Rabiner,  
755 1989), states are not necessarily associated with acoustic units, but instead represent the state of some  
756 possibly unknown and unobservable process. Thus, the system progresses from one state to another,  
757 where the nature of each state is unknown to the observer. Each of these states may generate a “signal”  
758 (i.e., a unit), but there is not necessarily a one-to-one mapping between state transitions and signals  
759 generated. For example, transitioning to state X might generate unit A, but the same might be true of  
760 transitioning to state Y. An observation is generated at each state according to a state-dependent  
761 probability density function, and state transitions are governed by a separate probability distribution  
762 (Figure 7). HMMs are particularly useful to model very complex systems, while still being  
763 computationally tractable.

764 Extensions to the HMM model also exist, in which the state transition probabilities are non-  
765 stationary. For example, the probability of remaining in the same state may decay with time e.g., due to  
766 neural depletion, as shown by Jin & Kozhevnikov (2011), or recurrent units may appear more often than  
767 expected by a Markov model, particularly where behavioural sequences are non-Markovian (Cane, 1959;  
768 Kershenbaum, 2013). Also, HMMs are popular in speech analysis (Rabiner, 1989), where emissions are  
769 continuous-valued, rather than discrete.

770 HMMs have been used fairly extensively in speaker recognition (Lee & Hon, 1989), the  
771 identification of acoustic units in birdsong (Trawicki, Johnson & Osiejuk, 2005), and other analyses of  
772 bird song sequences. ten Cate, Lachlan & Zuidema (2013) reviewed analytical methods for inferring the  
773 structure of birdsong and highlighted the idea that HMM states can be thought of as possibly modelling

1  
2  
3 774 an element of an animal's cognitive state. This makes it possible to build models that have multiple state  
4  
5 775 distributions for the same acoustic unit sequence. For instance, in the trigram AAC, the probability given  
6  
7 776 by the 2<sup>nd</sup> order Markov model,  $P(C|A, A)$  is fixed. There cannot be different distributions for observing  
8  
9 777 the unit  $C$ , if the previous two units are  $A$ . Yet cognitive state may have the potential to influence the  
10  
11 778 probability of observing  $C$ , even for identical sequence contexts ( $AA$ ). Another state variable ( $\theta$ ) exists  
12  
13  
14 779 unobserved, as it reflects cognitive state, rather than sequence history. In this case,  $P(C|A, A, \theta=0) \neq P(C|A,$   
15  
16 780  $A, \theta=1)$ . Hahnloser, Kozhevnikov & Fee (2002), Katahira et al. (2011) and Jin (2009) have used HMMs to  
17  
18 781 model the interaction between song and neural substrates in the brain. A more recent example of this can  
19  
20 782 be seen in the work of Jin & Kozhevnikov (2011), where they used states to model neural units in song  
21  
22 783 production of the Bengalese finch *Lonchura striata* ver. *domestica*, restricting each state to the emission  
23  
24 784 of a single acoustic unit, thus making acoustic units associated with each state deterministic while  
25  
26 785 retaining the stochastic nature of state transitions.

27  
28  
29 786 Because the states of a HMM represent an unobservable process, it is difficult to estimate the  
30  
31 787 number of states needed to describe the empirical data adequately. Model selection methods and criteria  
32  
33 788 (for example Akaike and Bayesian information criteria, and others) can be used to estimate model order  
34  
35 789 (see Hamaker, Ganapathiraju & Picone, 1998, and Zucchini & MacDonald, 2009 for a brief review), so  
36  
37 790 the number of states is often determined empirically. Increasing the number of states permits the  
38  
39 791 modelling of more complex underlying sequences (e.g., longer term dependencies), but increases the  
40  
41 792 amount of data required for proper estimation. The efficiency and accuracy of model fitting depends on  
42  
43 793 model complexity, so that models with many states, many acoustic units, and perhaps many covariates or  
44  
45 794 other conditions will take more time and require more data to fit.

46  
47  
48 795 During training, HMM parameters are estimated using an optimisation algorithm (Cappé,  
49  
50 796 Moulines & Rydén, 2005) that finds a combination of hidden states, state transition tables, and state-  
51  
52 797 dependent distributions that best describe the data. Software libraries for the training of HMMs are  
53  
54 798 available in many formats (e.g., the Matlab function *hmmtrain*, the R package HMM; R Development  
55  
56  
57  
58  
59  
60

Team, 2012, and the Hidden Markov Model Toolkit; Young & Young, 1994). Similar considerations of dataset completeness exist to those when generating regular Markov models, most importantly, that long sequences of data are required.

Although the states of a HMM are sometimes postulated to possess biologically relevant meaning, the internal states of the HMM represent a hidden process, and do not necessarily refer to concrete behavioural states. Specifically, the training algorithm does not contain an optimisation criterion that will necessarily associate model states with the functional or ecological states of the animal that a researcher is interested in observing (e.g., foraging, seeking a mate, etc.). While the functional/ecological state is likely related to the sequence, each model state may in fact represent a different subsequence of the data. Therefore, one cannot assume in general that there will be a one-to-one mapping between model and animal states. Network structures derived from different empirical data are often widely different, and it can be misleading to make comparisons between HMMs derived from different data sets. Furthermore, obtaining consistent states requires many examples with respect to the diversity of the sequence being modelled. An overtrained network will be highly dependent on the data presented to it and small changes in the training data can result in very different model parameters, making state-based inference questionable.

### (3) *Network models*

The structure of an acoustic sequence can also be described using a network approach (reviewed in Newman, 2003 and Baronchelli et al., 2013), as has been done for other behavioural sequences (e.g., pollen processing by honeybees; Fewell, 2003). A node in the network represents a type of unit, and a directional edge connecting two nodes means that one unit comes after the other in the acoustic sequence. For example, if a bird sings a song in the order: *ABCABC*; the network representing this song will have three nodes for *A*, *B*, and *C*, and three edges connecting *A* to *B*, *B* to *C*, and *C* to *A* (Figure 8). The edges may simply indicate association between units without order (undirected binary network), an ordered

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

sequence (directed binary network), or a probability of an ordered sequence (directed weighted network), the latter being equivalent to a Markov chain (Newman, 2009).

The network representation is fundamentally similar to the Markov model, and the basic input for constructing a binary network is a matrix of unit pairs within the repertoire, which corresponds to the transition matrix in a Markov model. However, the network representation may be more robust than a Markov analysis, particularly when a large number of distinct unit types exist, precluding accurate estimation of transition probabilities (e.g., Sasahara et al., 2012). In this case, binary or simple directed networks may capture pertinent properties of the sequence, even if transition probabilities are unknown.

One of the attractive features of network analysis is that a large number of quantitative network measures exist for comparison to other networks (e.g., from different individuals, populations, or species), or for testing hypotheses. We list a few of the popular algorithms that can be used to infer the structure of the acoustic sequence using a network approach. We refer the reader to introductory texts to network analysis for further details (Newman, 2009; Scott & Carrington, 2011).

*Degree centrality* measures the number of edges directly connected to each node. In a directed network, each node has an in-degree and an out-degree, corresponding to incoming and outgoing edges. The weighted version of degree centrality is termed strength centrality, which takes into account the weights of each edge (Barrat et al., 2004). Degree/strength centrality identifies the central nodes in the network, corresponding to central elements in the acoustic sequence. For example, in the mockingbird *Mimus polyglottos*, which imitates sounds of other species, its own song is central in the network, meaning that it usually separates between other sounds by singing its own song (Gammon & Altizer, 2011).

*Betweenness centrality* is a measure of the role a central node plays in connecting other nodes. For example, if an animal usually uses three units before moving to another group of units, a unit that lies between these groups in the acoustic sequence will have high betweenness centrality. A weighted version of betweenness centrality was described in Opsahl, Agneessens & Skvoretz (2010).

*Clustering coefficient* describes how many triads of nodes are closed in the network. For example, if unit *A* is connected to *B*, and *B* is connected to *C*, a cluster is formed if *A* is also connected to *C*. Directed and weighted versions of the clustering coefficient have been described (Barrat et al., 2004; Fagiolo, 2007).

*Mean path length* is defined as the average minimum number of connections to be crossed from any arbitrary node to any other. This measures the overall navigability in the network; as this value becomes large, a longer series of steps is required for any node to reach another.

*Small-world metric* measures the level of connectedness of a network and is the ratio of the clustering coefficient *C* to the mean path length *L* after normalising each with respect to the clustering coefficient and mean path length of a random network:  $S = (C/C_{rand}) / (L/L_{rand})$ . If  $S > 1$  the network is regarded as “small-world” (Watts & Strogatz, 1998; Humphries & Gurney, 2008), with the implication that nodes are reasonably well connected and that it does not take a large number of edges to connect most pairs of nodes. Sasahara et al. (2012) demonstrated that the network of California thrasher songs has a small-world structure, in which subsets of phrases are highly grouped and linked with a short mean path length.

*Network motifs* are recurring structures that serve as building blocks of the network (Milo et al., 2002). For example, a network may feature an overrepresentation of specific types of triads, tetrads, or feed-forward loops. Network motif analysis could be informative in comparing sequence networks from different individuals, populations or species. We refer the reader to three software packages available for motif analysis: FANMOD (Wernicke & Rasche, 2006); MAVisto (Schreiber & Schwöbbermeyer, 2005); and MFinder (Kashtan et al., 2002).

*Community detection* algorithms offer a method to detect network substructure objectively (Fortunato, 2010). These algorithms identify groups of nodes with dense connections between them but that are sparsely connected to other groups/nodes. Subgroups of nodes in a network can be considered somewhat independent components of it, offering insight into the different subunits of acoustic



1  
2  
3 874 sequences. Multi-scale community detection algorithms can be useful for detecting hierarchical sequence  
4  
5 875 structures (Fushing & McAssey, 2010; Chen & Fushing, 2012).  
6

7 876 *Exponential family Random Graph Models* (ERGMs) offer a robust analytic approach to evaluate  
8  
9  
10 877 the contribution of multiple factors to the network structure using statistical modelling (Snijders, 2002).  
11  
12 878 These factors may include structural factors (e.g., the tendency to have closed triads in the network), and  
13  
14 879 factors based on node or edge attributes (e.g., a tendency for connections between nodes that are  
15  
16 880 acoustically similar). The goal of ERGMs is to predict the joint probability that a set of edges exist on  
17  
18 881 nodes in a network. The R programming language package *statnet* has tools for model estimation and  
19  
20 882 evaluation, and for model-based network simulation and network visualisation (Handcock et al., 2008).  
21  
22

23 883 As with other models, many statistical tests for inference and model assessment require a  
24  
25 884 comparison of the observed network to a set of random networks. For example, the clustering coefficient  
26  
27 885 of an observed network can be compared to those of randomly generated networks, to test if it is  
28  
29 886 significantly smaller or larger than expected. A major concern when constructing random networks is  
30  
31 887 what properties of the observed network should be retained (Croft, James & Krause, 2008). The answer to  
32  
33 888 this question depends on the hypothesis being tested. For example, when testing the significance of the  
34  
35 889 clustering coefficient, it is reasonable to retain the original number of nodes and edges, density and  
36  
37 890 possibly also the degree distribution, such that the observed network is compared to random networks  
38  
39 891 with similar properties.  
40  
41

42 892 Several software packages exist that permit the computation of many of the metrics from this  
43  
44 893 section that can be used to make inferences about the network. Examples include UCINET (Borgatti,  
45  
46 894 Everett & Freeman, 2002) , Gephi (Bastian, Heymann & Jacomy, 2009), igraph (Csardi & Nepusz, 2006),  
47  
48 895 and Cytoscape (Shannon et al., 2003).  
49  
50

51 896  
52  
53 897 (4) *Formal grammars*  
54  
55 898 The structure of an acoustic sequence can be described using formal grammars. A grammar consists of a  
56  
57 899 set of rewrite rules (or “productions”), that define the ways in which units can be ordered. Grammar rules  
58  
59  
60



consist of operations performed on “terminals” (in our case, units), which are conventionally denoted with lower case letters, and non-terminals (symbols that must be replaced by terminals before the derivation is complete), conventionally denoted with upper case letters (note that this convention is inconsistent with the upper case convention used for acoustic unit labels). Grammars generate sequences iteratively, by applying rules repeatedly to a growing sequence. For example, the rule “ $U \rightarrow a W$ ” means that the nonterminal  $U$  can be rewritten with the symbols “ $a W$ .” The terminal  $a$  is a unit, as we are familiar with, but as  $W$  is a non-terminal, and may itself be rewritten by a different rule. For an example, see Figure 9.

Sequences that can be derived by a given grammar are called grammatical with respect to that grammar. The collection of all sequences that could possibly be generated by a grammar is called the language of the grammar. The validation of a grammar consists of verifying that the grammar’s language matches exactly the set of sequences to be modelled. If a species produces sequences that cannot be generated by the grammar, the grammar is deemed “overselective”. A grammar that is “overgeneralising” produces sequences not observed in the empirical data – although it is often unclear whether this represents a true failure of the grammar, or insufficient sampling of observed sequences. In the example given in Figure 9, the grammar is capable of producing the sequence  $ABBBBBBBBBBBBB$ , however, since blue whales have not been observed to produce similar sequences in decades of observation, we conclude that this grammar is overgeneralising. It is important to note, however, that formal grammars are deterministic, in contrast to the probabilistic models discussed previously (Markov model, HMM). If one assigned probabilities to each of the rewriting rules, the particular sequence shown above may not have been observed simply because it is very *unlikely*.

Algorithms known as parsers can be constructed from grammars to estimate the probability that a sequence belongs to the language for which the grammar has been inferred. Inferring a grammar from a collection of sequences is a difficult problem, which, as famously formulated by (Gold, 1967), is intractable for all but a number of restricted cases. Gold’s formulation, however, does not appear to preclude the learning of grammar in real-world examples, and is of questionable direct relevance to the understanding or modelling of the psychology of sequence processing (Johnson, 2004). When restated in

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

terms that arguably fit better the cognitive tasks faced by humans and other animals, grammar inference becomes possible (Clark, 2010; Clark, Eyraud & Habrard, 2010). Algorithms based on distributional learning, which seek probabilistically motivated phrase structure by recursively aligning and comparing input sequences, are becoming increasingly successful in sequence processing tasks such as modelling language acquisition (Solan et al., 2005; Kolodny, Lotem & Edelman, in press).

A grammar can be classified according to its place in a hierarchy of classes of formal grammars known as the Chomsky hierarchy (Chomsky, 2002) and illustrated in Figure 10. These classes differ in the complexity of languages that can be modelled. The simplest class of grammars are called regular grammars, which are capable of describing the generation of any finite set of sequences or repeating pattern, and are fundamentally similar to Markov models. Figure 9 is an example of a regular grammar. Kakishita et al. (2009) showed that Bengalese finch songs can be modelled by a restricted class of regular grammars, called “k-reversible regular grammars,” which is learnable from only positive samples, i.e., observed and hence permissible sequences, without information on those sequences that are not permissible in the grammar. Context-free grammars are more complex, and are able to retain state information that enable one part of the sequence to affect another; this is usually demonstrated through the ability to create sequences of symbols where each unit is repeated the same number of times  $A^n B^n$  where  $n$  denotes  $n$  repetitions of the terminal unit, e.g.,  $AAABBB$  ( $A^3 B^3$ ). Context sensitive languages allow context dependent rewrite rules that have few restrictions, permitting further reaching dependencies such as in the set of sequences  $A^n B^n C^n$ . The highest level in the Chomsky hierarchy, recursively enumerable grammars, are more complex still, and rarely have relevance to animal communication studies.

The level of a grammar within the Chomsky hierarchy can give an indication of the complexity of the communication system represented by that grammar. Most animal acoustic sequences are thought to be no more complex than regular grammars (Berwick et al., 2011), whereas complexity greater than the regular grammar is thought to be a unique feature of human language (Hauser, Chomsky & Fitch, 2002). Therefore, indication that any animal communication could not be represented by a regular grammar

would be considered an important discovery, and claims of context-free (but non-regular) sequences in European starlings *Sturnus vulgaris* (Gentner et al., 2006) have not been widely accepted (Van Heijningen et al., 2009; Beckers et al., 2012). The deterministic nature of regular grammars – or indeed any formal grammars – may explain why formal grammars are not sufficiently general to describe the sequences of many animal species, and formal grammars remain more popular in human linguistic studies than in animal communication research.

#### (5) *Temporal structure*

Information may exist in the relative or absolute timing of acoustic units in a sequence, rather than in the order of those units. In particular, timing and rhythm information may be of importance, and may be lost when acoustic sequences are represented as a series of symbols. This section describes two different approaches to quantifying the temporal structure in acoustic sequences: traditional techniques examining inter-event interval and pulse statistics (e.g., Randall, 1989; Narins et al., 1992), and recent multi-timescale rhythm analysis (Saar & Mitra, 2008).

Analyses of temporal structure can be applied to any audio recording, regardless of whether that recording contains recognisable sequences, individual sounds, or multiple simultaneously vocalising individuals. Such analyses are most likely to be informative, however, when recurring acoustic patterns are present, especially if those recurring patterns are rhythmic or produced at a predictable rate. Variations in interactive sound sequence production during chorusing and cross-individual synchronisation can be quantified through meter, or prosody analysis, and higher-order sequence structure can be identified through automated identification of repeating patterns. At the simplest level, it is possible to analyse the timing of sounds in a sequence, simply by recording when sound energy is above a fixed threshold. For instance, temporal patterns can be extracted automatically from simpler acoustic sequences by transforming recordings into sequences of numerical measures of the durations and silent intervals between sounds (Isaac & Marler, 1963; Catchpole, 1976; Mercado, Herman & Pack, 2003; Handel, Todd & Zoidis, 2009; Green et al., 2011), song bouts (Eens, Pinxten & Verheyen, 1989; Saar &

1  
2  
3 978 Mitra, 2008), or of acoustic energy within successive intervals (Murray, Mercado & Roitblat, 1998;  
4  
5 979 Mercado et al., 2010). Before the invention of the Kay sonograph, which led to the routine analysis of  
6  
7 980 audio spectrograms, temporal dynamics of bird song were often transcribed using musical notation  
8  
9 981 (Saunders, 1951; Nowicki & Marler, 1988).

11 982 Inter-pulse interval has been widely used to quantify temporal structure in animal acoustic  
12  
13 983 sequences, for example in kangaroo rats *Dipodomys spectabilis* (Randall, 1989), fruit flies *Drosophila*  
14  
15 984 *melanogaster* (Bennet-Clark & Ewing, 1969), and rhesus monkeys *Macaca mulatta* (Hauser, Agnetta &  
16  
17 985 Perez, 1998). Variations in pulse intervals can encode individual information such as identity and fitness  
18  
19 986 (Bennet-Clark & Ewing, 1969; Randall, 1989), as well species identity (Randall, 1997; Hauser, Agnetta  
20  
21 987 & Perez, 1998). In these examples, comparing the median inter-pulse interval between two sample  
22  
23 988 populations is often sufficient to uncover significant differences.

24  
25  
26  
27 989 More recently developed techniques for analysis of temporal structure require more detailed  
28  
29 990 processing. For example, periodic regularities and repetitions of patterns within recordings of musical  
30  
31 991 performances can be automatically detected and characterised (Paulus, Müller & Klapuri, 2010; Weiss &  
32  
33 992 Bello, 2011). The first step in modern approaches to analysing the temporal structure of sound sequences  
34  
35 993 involves segmenting the recording. The duration and distribution of individual segments can be fixed  
36  
37 994 (e.g., splitting a recording into 100 ms chunks/frames) or variable (e.g., using multiple frame sizes in  
38  
39 995 parallel or adjusting the frame size based on the rate and duration of acoustic events). The acoustic  
40  
41 996 features of individual frames can then be analysed using the same signal processing methods that are  
42  
43 997 applied when measuring the acoustic features of individual sounds, thereby transforming the smaller  
44  
45 998 waveform into a vector of elements that describe features of the segment. Sequences of such frame-  
46  
47 999 describing vectors then would typically be used to form a matrix representing the entire recording. In this  
48  
49 1000 matrix, the sequence of columns (or rows) corresponds to the temporal order of individual frames  
50  
51 1001 extracted from the recording.

52  
53  
54  
55 1002 Regularities within the feature matrix generated from frame-describing vectors reflect temporal  
56  
57 1003 regularities within the original recording. Thus, the problem of describing and detecting temporal patterns  
58  
59  
60

1  
2  
3 1004 within a recording is transformed into the more computationally tractable problem of detecting and  
4  
5 1005 identifying structure within a matrix of numbers (as opposed to a sequence of symbols). If each frame is  
6  
7 1006 described by a single number (e.g., mean amplitude), then the resulting sequence of numbers can be  
8  
9 1007 analysed using standard time-frequency analysis techniques to reveal rhythmic patterns (Saar & Mitra,  
10  
11 1008 2008). Alternatively, each frame can be compared with every other frame to detect similarities using  
12  
13 1009 standard measures for quantifying the distance between vectors (Paulus, Müller & Klapuri, 2010). These  
14  
15 1010 distances are then often collected within a second matrix called a self-distance matrix. Temporal  
16  
17 1011 regularities within the original feature matrix are visible as coherent patterns with the self-distance matrix  
18  
19 1012 (typically showing up as patterned blocks or diagonal stripes). Various methods used for describing and  
20  
21 1013 classifying patterns within matrices (or images) can then be used to classify these two-dimensional  
22  
23 1014 patterns.

24  
25  
26  
27 1015 Different patterns in these matrices can be associated with variations in the novelty or  
28  
29 1016 homogeneity of the temporal regularities over time, as well as the number of repetitions of particular  
30  
31 1017 temporal patterns (Paulus, Müller & Klapuri, 2010). Longitudinal analyses of time-series measures of  
32  
33 1018 temporal structure can also be used to describe the stability or dynamics of rhythmic pattern production  
34  
35 1019 over time (Saar & Mitra, 2008). An alternative approach to identifying temporal structure within the  
36  
37 1020 feature matrix is to decompose it into simpler component matrices that capture the most recurrent features  
38  
39 1021 within the recording (Weiss & Bello, 2011). Similar approaches are common in modern analyses of high-  
40  
41 1022 density EEG recordings (Makeig et al., 2004). Algorithms for analysing the temporal dynamics of brain  
42  
43 1023 waves may thus also be useful for analysing temporal structure within acoustic recordings.

## 44 45 46 1024 47 48 1025 49 50 51 1026 V. FUTURE DIRECTIONS

52  
53 1027  
54  
55 1028 Many of the central questions in animal communication research focus on the meaning of signals and on  
56  
57 1029 the role of natural, sexual, and social selection on the evolution of communication systems. As shown in  
58  
59  
60

1  
2  
3 1030 Figure 2, information can exist in a sequence simultaneously via diversity, and order, as well as other less  
4  
5 1031 well-studied phenomena. Both natural and sexual selection may act on this information, either through  
6  
7 1032 conspecifics or heterospecifics (e.g., predators). This is especially true for animal acoustic sequences  
8  
9 1033 because the potential complexity of a sequence may imply greater scope for both meaning and selective  
10  
11 1034 pressure. Many new questions – and several old and unanswered ones – can be addressed by the  
12  
13 1035 techniques that we have outlined in this review. Some of the most promising avenues for future research  
14  
15 1036 are outlined below. First, we illustrate the integration of our framework using two case studies: the songs  
16  
17 1037 of rock hyraxes, *Procavia capensis*, and California thrashers, *Toxostoma redivivum*. Then, we discuss  
18  
19 1038 four outstanding questions in animal acoustic sequences that can potentially be addressed more  
20  
21 1039 effectively using the approaches proposed in this review.  
22  
23  
24 1040  
25  
26  
27 1041 (1) Two case studies  
28  
29 1042 (a) Rock hyraxes  
30  
31 1043 Kershenbaum et al. (2012) examined the syntactic structure of hyrax vocal sequences. They selected units  
32  
33 1044 using the separation by silence approach (Figure 3a). A histogram showing the frequency of different  
34  
35 1045 lengths of silence in the recordings was strongly bimodal (Kershenbaum et al., 2012:Figure S3),  
36  
37 1046 supporting the separation by silence paradigm. From these recordings, five distinct syllable types were  
38  
39 1047 identified. Previous studies (Koren & Geffen, 2009; Koren & Geffen, 2011) had identified only three  
40  
41 1048 types of hyrax syllables, but Kershenbaum et al. (2012) felt that one of the syllable categories (“wail”)  
42  
43 1049 was too general for their data, and that continuous variation in the length and frequency modulation  
44  
45 1050 patterns in the “wail” syllable could be perceived as different units by the receiver (Figure 5b). To avoid  
46  
47 1051 observation bias, they used an automatic system to measure the length and bandwidth of the purportedly  
48  
49 1052 “wail” syllables, and divided them into two further categories, based on fixed thresholds. Inspection “by  
50  
51 1053 eye” showed that only 10% of automatically classified syllables needed manual reclassification.  
52  
53  
54  
55 1054 The analysis of these hyrax vocalisations followed the *Ordering* paradigm (Figure 2c). No  
56  
57 1055 attempt was made to fit these data to one of the models described in Section IV.  
58  
59  
60

The main finding of this study was to show the presence of geographical dialects, by quantifying the differences in order (syntax) using the edit (or Levenshtein) distance (Garland et al., 2012), via the Needleman-Wunsch algorithm (Needleman & Wunsch, 1970). Using a quantitative tool to compare sequence similarity, and demonstrating that hyrax vocalisations have syntactic structure, the study showed that complex vocalisation syntax is more common in mammals than commonly believed.

(b) *California thrashers*

Sasahara et al. (2012) examined the song sequences of the California thrasher, a bird with an extensive repertoire (over 180 song types in this study). Acoustic units were separated by silence (Figure 3a), and classified “by eye” against a library catalogue of phrase types. However, to assess the reliability of the classification criteria, the researchers trained a support vector machine (or a support vector network) on a sample of manually classified units, and ensured that the automatic and manual classifications were consistent. Although song phrases were typically monosyllabic, if the time between units was small, units were grouped into a larger “phrase” (Figure 5c). Similarly two or more units were considered a single unit if they were always associated with each other in a series (Figure 2c). Although not analysed in Sasahara et al. (2012), subsequent unpublished work indicated that sequences of phrases can be grouped into higher levels of organisation (Figure 3d), with longer intervals between them.

Although Markov and hidden Markov models work well for some species, these models are unlikely to work for California thrashers because of the very large phrase repertoire; nodes in the model appear to grow without limit, and estimation of the transition probabilities will be inaccurate. Sasahara et al. (2012) used a network model, to bypass these problems.

Figure 11 shows a sample of the analysed data, with the phrase types as nodes, and the colour of directed lines indicating the proportion of observed transitions. The network structure for California thrasher songs varies between individuals, as indicated by network metrics of mean path length, clustering coefficient, and degree centrality (Section IV.3). All individuals showed networks with a “small world”



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1081 structure. Individual networks also showed variation in network metrics with time, indicating that song  
1082 structure was non-stationary.

1083       Although traditional analytical techniques such as Markov models would have been inappropriate  
1084 for these data, by using new and alternative approaches (a network model), the authors extracted  
1085 quantitative metrics on song structure, and successfully used these to compare the songs of different  
1086 individuals. These techniques are a promising approach for the highly complex sequences of species such  
1087 as the thrasher, mockingbird, and humpback whales.

1088

1089 (2)     *Outstanding questions in animal vocal sequences*

1090       (a)     *How do we define communication complexity?*

1091 Perceptual and developmental constraints may drive selection for communication complexity. However,  
1092 complexity can exist at any one (or more) of the six levels of information encoding that we have detailed,  
1093 often leading to definitions of “communication complexity” that are inconsistent between researchers. As  
1094 it is likely that no one definition of communication complexity can be universally valid, any definition  
1095 has relevance only after choosing to which of the encoding paradigms described in Figure 2 – or  
1096 combination thereof – it applies. Complexity defined, say, for the *Repetition* paradigm (Figure 2a) and  
1097 quantified as pulse rate variation, is not easily compared with *Diversity* complexity (Figure 2b), typically  
1098 quantified as repertoire size.

1099       For example, one hotly debated subject is whether selection from increased social complexity or  
1100 sexual selection is associated with increased acoustic complexity (Pollard & Blumstein, 2012), with  
1101 “complexity” defined as repertoire size (Figure 2b). Some researchers have proposed the idea that  
1102 communicative complexity, again defined as repertoire size, has at least in some species been driven by  
1103 the need to encode more information, or redundant information, in a complex social environment  
1104 (Freeberg et al., 2012). Alternatively, complexity metrics that measure *Ordering* (Figure 2d), often based  
1105 on non-zero orders of entropy (McCowan, Hanser & Doyle, 1999; Kershenbaum, 2013), may be more  
1106 biologically relevant in species that use unit ordering to encode information. Understanding the variety of

sequence types is essential to choosing the relevant acoustic unit definitions, and without this, testing competitive evolutionary hypotheses becomes problematic.

(b) *What is the role of sequence dialects in speciation?*

It is tempting to speculate that acoustic sequences may have arisen from earlier selective forces acting on a communication system based on single units. Alternatively, however, sequences could arise by neutral processes analogous to drift. A complex interplay between production, perception, and encoding of information in sequence syntax (along with the large relative differences between different species in adaptive flexibility, Seyfarth & Cheney, 2010) could lead to adaptive pressures on communication structure. In a few species, geographic syntactic dialects (Nettle, 1999) have been demonstrated (e.g., Rhesus monkeys *Macaca mulatta*, Gouzoules, Gouzoules & Marler, 1984; chimpanzees *Pan troglodytes*, Arcadi, 1996; Mitani, Hunley & Murdoch, 1999; Crockford & Boesch, 2005; Carolina chickadees *Poecile carolinensis*, Freeberg, 2012; and rock hyraxes *Procavia capensis*, Kershenbaum et al., 2012) raising the question of whether sequence syntax has a role in speciation (Wiens, 1982; Nevo et al., 1987; Irwin, 2000; Lachlan et al., 2013). However, the definition of acoustic units is rarely considered. In particular, perceptual binding (Figure 5a) and the response of the focal species must be considered, as reproductive isolation cannot occur on the basis of differences that are not perceived by the receiver.

(c) *How do individual differences in acoustic sequences arise?*

The proximal processes driving individual differences in communicative sequences are rarely investigated. Likewise, although there is a decades-rich history of song learning studies in songbirds, the ontogenetic processes giving rise to communicative sequences *per se* have rarely been studied. Neural models (e.g., Jin, 2009) can provide probabilistic descriptions of sequence generation (e.g., Markov models, hidden Markov models), but the nature of the underlying stochasticity is unknown. The California thrasher example given above (Figure 11), shows how an appropriate choice of a model for sequence structure can allow quantitative comparisons between the parameters of different individuals.

1  
2  
3 1133 However, model fitting is only valid if unit selection is biologically appropriate (Section III). Other, more  
4  
5 1134 abstract, questions can also be addressed. Individual humans use language with varying degrees of  
6  
7 1135 efficiency, creativity, and effectiveness. Shakespearean sequences are radically unlike Haiku sequences,  
8  
9 1136 political speeches, or the babbling of infants, in part because their communicative purposes differ. While  
10  
11 1137 sexual selection and survival provide some purposive contexts through which we can approach meaning,  
12  
13 1138 additional operative contexts may suggest other purposes, and give us new frameworks through which to  
14  
15 1139 view vocal sequences (Waller, 2012).  
16  
17 1140  
18  
19  
20  
21 1141 (d) *How might information exist within units themselves?*  
22  
23 1142 Another promising direction lies in studying how animals use concatenated signals with multiple  
24  
25 1143 meanings. For example, Jansen, Cant & Manser (2012) provided evidence for temporal segregation of  
26  
27 1144 information within a syllable, where one segment of a banded mongoose *Mungos mungo* close call is  
28  
29 1145 individually distinct, while the other segment contains meaning about the caller's activity. Similar results  
30  
31 1146 have been demonstrated in the song of the white-crowned sparrow *Zonotrichia leucophrys* (Nelson &  
32  
33 1147 Poesel, 2007). Understanding how to divide acoustic units according to criteria other than silent gaps  
34  
35 1148 (Figure 3) can change the research approach, as well as the results of a study. The presence of information  
36  
37 1149 in sub-divisions of traditional acoustic units is a subject underexplored in the field of animal  
38  
39 1150 communication, and an understanding of the production and perceptual constraints on unit definition  
40  
41 1151 (Figure 5) is essential.  
42  
43 1152  
44  
45  
46 1153 We conclude by noting that more detailed and rigorous approaches to investigating animal acoustic  
47  
48 1154 sequences will allow us to investigate more complex systems that have not been formally studied. A  
49  
50 1155 number of directions lack even a basic framework as we have proposed in this review. For example, there  
51  
52 1156 is much to be learned from the detailed study of the sequences created by multiple animals vocalising  
53  
54 1157 simultaneously, and from the application of sequence analysis to multimodal communication with a  
55  
56 1158 combination of acoustic, visual, and perhaps other modalities (e.g., Partan & Marler, 1999; Bradbury &  
57  
58  
59  
60

Vehrencamp, 2011; Munoz & Blumstein, 2012). Eavesdropping, in which non-target receivers (such as predators) gain additional information from listening to the interaction between individuals, has only just begun to be studied in the context of sequence analysis. Finally, the study of non-stationary systems, where the statistical nature of the communicative sequences changes over long or short time scales (such as appears to occur in humpback whale songs) is ripe for exploration. We encourage researchers in these fields to extend treatments such as ours to cover these more complex directions in animal communication research, thereby facilitating quantitative comparisons between fields.

## VI. CONCLUSIONS

- (1) The use of acoustic sequences by animals is widespread across a large number of taxa. As diverse as the sequences themselves is the range of analytical approaches used by researchers. We have proposed a framework for analysing and interpreting such acoustic sequences, based around three central ideas of understanding the information content of sequences, defining the acoustic units that comprise sequences, and proposing analytical algorithms for testing hypotheses on empirical sequence data.
- (2) We propose use of the term “meaning” to refer to a feature of communication sequences that influences behavioural and evolutionary processes, and the term “information” to refer to the non-random statistical properties of sequences.
- (3) Information encoding in acoustic sequences can be classified into six non-mutually exclusive paradigms: *Repetition*, *Diversity*, *Combination*, *Ordering*, *Overlapping*, and *Timing*.
- (4) The constituent units of acoustic sequences can be classified according to production mechanisms, perception mechanisms, or analytical properties.
- (5) Discrete acoustic units are often delineated by silent intervals. However, changes in the acoustic properties of a continuous sound may also indicate a transition between discrete units, multiple repeated sounds may act as a discrete unit, and more complex hierarchical structure may also be present.

1  
2  
3 1183 (6) We have reviewed five approaches used for analysing the structure of animal acoustic sequences:  
4  
5 1184 Markov chains, hidden Markov models, network models, formal grammars, and temporal models,  
6  
7 1185 discussing their use and relative merits.  
8  
9  
10 1186 (7) Many important questions in the behavioural ecology of acoustic sequences remain to be answered,  
11  
12 1187 such as understanding the role of communication complexity, including multimodal sequences, the  
13  
14 1188 potential effect of communicative isolation on speciation, and the source of syntactic differences between  
15  
16 1189 individuals.  
17  
18 1190  
19  
20 1191 VII. ACKNOWLEDGEMENTS  
21  
22 1192  
23  
24  
25 1193 This review was developed at an investigative workshop, “Analyzing Animal Vocal Communication  
26  
27 1194 Sequences” that took place on October 21-23 2013 in Knoxville, Tennessee, sponsored by the National  
28  
29 1195 Institute for Mathematical and Biological Synthesis (NIMBioS). NIMBioS is an Institute sponsored by  
30  
31 1196 the National Science Foundation, the U.S. Department of Homeland Security, and the U.S. Department of  
32  
33 1197 Agriculture through NSF Awards #EF-0832858 and #DBI-1300426, with additional support from The  
34  
35 1198 University of Tennessee, Knoxville. In addition to the authors, Vincent Janik participated in the  
36  
37 1199 workshop. DTB’s research is currently supported by NSF DEB-1119660. MAB’s research is currently  
38  
39 1200 supported by NSF IOS-0842759 and NIH R01DC009582. MAR’s research is supported by ONR  
40  
41 1201 N0001411IP20086 and NOPP (ONR/BOEM) N00014-11-1-0697. SLDeR’s research is supported by the  
42  
43 1202 U.S. Office of Naval Research. RF-i-C’s research was supported by the grant BASMATI (TIN2011-  
44  
45 1203 27479-C04-03) from the Spanish Ministry of Science and Innovation. ECG’s research is currently  
46  
47 1204 supported by a National Research Council postdoctoral fellowship. EEV’s research is supported by  
48  
49 1205 CONACYT, Mexico, award number I010/214/2012.  
50  
51  
52 1206  
53  
54  
55 1207  
56  
57  
58  
59  
60

# VIII. REFERENCES

- ADAMS, D. C. & ANTHONY, C. D. (1996). Using randomization techniques to analyse behavioural data. *Animal Behaviour* **51**, 733-738.
- AKÇAY, Ç, TOM, M. E., CAMPBELL, S. E. & BEECHER, M. D. (2013). Song type matching is an honest early threat signal in a hierarchical animal communication system. *Proceedings of the Royal Society B: Biological Sciences* **280**.
- ANDERSON, T. W. & GOODMAN, L. A. (1957). Statistical inference about Markov chains. *The Annals of Mathematical Statistics* **28**, 89-110.
- ARAYA-SALAS, M. (2012). Is birdsong music? *Significance* **9**, 4-7.
- ARCADI, A. C. (1996). Phrase structure of wild chimpanzee pant hoots: patterns of production and interpopulation variability. *American Journal of Primatology* **39**, 159-178.
- ARLOT, S. & CELISSE, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys* **4**, 40-79.
- ARNOLD, K. & ZUBERBÜHLER, K. (2013). Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PloS One* **8**, e65660.
- ARNOLD, K. & ZUBERBÜHLER, K. (2008). Meaningful call combinations in a non-human primate. *Current Biology* **18**, R202-R203.
- ARNOLD, K. & ZUBERBÜHLER, K. (2006a). The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour* **72**, 643-653.
- ARNOLD, K. & ZUBERBÜHLER, K. (2006b). Language evolution: semantic combinations in primate calls. *Nature* **441**, 303-303.
- BAKER, M. C. & BECKER, A. M. (2002). Mobbing calls of black-capped chickadees: effects of urgency on call production. *The Wilson Bulletin* **114**, 510-516.
- BAKER, M. C. & LOGUE, D. M. (2003). Population differentiation in a complex bird sound: a comparison of three bioacoustical analysis procedures. *Ethology* **109**, 223-242.
- BAPTISTA, L. F. & KEISTER, R. A. (2005). Why birdsong is sometimes like music. *Perspectives in Biology and Medicine* **48**, 426-443.
- BARONCHELLI, A., FERRER-I-CANCHO, R., PASTOR-SATORRAS, R., CHATER, N. & CHRISTIANSEN, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences* **17**, 348-360.
- BARRAT, A., BARTHELEMY, M., PASTOR-SATORRAS, R. & VESPIGNANI, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3747-3752.

1  
2  
3 1241 BASTIAN, M., HEYMANN, S. & JACOMY, M. (2009). Gephi: an open source software for exploring  
4 1242 and manipulating networks. *Proceedings of the Third International ICWSM Conference* .  
5  
6  
7 1243 BAUGH, A., AKRE, K. & RYAN, M. (2008). Categorical perception of a natural, multivariate signal:  
8 1244 mating call recognition in túngara frogs. *Proceedings of the National Academy of Sciences* **105**, 8985-  
9 1245 8988.  
10  
11 1246 BAZÚA-DURÁN, C. & AU, W. W. (2002). The whistles of Hawaiian spinner dolphins. *The Journal of*  
12 1247 *the Acoustical Society of America* **112**, 3064-3072.  
13  
14 1248 BECKERS, G. J., BOLHUIS, J. J., OKANOYA, K. & BERWICK, R. C. (2012). Birdsong  
15 1249 neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport* **23**, 139-145.  
16  
17  
18 1250 BEE, M. A. & MICHEYL, C. (2008). The cocktail party problem: what is it? How can it be solved? And  
19 1251 why should animal behaviorists study it? *Journal of Comparative Psychology* **122**, 235-251.  
20  
21 1252 BEN-DAVID, S., VON LUXBURG, U. & PÁL, D. (2006). A sober look at clustering stability. In  
22 1253 *Learning theory*, pp. 5-19. Springer.  
23  
24 1254 BENNET-CLARK, H. & EWING, A. (1969). Pulse interval as a critical parameter in the courtship song  
25 1255 of *Drosophila melanogaster*. *Animal Behaviour* **17**, 755-759.  
26  
27  
28 1256 BERGMAN, T. J., BEEHNER, J. C., CHENEY, D. L. & SEYFARTH, R. M. (2003). Hierarchical  
29 1257 classification by rank and kinship in baboons. *Science* **302**, 1234-1236.  
30  
31 1258 BERWICK, R. C., OKANOYA, K., BECKERS, G. J. L. & BOLHUIS, J. J. (2011). Songs to syntax: the  
32 1259 linguistics of birdsong. *Trends in Cognitive Sciences* **15**, 113-121.  
33  
34  
35 1260 BIZLEY, J. K. & COHEN, Y. E. (2013). The what, where and how of auditory-object perception. *Nature*  
36 1261 *Reviews Neuroscience* **14**, 693-707.  
37  
38 1262 BLUMSTEIN, D. T. (2007). The evolution, function, and meaning of marmot alarm communication.  
39 1263 *Advances in the Study of Behavior* **37**, 371-401.  
40  
41 1264 BLUMSTEIN, D. T. (1995). Golden-marmot alarm calls. II. Asymmetrical production and perception of  
42 1265 situationally specific vocalizations? *Ethology* **101**, 25-32.  
43  
44  
45 1266 BLUMSTEIN, D. T., VERNEYRE, L. & DANIEL, J. C. (2004). Reliability and the adaptive utility of  
46 1267 discrimination among alarm callers. *Proceedings of the Royal Society of London. Series B: Biological*  
47 1268 *Sciences* **271**, 1851-1857.  
48  
49 1269 BLUMSTEIN, D. T. & MUNOS, O. (2005). Individual, age and sex-specific information is contained in  
50 1270 yellow-bellied marmot alarm calls. *Animal Behaviour* **69**, 353-361.  
51  
52  
53 1271 BOHM, D. (1989). Meaning and information. In *The Search for Meaning*, pp. 43-62.  
54  
55 1272 BOHN, K. M., SCHMIDT-FRENCH, B., MA, S. T. & POLLAK, G. D. (2008). Syllable acoustics,  
56 1273 temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats. *The*  
57 1274 *Journal of the Acoustical Society of America* **124**, 1838-1848.  
58  
59  
60



- 1275 BOHN, K. M., SCHMIDT-FRENCH, B., SCHWARTZ, C., SMOTHERMAN, M. & POLLAK, G. D.  
1276 (2009). Versatility and stereotypy of free-tailed bat songs. *PloS One* **4**, e6746.
- 1277 BORGATTI, S. P., EVERETT, M. G. & FREEMAN, L. C. (2002). *Ucinet for Windows: Software for*  
1278 *social network analysis*. Analytic Technologies.
- 1279 BRADBURY, J. W. & VEHRENCAMP, S. L. (2011). *Principles of Animal Communication*. Sinauer,  
1280 Sunderland, MA.
- 1281 BRIEFER, E. F., RYBAK, F. & AUBIN, T. (2013). Does true syntax or simple auditory object support  
1282 the role of skylark song dialect? *Animal Behaviour* **86**, 1131-1137.
- 1283 BROWN, S. D., DOOLING, R. J. & O'GRADY, K. E. (1988). Perceptual organization of acoustic stimuli  
1284 by budgerigars (*Melopsittacus undulatus*): III. Contact calls. *Journal of Comparative Psychology* **102**,  
1285 236-247.
- 1286 BRUMM, H. & SLATER, P. J. (2006). Ambient noise, motor fatigue, and serial redundancy in chaffinch  
1287 song. *Behavioral Ecology and Sociobiology* **60**, 475-481.
- 1288 BUCK, J. R. & TYACK, P. L. (1993). A quantitative measure of similarity for *Tursiops truncatus*  
1289 signature whistles. *The Journal of the Acoustical Society of America* **94**, 2497-2506.
- 1290 BURGHARDT, G. M., BARTMESS-LEVASSEUR, J. N., BROWNING, S. A., MORRISON, K. E.,  
1291 STEC, C. L., ZACHAU, C. E. & FREEBERG, T. M. (2012). Perspectives—minimizing observer bias in  
1292 behavioral studies: A review and recommendations. *Ethology* **118**, 511-517.
- 1293 CALDWELL, M. C., CALDWELL, D. K. & TYACK, P. L. (1990). Review of the signature-whistle  
1294 hypothesis for the Atlantic bottlenose dolphin. In *The Bottlenose Dolphin* (eds. S. Leatherwood and R. R.  
1295 Reeves), pp. 199-234. Academic Press, San Diego.
- 1296 CALDWELL, M. C. (1965). Individualized whistle contours in bottle-nosed dolphins (*Tursiops*  
1297 *truncatus*). *Nature* **207**, 434-435.
- 1298 CALL, J. E. & TOMASELLO, M. E. (2007). *The gestural communication of apes and monkeys*. Taylor  
1299 & Francis Group/Lawrence Erlbaum Associates.
- 1300 CANE, V. R. (1959). Behaviour sequences as semi-Markov chains. *Journal of the Royal Statistical*  
1301 *Society. Series B (Methodological)* **21**, 36-58.
- 1302 CAPPÉ, O., MOULINES, E. & RYDÉN, T. (2005). *Inference in hidden Markov models*. Springer  
1303 Science Business Media, New York.
- 1304 CARTER, G. G., SKOWRONSKI, M. D., FAURE, P. A. & FENTON, B. (2008). Antiphonal calling  
1305 allows individual discrimination in white-winged vampire bats. *Animal Behaviour* **76**, 1343-1355.
- 1306 CĂSAR, C., BYRNE, R. W., HOPPITT, W., YOUNG, R. J. & ZUBERBÜHLER, K. (2012a). Evidence  
1307 for semantic communication in titi monkey alarm calls. *Animal Behaviour* **84**, 405-411.

1  
2  
3 1308 CÄSAR, C., BYRNE, R., YOUNG, R. J. & ZUBERBÜHLER, K. (2012b). The alarm call system of wild  
4 1309 black-fronted titi monkeys, *Callicebus nigrifrons*. *Behavioral Ecology and Sociobiology* **66**, 653-667.  
5  
6 1310 CÄSAR, C., ZUBERBÜHLER, K., YOUNG, R. J. & BYRNE, R. W. (2013). Titi monkey call sequences  
7 1311 vary with predator location and type. *Biology letters* **9**, 20130535.  
8  
9  
10 1312 CATCHPOLE, C. K. & SLATER, P. J. B. (2003). *Bird song: biological themes and variations*.  
11 1313 Cambridge Univ Press, Cambridge.  
12  
13 1314 CATCHPOLE, C. K. (1976). Temporal and sequential organisation of song in the sedge warbler  
14 1315 (*Acrocephalus schoenobaenus*). *Behaviour* **59**, 226-246.  
15  
16  
17 1316 CHATFIELD, C. & LEMON, R. E. (1970). Analysing sequences of behavioural events. *Journal of*  
18 1317 *Theoretical Biology* **29**, 427-445.  
19  
20 1318 CHATTERJEE, A. (2005). A madness to the methods in cognitive neuroscience? *Journal of Cognitive*  
21 1319 *Neuroscience* **17**, 847-849.  
22  
23  
24 1320 CHEN, C. & FUSHING, H. (2012). Multiscale community geometry in a network and its application.  
25 1321 *Physical Review E* **86**, 041120.  
26  
27 1322 CHOLEWIAK, D. M., SOUSA-LIMA, R. S. & CERCHIO, S. (2012). Humpback whale song  
28 1323 hierarchical structure: Historical context and discussion of current classification issues. *Marine Mammal*  
29 1324 *Science* **29**, E312-E332.  
30  
31 1325 CHOMSKY, N. (2002). *Syntactic structures, 9th edition*. de Gruyter Mouton, The Hague.  
32  
33  
34 1326 CHRISTIANSEN, M. H. & CHATER, N. (2001). Connectionist psycholinguistics: Capturing the  
35 1327 empirical data. *Trends in Cognitive Sciences* **5**, 82-88.  
36  
37 1328 CLARK, A. (2010). Towards general algorithms for grammatical inference. *Lecture Notes in Computer*  
38 1329 *Science* **6331**, 11-30.  
39  
40 1330 CLARK, A., EYRAUD, R. & HABRARD, A. (2010). Using contextual representations to efficiently  
41 1331 learn context-free languages. *The Journal of Machine Learning Research* **11**, 2707-2744.  
42  
43 1332 CLARK, C. W., MARLER, P. & BEEMAN, K. (1987). Quantitative analysis of animal vocal phonology:  
44 1333 an application to swamp sparrow song. *Ethology* **76**, 101-115.  
45  
46  
47 1334 CLEMINS, P. J. & JOHNSON, M. T. (2006). Generalized perceptual linear prediction features for animal  
48 1335 vocalization analysis. *The Journal of the Acoustical Society of America* **120**, 527-534.  
49  
50 1336 COHN, L. (1995). *Time-Frequency Analysis: Theory and Applications, 1995*. Prentice Hall.  
51  
52  
53 1337 CROCKFORD, C. & BOESCH, C. (2005). Call combinations in wild chimpanzees. *Behaviour* **142**, 397-  
54 1338 421.  
55  
56 1339 CROFT, D. P., JAMES, R. & KRAUSE, J. (2008). *Exploring Animal Social Networks*. Princeton  
57 1340 University Press.  
58  
59  
60

- 1  
2  
3 1341 CSARDI, G. & NEPUSZ, T. (2006). The igraph software package for complex network research.  
4 1342 *International Journal of Complex Systems* **1695**, 1-9.  
5  
6  
7 1343 CYNX, J. (1990). Experimental determination of a unit of song production in the zebra finch  
8 1344 (*Taeniopygia guttata*). *Journal of Comparative Psychology* **104**, 3-10.  
9  
10 1345 DARLING, J. D., JONES, M. E. & NICKLIN, C. P. (2006). Humpback whale songs: Do they organize  
11 1346 males during the breeding season? *Behaviour* **143**, 1051-1102.  
12  
13 1347 DAWKINS, R. & KREBS, J. R. (1978). Animal signals: information or manipulation. *Behavioural*  
14 1348 *ecology: An evolutionary approach* **2**, 282-309.  
15  
16  
17 1349 DEECKE, V. B. & JANIK, V. M. (2006). Automated categorization of bioacoustic signals: avoiding  
18 1350 perceptual pitfalls. *The Journal of the Acoustical Society of America* **119**, 645-653.  
19  
20 1351 DENT, M. L., BRITTAN-POWELL, E. F., DOOLING, R. J. & PIERCE, A. (1997). Perception of  
21 1352 synthetic/ba/-/wa/speech continuum by budgerigars (*Melopsittacus undulatus*). *The Journal of the*  
22 1353 *Acoustical Society of America* **102**, 1891-1897.  
23  
24 1354 DOOLING, R. J., PARK, T. J., BROWN, S. D., OKANOYA, K. & SOLI, S. D. (1987). Perceptual  
25 1355 organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals. *Journal of*  
26 1356 *Comparative Psychology* **101**, 367-381.  
27  
28  
29 1357 DUDA, R. O., HART, P. E. & STORK, D. G. (2012). *Pattern Classification*. John Wiley & Sons, New  
30 1358 York.  
31  
32 1359 EENS, M. (1997). Understanding the complex song of the European starling: an integrated ethological  
33 1360 approach. *Advances in the Study of Behavior* **26**, 355-434.  
34  
35  
36 1361 EENS, M., PINXTEN, R. & VERHEYEN, R. F. (1993). Function of the song and song repertoire in the  
37 1362 European starling (*Sturnus vulgaris*): an aviary experiment. *Behaviour* **125**, 51-66.  
38  
39 1363 EENS, M., PINXTEN, R. & VERHEYEN, R. F. (1991). Male song as a cue for mate choice in the  
40 1364 European starling. *Behaviour* **116**, 210-238.  
41  
42 1365 EENS, M., PINXTEN, R. & VERHEYEN, R. F. (1989). Temporal and sequential organization of song  
43 1366 bouts in the starling. *Ardea* **77**, 75-86.  
44  
45  
46 1367 ESSER, K., CONDON, C. J., SUGA, N. & KANWAL, J. S. (1997). Syntax processing by auditory  
47 1368 cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proceedings of the*  
48 1369 *National Academy of Sciences* **94**, 14019-14024.  
49  
50 1370 EVANS, C. S. (1997). Referential signals. *Perspectives in Ethology* **12**, 99-143.  
51  
52 1371 FAGIOLO, G. (2007). Clustering in complex directed networks. *Physical Review E* **76**, 026107.  
53  
54  
55 1372 FALLS, J. B. (1982). Individual recognition by sounds in birds. *Acoustic Communication in Birds* **2**, 237-  
56 1373 278.  
57  
58  
59  
60

1  
2  
3 1374 FARABAUGH, S. & DOOLING, R. (1996). Acoustic communication in parrots: laboratory and field  
4 1375 studies of budgerigars, *Melopsittacus undulatus*. In *Ecology and Evolution of Acoustic Communication in*  
5 1376 *Birds*, pp. 97-117. Cornell University Press Ithaca, New York, USA.

7  
8 1377 FERRER-I-CANCHO, R. (2005). Zipf's law from a communicative phase transition. *The European*  
9 1378 *Physical Journal B-Condensed Matter and Complex Systems* **47**, 449-457.

11 1379 FERRER-I-CANCHO, R. & MCCOWAN, B. (2012). The span of correlations in dolphin whistle  
12 1380 sequences. *Journal of Statistical Mechanics: Theory and Experiment* **2012**, P06002.

14 1381 FERRER-I-CANCHO, R., HERNÁNDEZ-FERNÁNDEZ, A., LUSSEAU, D., AGORAMOORTHY, G.,  
15 1382 HSU, M. J. & SEMPLE, S. (2013). Compression as a universal principle of animal behavior. *Cognitive*  
16 1383 *Science* **37**, 1565-1578.

18  
19 1384 FEWELL, J. H. (2003). Social insect networks. *Science* **301**, 1867-1870.

21 1385 FISCHER, J., NOSER, R. & HAMMERSCHMIDT, K. (2013). Bioacoustic field research: A primer to  
22 1386 acoustic analyses and playback experiments with primates. *American Journal of Primatology* **75**, 643-  
23 1387 663.

25 1388 FITCH, W. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences* **4**, 258-  
26 1389 267.

28  
29 1390 FITCH, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in  
30 1391 rhesus macaques. *The Journal of the Acoustical Society of America* **102**, 1213-1222.

32 1392 FORD, J. K. (1989). Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island,  
33 1393 British Columbia. *Canadian Journal of Zoology* **67**, 727-745.

35  
36 1394 FORTUNATO, S. (2010). Community detection in graphs. *Physics Reports* **486**, 75-174.

38 1395 FRED, A. L. & JAIN, A. K. (2005). Combining multiple clusterings using evidence accumulation.  
39 1396 *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **27**, 835-850.

41 1397 FREEBERG, T. M., DUNBAR, R. I. M., ORD, T. J., FREEBERG, T. M., DUNBAR, R. I. M. & ORD,  
42 1398 T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity.  
43 1399 *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1785-1801.

45  
46 1400 FREEBERG, T. M. (2012). Geographic variation in note composition and use of chick-a-dee calls of  
47 1401 Carolina chickadees (*Poecile carolinensis*). *Ethology* **118**, 555-565.

49 1402 FREEBERG, T. M. & LUCAS, J. R. (2012). Information theoretical approaches to chick-a-dee calls of  
50 1403 Carolina chickadees (*Poecile carolinensis*). *Journal of Comparative Psychology* **126**, 68-81.

52  
53 1404 FUSHING, H. & MCASSEY, M. P. (2010). Time, temperature, and data cloud geometry. *Physical*  
54 1405 *Review E* **82**, 061110.

56 1406 GALE, W. & SAMPSON, G. (1995). Good-Turing smoothing without tears. *Journal of Quantitative*  
57 1407 *Linguistics* **2**, 217-237.

58  
59  
60

- 1  
2  
3 1408 GAMMON, D. E. & ALTIZER, C. E. (2011). Northern Mockingbirds produce syntactical patterns of  
4 1409 vocal mimicry that reflect taxonomy of imitated species. *Journal of Field Ornithology* **82**, 158-164.
- 6 1410 GARLAND, E. C., LILLEY, M. S., GOLDIZEN, A. W., REKDAHL, M. L., GARRIGUE, C. & NOAD,  
7 1411 M. J. (2012). Improved versions of the Levenshtein distance method for comparing sequence information  
8 1412 in animals' vocalisations: tests using humpback whale song. *Behaviour* **149**, 1413-1441.
- 10  
11 1413 GENTNER, T. Q. & HULSE, S. H. (1998). Perceptual mechanisms for individual vocal recognition in  
12 1414 European starlings, *Sturnus vulgaris*. *Animal Behaviour* **56**, 579-594.
- 14 1415 GENTNER, T. Q., FENN, K. M., MARGOLIASH, D. & NUSBAUM, H. C. (2006). Recursive syntactic  
15 1416 pattern learning by songbirds. *Nature* **440**, 1204-1207.
- 17  
18 1417 GERHARDT, H. C. (2001). Acoustic communication in two groups of closely related treefrogs.  
19 1418 *Advances in the Study of Behavior* **30**, 99-167.
- 20  
21 1419 GERHARDT, H. C. & HUBER, F. (2002). *Acoustic Communication in Insects and Anurans: Common*  
22 1420 *Problems and Diverse Solutions*. University of Chicago Press.
- 23  
24 1421 GHAZANFAR, A. A., FLOMBAUM, J. I., MILLER, C. T. & HAUSER, M. D. (2001). The units of  
25 1422 perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): playback  
26 1423 experiments with long calls. *Journal of Comparative Physiology A* **187**, 27-35.
- 28  
29 1424 GIRGENRATH, M. & MARSH, R. (1997). In vivo performance of trunk muscles in tree frogs during  
30 1425 calling. *Journal of Experimental Biology* **200**, 3101-3108.
- 31  
32 1426 GOLD, E. M. (1967). Language identification in the limit. *Information and control* **10**, 447-474.
- 33  
34 1427 GOLDBOGEN, J. A., SOUTHALL, B. L., DERUITER, S. L., CALAMBOKIDIS, J., FRIEDLAENDER,  
35 1428 A. S., HAZEN, E. L., FALCONE, E. A., SCHORR, G. S., DOUGLAS, A. & MORETTI, D. J. (2013).  
36 1429 Blue whales respond to simulated mid-frequency military sonar. *Proceedings of the Royal Society B:*  
37 1430 *Biological Sciences* **280**, 20130657.
- 38  
39 1431 GOUZOULES, S., GOUZOULES, H. & MARLER, P. (1984). Rhesus monkey (*Macaca mulatta*)  
40 1432 screams: Representational signalling in the recruitment of agonistic aid. *Animal Behaviour* **32**, 182-193.
- 42  
43 1433 GREEN, S. R., MERCADO III, E., PACK, A. A. & HERMAN, L. M. (2011). Recurring patterns in the  
44 1434 songs of humpback whales (*Megaptera novaeangliae*). *Behavioural Processes* **86**, 284-294.
- 46  
47 1435 GRIFFITHS, T. D. & WARREN, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*  
48 1436 **5**, 887-892.
- 49  
50 1437 HAHNLOSER, R. H., KOZHEVNIKOV, A. A. & FEE, M. S. (2002). An ultra-sparse code underlies the  
51 1438 generation of neural sequences in a songbird. *Nature* **419**, 65-70.
- 52  
53 1439 HAILMAN, J. P. (2008). *Coding and Redundancy: Man-Made and Animal-Evolved Signals*. Harvard  
54 1440 University Press.
- 55  
56  
57  
58  
59  
60



1  
2  
3 1441 HAMAKER, J., GANAPATHIRAJU, A. & PICONE, J. (1998). Information theoretic approaches to  
4 1442 model selection. *Proceedings of the International Conference on Spoken Language Processing* **7**, 2931-  
5 1443 2934.  
6  
7  
8 1444 HANDCOCK, M. S., HUNTER, D. R., BUTTS, C. T., GOODREAU, S. M. & MORRIS, M. (2008).  
9 1445 statnet: Software tools for the representation, visualization, analysis and simulation of network data.  
10 1446 *Journal of Statistical Software* **24**, 1548-1557.  
11  
12 1447 HANDEL, S., TODD, S. K. & ZOIDIS, A. M. (2009). Rhythmic structure in humpback whale  
13 1448 (*Megaptera novaeangliae*) songs: Preliminary implications for song production and perception. *The*  
14 1449 *Journal of the Acoustical Society of America* **125**, EL225-EL230.  
15  
16  
17 1450 HARLEY, H. E. (2008). Whistle discrimination and categorization by the Atlantic bottlenose dolphin  
18 1451 (*Tursiops truncatus*): A review of the signature whistle framework and a perceptual test. *Behavioural*  
19 1452 *processes* **77**, 243-268.  
20  
21 1453 HARNAD, S. R. (1990). *Categorical Perception: The Groundwork of Cognition*. Cambridge University  
22 1454 Press.  
23  
24 1455 HARRINGTON, F. H., ASA, C. S., MECH, L. & BOITANI, L. (2003). Wolf communication. In *Wolves:*  
25 1456 *Behavior, ecology, and conservation*, pp. 66-103. University of Chicago Press, Chicago.  
26  
27  
28 1457 HARTSHORNE, C. (1973). *Born to sing. An interpretation and world survey of bird songs*. London:  
29 1458 Indiana University Press.  
30  
31 1459 HARTSHORNE, C. (1956). The monotony-threshold in singing birds. *The Auk* **73**, 176-192.  
32  
33 1460 HASSELQUIST, D., BENSCH, S. & VON SCHANTZ, T. (1996). Correlation between male song  
34 1461 repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature* **381**, 229-232.  
35  
36  
37 1462 HAUSER, M. D., CHOMSKY, N. & FITCH, W. (2002). The faculty of language: What is it, who has it,  
38 1463 and how did it evolve? *Science* **298**, 1569-1579.  
39  
40 1464 HAUSER, M. D., AGNETTA, B. & PEREZ, C. (1998). Orienting asymmetries in rhesus monkeys: the  
41 1465 effect of time-domain changes on acoustic perception. *Animal Behaviour* **56**, 41-47.  
42  
43 1466 HEBETS, E. A. & PAPAJ, D. R. (2005). Complex signal function: developing a framework of testable  
44 1467 hypotheses. *Behavioral Ecology and Sociobiology* **57**, 197-214.  
45  
46  
47 1468 HELBLE, T. A., IERLEY, G. R., GERALD, L., ROCH, M. A. & HILDEBRAND, J. A. (2012). A  
48 1469 generalized power-law detection algorithm for humpback whale vocalizations. *The Journal of the*  
49 1470 *Acoustical Society of America* **131**, 2682-2699.  
50  
51 1471 HENDERSON, E. E., HILDEBRAND, J. A. & SMITH, M. H. (2011). Classification of behavior using  
52 1472 vocalizations of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*). *The Journal of the*  
53 1473 *Acoustical Society of America* **130**, 557-567.  
54  
55  
56  
57  
58  
59  
60

- 1474 HENRY, K. S., GALL, M. D., BIDELMAN, G. M. & LUCAS, J. R. (2011). Songbirds tradeoff auditory  
1475 frequency resolution and temporal resolution. *Journal of Comparative Physiology A: Neuroethology,*  
1476 *Sensory, Neural, and Behavioral Physiology* **197**, 351-359.
- 1477 HENRY, K. S. & LUCAS, J. R. (2010). Habitat-related differences in the frequency selectivity of  
1478 auditory filters in songbirds. *Functional Ecology* **24**, 614-624.
- 1479 HERZING, D. L. (1996). Vocalizations and associated underwater behavior of free-ranging Atlantic  
1480 spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*. *Aquatic Mammals* **22**,  
1481 61-80.
- 1482 HULSE, S. H. (2002). Auditory scene analysis in animal communication. *Advances in the Study of*  
1483 *Behavior* **31**, 163-200.
- 1484 HUMPHRIES, M. D. & GURNEY, K. (2008). Network 'small-world-ness': a quantitative method for  
1485 determining canonical network equivalence. *PLoS One* **3**, e0002051.
- 1486 ILANY, A., BAROCAS, A., KOREN, L., KAM, M. & GEFFEN, E. (2011). Do singing rock hyraxes  
1487 exploit conspecific calls to gain attention? *PloS One* **6**, e28612.
- 1488 ILANY, A., BAROCAS, A., KAM, M., ILANY, T. & GEFFEN, E. (2013). The energy cost of singing in  
1489 wild rock hyrax males: evidence for an index signal. *Animal Behaviour* **85**, 995-1001.
- 1490 IRWIN, D. E. (2000). Song variation in an avian ring species. *Evolution* **54**, 998-1010.
- 1491 ISAAC, D. & MARLER, P. (1963). Ordering of sequences of singing behaviour of mistle thrushes in  
1492 relationship to timing. *Animal Behaviour* **11**, 179-188.
- 1493 JANI, V. M. (1999). Pitfalls in the categorization of behaviour: a comparison of dolphin whistle  
1494 classification methods. *Animal Behaviour* **57**, 133-143.
- 1495 JANI, V. M., SAYIGH, L. & WELLS, R. (2006). Signature whistle shape conveys identity information  
1496 to bottlenose dolphins. *Proceedings of the National Academy of Sciences* **103**, 8293-8297.
- 1497 JANSEN, D. A., CANT, M. A. & MANSER, M. B. (2012). Segmental concatenation of individual  
1498 signatures and con-text cues in banded mongoose (*Mungos mungo*) close calls. *BMC biology* **10**, 97.
- 1499 JI, A., JOHNSON, M. T., WALSH, E. J., MCGEE, J. & ARMSTRONG, D. L. (2013). Discrimination of  
1500 individual tigers (*Panthera tigris*) from long distance roars. *The Journal of the Acoustical Society of*  
1501 *America* **133**, 1762-1769.
- 1502 JIN, D. Z. & KOZHEVNIKOV, A. A. (2011). A compact statistical model of the song syntax in  
1503 Bengalese finch. *PLoS Computational Biology* **7**, e1001108.
- 1504 JIN, D. Z. (2009). Generating variable birdsong syllable sequences with branching chain networks in  
1505 avian premotor nucleus HVC. *Physical Review E* **80**, 051902.
- 1506 JOHNSON, K. (2004). Gold's Theorem and cognitive science. *Philosophy of Science* **71**, 571-592.



1  
2  
3 1507 JONES, A. E., TEN CATE, C. & BIJLEVELD, C. C. (2001). The interobserver reliability of scoring  
4 1508 sonagrams by eye: a study on methods, illustrated on zebra finch songs. *Animal Behaviour* **62**, 791-801.  
5  
6 1509 KAKISHITA, Y., SASAHARA, K., NISHINO, T., TAKAHASI, M. & OKANOYA, K. (2009).  
7 1510 Ethological data mining: an automata-based approach to extract behavioral units and rules. *Data Mining*  
8 1511 *and Knowledge Discovery* **18**, 446-471.  
9  
10 1512 KANWAL, J. S., MATSUMURA, S., OHLEMILLER, K. & SUGA, N. (1994). Analysis of acoustic  
11 1513 elements and syntax in communication sounds emitted by mustached bats. *The Journal of the Acoustical*  
12 1514 *Society of America* **96**, 1229-1254.  
13  
14 1515 KASHTAN, N., ITZKOVITZ, S., MILO, R. & ALON, U. (2002). Mfinder tool guide. *Department of*  
15 1516 *Molecular Cell Biology and Computer Science and Applied Mathematics, Weizmann Institute of Science,*  
16 1517 *Rehovot Israel, Tech.Rep* .  
17  
18 1518 KATAHIRA, K., SUZUKI, K., OKANOYA, K. & OKADA, M. (2011). Complex sequencing rules of  
19 1519 birdsong can be explained by simple hidden Markov processes. *PloS One* **6**, e24516.  
20  
21 1520 KATZ, S. (1987). Estimation of probabilities from sparse data for the language model component of a  
22 1521 speech recognizer. *Acoustics, Speech and Signal Processing, IEEE Transactions on* **35**, 400-401.  
23  
24 1522 KERSHENBAUM, A. (2013). Entropy rate as a measure of animal vocal complexity. *Bioacoustics* .  
25  
26 1523 KERSHENBAUM, A., ILANY, A., BLAUSTEIN, L. & GEFFEN, E. (2012). Syntactic structure and  
27 1524 geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological*  
28 1525 *Sciences* **279**, 2974-2981.  
29  
30 1526 KERSHENBAUM, A. & ROCH, M. A. (2013). An image processing based paradigm for the extraction  
31 1527 of tonal sounds in cetacean communications. *Journal of the Acoustical Society of America* **134**, 4435-  
32 1528 4445.  
33  
34 1529 KERSHENBAUM, A., SAYIGH, L. S. & JANIK, V. M. (2013). The encoding of individual identity in  
35 1530 dolphin signature whistles: how much information is needed? *PLoS One* **8**, e77671.  
36  
37 1531 KIPPER, S., MUNDRY, R., SOMMER, C., HULTSCH, H. & TODT, D. (2006). Song repertoire size is  
38 1532 correlated with body measures and arrival date in common nightingales, *Luscinia megarhynchos*. *Animal*  
39 1533 *Behaviour* **71**, 211-217.  
40  
41 1534 KITCHEN, D. M. (2004). Alpha male black howler monkey responses to loud calls: effect of numeric  
42 1535 odds, male companion behaviour and reproductive investment. *Animal Behaviour* **67**, 125-139.  
43  
44 1536 KLINCK, H., KINDERMANN, L. & BOEBEL, O. (2008). Detection of leopard seal (*Hydrurga leptonyx*)  
45 1537 vocalizations using the Envelope-Spectrogram Technique (TEST) in combination with a Hidden Markov  
46 1538 model. *Canadian Acoustics* **36**, 118-124.  
47  
48 1539 KOGAN, J. A. & MARGOLIASH, D. (1998). Automated recognition of bird song elements from  
49 1540 continuous recordings using dynamic time warping and hidden Markov models: A comparative study.  
50 1541 *The Journal of the Acoustical Society of America* **103**, 2185-2196.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1542 KOLODNY, O., LOTEM, A. & EDELMAN, S. (in press). Learning a generative probabilistic grammar  
1543 of experience: a process-level model of language acquisition.
- 1544 KOREN, L. & GEFFEN, E. (2011). Individual identity is communicated through multiple pathways in  
1545 male rock hyrax (*Procavia capensis*) songs. *Behavioral Ecology and Sociobiology* **65**, 675-684.
- 1546 KOREN, L. & GEFFEN, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): a multi-  
1547 information distributing channel. *Behavioral Ecology and Sociobiology* **63**, 581-590.
- 1548 KRAMS, I., KRAMA, T., FREEBERG, T. M., KULLBERG, C. & LUCAS, J. R. (2012). Linking social  
1549 complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society B:*  
1550 *Biological Sciences* **367**, 1879-1891.
- 1551 KREBS, J. R., ASHCROFT, R. & ORSDOL, K. V. (1981). Song matching in the great tit *Parus major* L.  
1552 *Animal Behaviour* **29**, 918-923.
- 1553 KREBS, J. (1976). Habituation and song repertoires in the great tit. *Behavioral Ecology and Sociobiology*  
1554 **1**, 215-227.
- 1555 KROODSMA, D. E. (1989). Suggested experimental designs for song playbacks. *Animal Behaviour* **37**,  
1556 600-609.
- 1557 KROODSMA, D. E. (1985). Development and use of two song forms by the Eastern Phoebe. *The Wilson*  
1558 *Bulletin* **97**, 21-29.
- 1559 KROODSMA, D. E. (1990). Patterns in songbird singing behaviour: Hartshorne vindicated. *Animal*  
1560 *Behaviour* **39**, 994-996.
- 1561 LACHLAN, R. F., VERZIJDEN, M. N., BERNARD, C. S., JONKER, P., KOESE, B., JAARSMA, S.,  
1562 SPOOR, W., SLATER, P. J. & TEN CATE, C. (2013). The progressive loss of syntactical structure in  
1563 bird song along an island colonization chain. *Current Biology* **19**, 1896-1901.
- 1564 LAMBRECHTS, M. & DHONDT, A. (1987). Differences in singing performance between male great  
1565 tits. *Ardea* **75**, 43-52.
- 1566 LAMBRECHTS, M. & DHONDT, A. A. (1988). The anti-exhaustion hypothesis: a new hypothesis to  
1567 explain song performance and song switching in the great tit. *Animal Behaviour* **36**, 327-334.
- 1568 LARSON, K. A. (2004). Advertisement call complexity in northern leopard frogs, *Rana pipiens*. *Journal*  
1569 *Information* **2004**.
- 1570 LEE, K. & HON, H. (1989). Speaker-independent phone recognition using hidden Markov models.  
1571 *Acoustics, Speech and Signal Processing, IEEE Transactions on* **37**, 1641-1648.
- 1572 LEMASSON, A., OUATTARA, K., BOUCHET, H. & ZUBERBÜHLER, K. (2010). Speed of call  
1573 delivery is related to context and caller identity in Campbell's monkey males. *Naturwissenschaften* **97**,  
1574 1023-1027.

1  
2  
3 1575 LIPKIND, D., MARCUS, G. F., BEMIS, D. K., SASAHARA, K., JACOBY, N., TAKAHASI, M.,  
4 1576 SUZUKI, K., FEHER, O., RAVBAR, P. & OKANOYA, K. (2013). Stepwise acquisition of vocal  
5 1577 combinatorial capacity in songbirds and human infants. *Nature* **498**, 104-108.

7  
8 1578 LYNCH, K. S., STANELY RAND, A., RYAN, M. J. & WILCZYNSKI, W. (2005). Plasticity in female  
9 1579 mate choice associated with changing reproductive states. *Animal Behaviour* **69**, 689-699.

11 1580 MACEDONIA, J. M. (1990). What is communicated in the antipredator calls of lemurs: evidence from  
12 1581 playback experiments with ringtailed and ruffed lemurs. *Ethology* **86**, 177-190.

14 1582 MAHURIN, E. J. & FREEBERG, T. M. (2009). Chick-a-dee call variation in Carolina chickadees and  
15 1583 recruiting flockmates to food. *Behavioral Ecology* **20**, 111-116.

17  
18 1584 MAKEIG, S., DEBENER, S., ONTON, J. & DELORME, A. (2004). Mining event-related brain  
19 1585 dynamics. *Trends in Cognitive Sciences* **8**, 204-210.

21 1586 MALLAT, S. (1999). *A Wavelet Tour of Signal Processing*. Access Online via Elsevier.

23 1587 MANNING, C. D., RAGHAVAN, P. & SCHÜTZE, H. (2008). *Introduction to information retrieval*.  
24 1588 Cambridge University Press, Cambridge.

26  
27 1589 MANSER, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the  
28 1590 level of response urgency. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**,  
29 1591 2315-2324.

31 1592 MARLER, P. (1977). The structure of animal communication sounds. In *Recognition of Complex*  
32 1593 *Acoustic Signals* (ed. T. H. Bullock ), pp. 17-35. Springer Verlag, Berlin.

34  
35 1594 MARLER, P. & PICKERT, R. (1984). Species-universal microstructure in the learned song of the swamp  
36 1595 sparrow (*Melospiza georgiana*). *Animal Behaviour* **32**, 673-689.

38 1596 MARLER, P. & SHERMAN, V. (1985). Innate differences in singing behaviour of sparrows reared in  
39 1597 isolation from adult conspecific song. *Animal Behaviour* **33**, 57-71.

41 1598 MARTIN, W. F. & GANS, C. (1972). Muscular control of the vocal tract during release signaling in the  
42 1599 toad *Bufo valliceps*. *Journal of Morphology* **137**, 1-27.

44  
45 1600 MARTIN, W. P. (1972). Evolution of vocalization in the genus *Bufo*. In *Evolution in the Genus Bufo* (ed.  
46 1601 W. F. Blair ), pp. 37-70. University of Texas Press, Austin, TX.

48 1602 MARTIN, W. & FLANDRIN, P. (1985). Wigner-Ville spectral analysis of nonstationary processes.  
49 1603 *Acoustics, Speech and Signal Processing, IEEE Transactions on* **33**, 1461-1470.

51 1604 MARTINICH, A. & SOSA, D. (2013). *The Philosophy of Language*. Oxford University Press, Oxford.

53  
54 1605 MATTHEWS, J., RENDELL, L., GORDON, J. & MACDONALD, D. (1999). A review of frequency and  
55 1606 time parameters of cetacean tonal calls. *Bioacoustics* **10**, 47-71.

56  
57  
58  
59  
60

- 1607 MCCOWAN, B., DOYLE, L. R. & HANSER, S. F. (2002). Using information theory to assess the  
1608 diversity, complexity, and development of communicative repertoires. *Journal of Comparative*  
1609 *Psychology* **116**, 166-172.
- 1610 MCCOWAN, B., HANSER, S. F. & DOYLE, L. R. (1999). Quantitative tools for comparing animal  
1611 communication systems: information theory applied to bottlenose dolphin whistle repertoires. *Animal*  
1612 *Behaviour* **57**, 409-419.
- 1613 MENNILL, D. J. & RATCLIFFE, L. M. (2004). Overlapping and matching in the song contests of black-  
1614 capped chickadees. *Animal Behaviour* **67**, 441-450.
- 1615 MENNILL, D. J., RATCLIFFE, L. M. & BOAG, P. T. (2002). Female eavesdropping on male song  
1616 contests in songbirds. *Science* **296**, 873-873.
- 1617 MERCADO, E. I., HERMAN, L. M. & PACK, A. A. (2003). Stereotypical sound patterns in humpback  
1618 whale songs: Usage and function. *Aquatic Mammals* **29**, 37-52.
- 1619 MERCADO, E. I., SCHNEIDER, J. N., PACK, A. A. & HERMAN, L. M. (2010). Sound production by  
1620 singing humpback whales. *The Journal of the Acoustical Society of America* **127**, 2678-2691.
- 1621 MILLER, C. T. & BEE, M. A. (2012). Receiver psychology turns 20: is it time for a broader approach?  
1622 *Animal Behaviour* **83**, 331-343.
- 1623 MILLER, C. T. & COHEN, Y. E. (2010). Vocalizations as auditory objects: behavior and  
1624 neurophysiology. In *Primate Neuroethology* (eds. M. L. Platt and A. A. Ghazanfar), pp. 237-255. Oxford  
1625 University Press, Oxford.
- 1626 MILO, R., SHEN-ORR, S., ITZKOVITZ, S., KASHTAN, N., CHKLOVSKII, D. & ALON, U. (2002).  
1627 Network motifs: simple building blocks of complex networks. *Science* **298**, 824-827.
- 1628 MITANI, J. C., HUNLEY, K. & MURDOCH, M. (1999). Geographic variation in the calls of wild  
1629 chimpanzees: a reassessment. *American Journal of Primatology* **47**, 133-151.
- 1630 MOORE, B. C. & MOORE, B. C. (2003). *An Introduction to the Psychology of Hearing*. Academic  
1631 Press, San Diego.
- 1632 MULLIGAN, J. A. (1966). *Singing Behavior and its Development in the Song Sparrow Melospiza*  
1633 *melodia*. University of California Press.
- 1634 MUMFORD, D. & DESOLNEUX, A. (2010). *Pattern Theory: The Stochastic Analysis of Real-World*  
1635 *Signals*. A. K. Peters Ltd.
- 1636 MUNOZ, N. E. & BLUMSTEIN, D. T. (2012). Multisensory perception in uncertain environments.  
1637 *Behavioral Ecology* **23**, 457-462.
- 1638 MURRAY, S. O., MERCADO, E. I. & ROITBLAT, H. L. (1998). Characterizing the graded structure of  
1639 false killer whale (*Pseudorca crassidens*) vocalizations. *The Journal of the Acoustical Society of America*  
1640 **104**, 1679-1688.

1  
2  
3 1641 NARINS, P. M., REICHMAN, O., JARVIS, J. U. & LEWIS, E. R. (1992). Seismic signal transmission  
4 1642 between burrows of the Cape mole-rat, *Georychus capensis*. *Journal of Comparative Physiology A* **170**,  
5 1643 13-21.  
6  
7  
8 1644 NARINS, P.M., LEWIS, E.R. & MCCLELLAND, B.E. (2000). Hyperextended call note repertoire of the  
9 1645 endemic Madagascar treefrog *Boophis madagascariensis* (Rhacophoridae). *Journal of Zoology* **250**, 283-  
10 1646 298.  
11  
12 1647 NEEDLEMAN, S. B. & WUNSCH, C. D. (1970). A general method applicable to the search for  
13 1648 similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**, 443-453.  
14  
15  
16 1649 NELSON, D. A. & MARLER, P. (1989). Categorical perception of a natural stimulus continuum:  
17 1650 birdsong. *Science* **244**, 976-978.  
18  
19 1651 NELSON, D. A. & POESEL, A. (2007). Segregation of information in a complex acoustic signal:  
20 1652 individual and dialect identity in white-crowned sparrow song. *Animal Behaviour* **74**, 1073-1084.  
21  
22 1653 NETTLE, D. (1999). Language variation and the evolution of societies. In *The Evolution of Culture: An*  
23 1654 *Interdisciplinary View* (eds. R. I. M. Dunbar, C. Knight and C. Power), pp. 214-227. Rutgers University  
24 1655 Press.  
25  
26  
27 1656 NEVO, E., HETH, G., BEILES, A. & FRANKENBERG, E. (1987). Geographic dialects in blind mole  
28 1657 rats: role of vocal communication in active speciation. *Proceedings of the National Academy of Sciences*  
29 1658 **84**, 3312-3315.  
30  
31 1659 NEWMAN, M. (2009). *Networks: An Introduction*. Oxford University Press, Oxford.  
32  
33 1660 NEWMAN, M. E. (2003). The structure and function of complex networks. *SIAM Review* **45**, 167-256.  
34  
35  
36 1661 NITYANANDA, V. & BEE, M. A. (2011). Finding your mate at a cocktail party: frequency separation  
37 1662 promotes auditory stream segregation of concurrent voices in multi-species frog choruses. *PloS One* **6**,  
38 1663 e21191.  
39  
40 1664 NOTMAN, H. & RENDALL, D. (2005). Contextual variation in chimpanzee pant hoots and its  
41 1665 implications for referential communication. *Animal Behaviour* **70**, 177-190.  
42  
43  
44 1666 NOWICKI, S. & MARLER, P. (1988). How do birds sing? *Music Perception* , 391-426.  
45  
46 1667 NOWICKI, S. & NELSON, D. A. (1990). Defining natural categories in acoustic signals: comparison of  
47 1668 three methods applied to ‘chick-a-dee’ call notes. *Ethology* **86**, 89-101.  
48  
49 1669 OHMS, V. R., ESCUDERO, P., LAMMERS, K. & TEN CATE, C. (2012). Zebra finches and Dutch  
50 1670 adults exhibit the same cue weighting bias in vowel perception. *Animal Cognition* **15**, 155-161.  
51  
52  
53 1671 OLESON, E. M., WIGGINS, S. M. & HILDEBRAND, J. A. (2007). Temporal separation of blue whale  
54 1672 call types on a southern California feeding ground. *Animal Behaviour* **74**, 881-894.  
55  
56 1673 OPPENHEIM, A. V. & SCHAFER, R. W. (2004). From frequency to quefrequency: A history of the  
57 1674 cepstrum. *Signal Processing Magazine, IEEE* **21**, 95-106.  
58  
59  
60



- 1  
2  
3 1675 OPSAHL, T., AGNEESSENS, F. & SKVORETZ, J. (2010). Node centrality in weighted networks:  
4 1676 Generalizing degree and shortest paths. *Social Networks* **32**, 245-251.
- 5  
6 1677 OSWALD, J. N., RANKIN, S., BARLOW, J. & LAMMERS, M. O. (2007). A tool for real-time acoustic  
7 1678 species identification of delphinid whistles. *The Journal of the Acoustical Society of America* **122**, 587-  
8 1679 595.
- 9  
10  
11 1680 OTTER, K., MCGREGOR, P. K., TERRY, A. M., BURFORD, F. R., PEAKE, T. M. & DABELSTEEN,  
12 1681 T. (1999). Do female great tits (*Parus major*) assess males by eavesdropping? A field study using  
13 1682 interactive song playback. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **266**,  
14 1683 1305-1309.
- 15  
16  
17 1684 OWREN, M. J., RENDALL, D. & RYAN, M. J. (2010). Redefining animal signaling: influence versus  
18 1685 information in communication. *Biology & Philosophy* **25**, 755-780.
- 19  
20 1686 PARTAN, S. & MARLER, P. (1999). Communication goes multimodal. *Science* **283**, 1272-1273.
- 21  
22 1687 PAULUS, J., MÜLLER, M. & KLAPURI, A. (2010). State of the Art Report: Audio-Based Music  
23 1688 Structure Analysis. *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*  
24 1689 , 625-636.
- 25  
26  
27 1690 PAYNE, K. B., LANGBAUER JR, W. R. & THOMAS, E. M. (1986). Infrasonic calls of the Asian  
28 1691 elephant (*Elephas maximus*). *Behavioral Ecology and Sociobiology* **18**, 297-301.
- 29  
30 1692 PAYNE, R. S. & MCVAY, S. (1971). Songs of humpback whales. *Science* **173**, 585-597.
- 31  
32 1693 PETERSON, G. E. & BARNEY, H. L. (1952). Control methods used in a study of the vowels. *The*  
33 1694 *Journal of the Acoustical Society of America* **24**, 175-184.
- 34  
35  
36 1695 PICONE, J. W. (1993). Signal modeling techniques in speech recognition. *Proceedings of the IEEE* **81**,  
37 1696 1215-1247.
- 38  
39 1697 PODOS, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird family  
40 1698 (Passeriformes: Emberizidae). *Evolution* , 537-551.
- 41  
42 1699 PODOS, J., PETERS, S., RUDNICKY, T., MARLER, P. & NOWICKI, S. (1992). The organization of  
43 1700 song repertoires in song sparrows: themes and variations. *Ethology* **90**, 89-106.
- 44  
45  
46 1701 POLLARD, K. A. & BLUMSTEIN, D. T. (2012). Evolving communicative complexity: insights from  
47 1702 rodents and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1869-  
48 1703 1878.
- 49  
50 1704 POOLE, J. H. (1999). Signals and assessment in African elephants: evidence from playback experiments.  
51 1705 *Animal Behaviour* **58**, 185-193.
- 52  
53  
54 1706 QUICK, N. J. & JANIK, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at  
55 1707 sea. *Proceedings of the Royal Society B: Biological Sciences* **279**, 2539-2545.
- 56  
57  
58  
59  
60

1  
2  
3 1708 R DEVELOPMENT TEAM (2012) *R: A Language and Environment for Statistical Computing*. R  
4 1709 Foundation for Statistical Computing, Vienna, Austria, 2007 .  
5  
6  
7 1710 RABINER, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech  
8 1711 recognition. *Proceedings of the IEEE* **77**, 257-286.  
9  
10 1712 RAEMAEEKERS, J. J., RAEMAEEKERS, P. M. & HAIMOFF, E. H. (1984). Loud calls of the gibbon  
11 1713 (*Hylobates lar*): repertoire, organisation and context. *Behaviour* **91**, 146-189.  
12  
13 1714 RAGGE, D. & REYNOLDS, W. (1988). The songs and taxonomy of the grasshoppers of the *Chorthippus*  
14 1715 *biguttulus* group in the Iberian Peninsula (Orthoptera: Acrididae). *Journal of Natural History* **22**, 897-  
15 1716 929.  
16  
17  
18 1717 RANDALL, J. A. (1997). Species-specific footdrumming in kangaroo rats: *Dipodomys ingens*, *D. deserti*,  
19 1718 *D. spectabilis*. *Animal Behaviour* **54**, 1167-1175.  
20  
21 1719 RANDALL, J. A. (1989). Individual footdrumming signatures in banner-tailed kangaroo rats *Dipodomys*  
22 1720 *spectabilis*. *Animal Behaviour* **38**, 620-630.  
23  
24 1721 RECANZONE, G. H. & SUTTER, M. L. (2008). The biological basis of audition. *Annu.Rev.Psychol.* **59**,  
25 1722 119-142.  
26  
27  
28 1723 REMEZ, R. E., RUBIN, P. E., BERNS, S. M., PARDO, J. S. & LANG, J. M. (1994). On the perceptual  
29 1724 organization of speech. *Psychological Review* **101**, 129.  
30  
31 1725 RICHARDS, D. G. (1981). Alerting and message components in songs of rufous-sided towhees.  
32 1726 *Behaviour* **76**, 223-249.  
33  
34  
35 1727 RIEBEL, K. & SLATER, P. (2003). Temporal variation in male chaffinch song depends on the singer and  
36 1728 the song type. *Behaviour* **140**, 269-288.  
37  
38 1729 RIPLEY, B. D. (2007). *Pattern recognition and neural networks*. Cambridge University Press.  
39  
40 1730 ROBINSON, J. G. (1979). An analysis of the organization of vocal communication in the titi monkey  
41 1731 *Callicebus moloch*. *Zeitschrift für Tierpsychologie* **49**, 381-405.  
42  
43 1732 ROOT-GUTTERIDGE, H., BENCSIK, M., CHEBLI, M., GENTLE, L. K., TERRELL-NIELD, C.,  
44 1733 BOURIT, A. & YARNELL, R. W. (2014). Identifying individual wild Eastern grey wolves (*Canis lupus*  
45 1734 *lycaon*) using fundamental frequency and amplitude of howls. *Bioacoustics* **23**, 55-66.  
46  
47  
48 1735 ROTHENBERG, D., ROESKE, T. C., VOSS, H. U., NAGUIB, M. & TCHERNICHOVSKI, O. (2013).  
49 1736 Investigation of musicality in birdsong. *Hearing Research* **308**, 71-83.  
50  
51 1737 RUXTON, G. D. & SCHAEFER, H. M. (2011). Resolving current disagreements and ambiguities in the  
52 1738 terminology of animal communication. *Journal of Evolutionary Biology* **24**, 2574-2585.  
53  
54  
55 1739 SAAR, S. & MITRA, P. P. (2008). A technique for characterizing the development of rhythms in bird  
56 1740 song. *PLoS One* **3**, e1461.  
57  
58  
59  
60



- 1741 SALMI, R., HAMMERSCHMIDT, K. & DORAN-SHEEHY, D. M. (2013). Western Gorilla Vocal  
1742 Repertoire and Contextual Use of Vocalizations. *Ethology* **119**, 831-847.
- 1743 SASAHARA, K., CODY, M. L., COHEN, D. & TAYLOR, C. E. (2012). Structural design principles of  
1744 complex bird songs: a network-based approach. *PloS One* **7**, e44436.
- 1745 SAUNDERS, A. A. (1951). The song of the Song Sparrow. *The Wilson Bulletin*, 99-109.
- 1746 SAYIGH, L., QUICK, N., HASTIE, G. & TYACK, P. (2012). Repeated call types in short-finned pilot  
1747 whales, *Globicephala macrorhynchus*. *Marine Mammal Science* **29**, 312-324.
- 1748 SAYIGH, L. S., ESCH, H. C., WELLS, R. S. & JANIK, V. M. (2007). Facts about signature whistles of  
1749 bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour* **74**, 1631-1642.
- 1750 SAYIGH, L. S., TYACK, P. L., WELLS, R. S., SOLOW, A. R., SCOTT, M. D. & IRVINE, A. B.  
1751 (1999). Individual recognition in wild bottlenose dolphins: a field test using playback experiments.  
1752 *Animal Behaviour* **57**, 41-50.
- 1753 SCHEL, A. M., CANDIOTTI, A. & ZUBERBÜHLER, K. (2010). Predator-detering alarm call  
1754 sequences in Guereza colobus monkeys are meaningful to conspecifics. *Animal Behaviour* **80**, 799-808.
- 1755 SCHEL, A. M., TRANQUILLI, S. & ZUBERBÜHLER, K. (2009). The alarm call system of two species  
1756 of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*). *Journal of Comparative*  
1757 *Psychology* **123**, 136-150.
- 1758 SCHREIBER, F. & SCHWÖBBERMEYER, H. (2005). MAVisto: a tool for the exploration of network  
1759 motifs. *Bioinformatics* **21**, 3572-3574.
- 1760 SCHULZ, T. M., WHITEHEAD, H., GERO, S. & RENDELL, L. (2008). Overlapping and matching of  
1761 codas in vocal interactions between sperm whales: insights into communication function. *Animal*  
1762 *Behaviour* **76**, 1977-1988.
- 1763 SCOTT, J. & CARRINGTON, P. J. (2011). *The SAGE Handbook of Social Network Analysis*. SAGE  
1764 Publications.
- 1765 SEARCY, W. A. & BEECHER, M. D. (2011). Continued scepticism that song overlapping is a signal.  
1766 *Animal Behaviour* **81**, e1-e4.
- 1767 SEARCY, W. A., NOWICKI, S. & PETERS, S. (1999). Song types as fundamental units in vocal  
1768 repertoires. *Animal Behaviour* **58**, 37-44.
- 1769 SEARCY, W. A., PODOS, J., PETERS, S. & NOWICKI, S. (1995). Discrimination of song types and  
1770 variants in song sparrows. *Animal Behaviour* **49**, 1219-1226.
- 1771 SEYFARTH, R. M. & CHENEY, D. L. (2010). Production, usage, and comprehension in animal  
1772 vocalizations. *Brain and Language* **115**, 92-100.
- 1773 SEYFARTH, R. M., CHENEY, D. L. & BERGMAN, T. J. (2005). Primate social cognition and the  
1774 origins of language. *Trends in Cognitive Sciences* **9**, 264-266.

1  
2  
3 1775 SEYFARTH, R. M., CHENEY, D. L., BERGMAN, T., FISCHER, J., ZUBERBÜHLER, K. &  
4 1776 HAMMERSCHMIDT, K. (2010). The central importance of information in studies of animal  
5 1777 communication. *Animal Behaviour* **80**, 3-8.  
6  
7  
8 1778 SHANNON, C. E., WEAVER, W., BLAHUT, R. E. & HAJEK, B. (1949). *The Mathematical Theory of*  
9 1779 *Communication*. University of Illinois Press, Urbana.  
10  
11 1780 SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N.,  
12 1781 SCHWIKOWSKI, B. & IDEKER, T. (2003). Cytoscape: a software environment for integrated models of  
13 1782 biomolecular interaction networks. *Genome Research* **13**, 2498-2504.  
14  
15  
16 1783 SHAPIRO, A. D., TYACK, P. L. & SENEFF, S. (2010). Comparing call-based versus subunit-based  
17 1784 methods for categorizing Norwegian killer whale, *Orcinus orca*, vocalizations. *Animal Behaviour* **81**,  
18 1785 377-386.  
19  
20 1786 SLOCOMBE, K. E. & ZUBERBÜHLER, K. (2006). Food-associated calls in chimpanzees: responses to  
21 1787 food types or food preferences? *Animal Behaviour* **72**, 989-999.  
22  
23  
24 1788 SMITH, J. N., GOLDIZEN, A. W., DUNLOP, R. A. & NOAD, M. J. (2008). Songs of male humpback  
25 1789 whales, *Megaptera novaeangliae*, are involved in intersexual interactions. *Animal Behaviour* **76**, 467-477.  
26  
27 1790 SNIJDERS, T. A. (2002). Markov chain Monte Carlo estimation of exponential random graph models.  
28 1791 *Journal of Social Structure* **3**, 1-40.  
29  
30 1792 SOLAN, Z., HORN, D., RUPPIN, E. & EDELMAN, S. (2005). Unsupervised learning of natural  
31 1793 languages. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11629-  
32 1794 11634.  
33  
34  
35 1795 STEGMANN, U. (2013). *Animal Communication Theory: Information and Influence*. Cambridge  
36 1796 University Press.  
37  
38 1797 STODDARD, P. (1996). Vocal recognition of neighbors by territorial passerines. In *Ecology and*  
39 1798 *evolution of acoustic communication in birds* (eds. D. E. Kroodsma and E. L. Miller), pp. 356-374.  
40 1799 Cornell University Press, Ithaca, New York.  
41  
42 1800 STOLCKE, A. (2002). SRILM-an extensible language modeling toolkit. *International Conference on*  
43 1801 *Spoken Language Processing* **2**, 901-904.  
44  
45  
46 1802 STURDY, C. B., PHILLMORE, L. S. & WEISMAN, R. G. (2000). Call-note discriminations in black-  
47 1803 capped chickadees (*Poecile atricapillus*). *Journal of Comparative Psychology* **114**, 357.  
48  
49 1804 SUTHERS, R. A. (2004). How birds sing and why it matters. In *Nature's Music: The Science of Birdsong*  
50 1805 (eds. P. R. Marler and H. Slabbekoorn), pp. 272-295. Elsevier Academic Press, San Diego.  
51  
52 1806 SUTHERS, R. A. (1997). Peripheral control and lateralization of birdsong. *Journal of Neurobiology* **33**,  
53 1807 632-652.  
54  
55  
56 1808 TAYLOR, A. M., REBY, D. & MCCOMB, K. (2008). Human listeners attend to size information in  
57 1809 domestic dog growls. *The Journal of the Acoustical Society of America* **123**, 2903-2909.  
58  
59  
60

- 1810 TCHERNICHOVSKI, O., NOTTEBOHM, F., HO, C. E., PESARAN, B. & MITRA, P. P. (2000). A  
1811 procedure for an automated measurement of song similarity. *Animal Behaviour* **59**, 1167-1176.
- 1812 TEN CATE, C., LACHLAN, R. & ZUIDEMA, W. (2013). Analyzing the Structure of Bird Vocalizations  
1813 and Language: Finding Common Ground. In *Birdsong, Speech, and Language: Exploring the Evolution*  
1814 *of Mind and Brain* (eds. J. J. Bolhuis and M. Everaert), pp. 243-260. MIT Press.
- 1815 THOMPSON, A. B. & HARE, J. F. (2010). Neighbourhood watch: multiple alarm callers communicate  
1816 directional predator movement in Richardson's ground squirrels, *Spermophilus richardsonii*. *Animal*  
1817 *Behaviour* **80**, 269-275.
- 1818 THOMSON, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* **70**,  
1819 1055-1096.
- 1820 TITZE, I. R. (1994). *Principles of Voice Production*. Prentice Hall Englewood Cliffs.
- 1821 TRAWICKI, M. B., JOHNSON, M. & OSIEJUK, T. (2005). Automatic song-type classification and  
1822 speaker identification of Norwegian Ortolan Bunting (*Emberiza hortulana*) vocalizations. *IEEE*  
1823 *Workshop on Machine Learning for Signal Processing*, 277-282.
- 1824 TU, H. & DOOLING, R. J. (2012). Perception of warble song in budgerigars (*Melopsittacus undulatus*):  
1825 evidence for special processing. *Animal Cognition* **15**, 1151-1159.
- 1826 TU, H., SMITH, E. W. & DOOLING, R. J. (2011). Acoustic and perceptual categories of vocal elements  
1827 in the warble song of budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology* **125**,  
1828 420-430.
- 1829 TYACK, P. (1983). Differential response of humpback whales, *Megaptera novaeangliae*, to playback of  
1830 song or social sounds. *Behavioral Ecology and Sociobiology* **13**, 49-55.
- 1831 VAN HEIJNINGEN, C. A., DE VISSER, J., ZUIDEMA, W. & TEN CATE, C. (2009). Simple rules can  
1832 explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the*  
1833 *National Academy of Sciences* **106**, 20538-20543.
- 1834 VAN SCHAIK, C. P., DAMERIUS, L. & ISLER, K. (2013). Wild orangutan males plan and  
1835 communicate their travel direction one day in advance. *PloS One* **8**, e74896.
- 1836 WALLER, S. (2012). Science of the Monkey Mind: Primate Penchants and Human Pursuits. In  
1837 *Experiencing Animal Minds: An Anthology of Animal-human Encounters* (eds. J. A. Smith and R. W.  
1838 Mitchell).
- 1839 WATKINS, W. A. & SCHEVILL, W. E. (1977). Sperm whale codas. *The Journal of the Acoustical*  
1840 *Society of America* **62**, 1485-1490.
- 1841 WATTS, D. J. & STROGATZ, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* **393**,  
1842 440-442.
- 1843 WEARY, D. M., LAMBRECHTS, M. & KREBS, J. (1991). Does singing exhaust male great tits? *Animal*  
1844 *Behaviour* **41**, 540-542.

1  
2  
3 1845 WEARY, D. M. & LEMON, R. E. (1990). Kroodsma refuted. *Animal Behaviour* **39**, 996-998.  
4  
5 1846 WEARY, D. M. & LEMON, R. E. (1988). Evidence against the continuity-versatility relationship in bird  
6 1847 song. *Animal Behaviour* **36**, 1379-1383.  
7  
8  
9 1848 WEARY, D., KREBS, J., EDDYSHAW, R., MCGREGOR, P. & HORN, A. (1988). Decline in song  
10 1849 output by great tits: Exhaustion or motivation? *Animal Behaviour* **36**, 1242-1244.  
11  
12 1850 WEISS, D. J. & HAUSER, M. D. (2002). Perception of harmonics in the combination long call of  
13 1851 cottontop tamarins, *Saguinus oedipus*. *Animal Behaviour* **64**, 415-426.  
14  
15  
16 1852 WEISS, R. J. & BELLO, J. P. (2011). Unsupervised discovery of temporal structure in music. *Selected*  
17 1853 *Topics in Signal Processing, IEEE Journal of* **5**, 1240-1251.  
18  
19 1854 WERNICKE, S. & RASCHE, F. (2006). FANMOD: a tool for fast network motif detection.  
20 1855 *Bioinformatics* **22**, 1152-1153.  
21  
22 1856 WHEELER, B. C. & HAMMERSCHMIDT, K. (2012). Proximate factors underpinning receiver  
23 1857 responses to deceptive false alarm calls in wild tufted capuchin monkeys: Is it counterdeception?  
24 1858 *American Journal of Primatology* **75**, 715-725.  
25  
26  
27 1859 WHEELER, B. C. (2010a). Decrease in alarm call response among tufted capuchins in competitive  
28 1860 feeding contexts: possible evidence for counterdeception. *International Journal of Primatology* **31**, 665-  
29 1861 675.  
30  
31 1862 WHEELER, B. C. (2010b). Production and perception of situationally variable alarm calls in wild tufted  
32 1863 capuchin monkeys (*Cebus apella nigratus*). *Behavioral Ecology and Sociobiology* **64**, 989-1000.  
33  
34  
35 1864 WIENS, J. A. (1982). Song pattern variation in the sage sparrow (*Amphispiza belli*): dialects or  
36 1865 epiphenomena? *The Auk* **99**, 208-229.  
37  
38 1866 WILEY, R. H. (1983). The evolution of communication: information and manipulation. *Animal*  
39 1867 *Behaviour* **2**, 156-189.  
40  
41 1868 WOLPERT, D. H. & MACREADY, W. G. (1997). No free lunch theorems for optimization.  
42 1869 *Evolutionary Computation, IEEE Transactions on* **1**, 67-82.  
43  
44  
45 1870 WYTENBACH, R. A., MAY, M. L. & HOY, R. R. (1996). Categorical perception of sound frequency  
46 1871 by crickets. *Science* **273**, 1542-1544.  
47  
48 1872 YOUNG, S. J. & YOUNG, S. (1994). The htk hidden Markov model toolkit: Design and philosophy.  
49 1873 *Entropic Cambridge Research Laboratory, Ltd* **2**, 2-44.  
50  
51 1874 ZHONG, S. & GHOSH, J. (2005). Generative model-based document clustering: a comparative study.  
52 1875 *Knowledge and Information Systems* **8**, 374-384.  
53  
54  
55 1876 ZIPF, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley press, Oxford.  
56  
57  
58  
59  
60

1877 ZUCCHINI, W. & MACDONALD, I. L. (2009). *Hidden Markov models for time series: an introduction*  
1878 *using R*. CRC Press.

1879

1880

For Review Only

FIGURE LEGENDS

Figure 1. Overall flowchart showing a typical analysis of animal acoustic sequences. In this review, we discuss Identifying units, Characterising sequences, and Identifying meaning.

Figure 2. Different ways that units can be combined to encode information in a sequence.

Figure 3. Examples of the different criteria for dividing a spectrogram into units. (a) Separating units by silent gaps is probably the most commonly used criterion. (b) An acoustic signal may changes its properties at a certain time, without the presence of a silent “gap”, for instance becoming harmonic or noisy. (c) A series of similar sounds may be grouped together as a single unit, regardless of silent gaps between them; a chirp sequence is labelled as **C**. (d) A complex hierarchical structure to the sequence, combining sounds that might otherwise be considered fundamental units.

Figure 4. Example of cepstral processing of a grey wolf *Canis lupis* howl and crickets chirping. Recording was sampled at  $F_s = 16$  kHz, 8 bit quantization. (a) Standard spectrogram analyzed with a 15 ms Blackman-Harris window. (b) Plot of transform to cepstral domain. Lower quefrequencies are related to vocal tract information.  $F_0$  can be determined from the "cepstral bump" apparent between quefrequencies 25-45 and can be derived by  $F_s/\text{quefrequency}$ . (c) Cepstrum (inset) of the frame indicated by an arrow (2.5 s) along with reconstructions of the spectrum created from truncated cepstral sequences. Fidelity improves as the number of cepstra are increased.

Figure 5. Perceptual constraints for the definition of sequence units. (a) Perceptual binding, where two discrete acoustic elements may be perceived by the receiver either as a single element, or as two separate ones. (b) Categorical perception, where continuous variation in acoustic signals may be interpreted by the



receiver as discrete categories. (c) Spectrotemporal constraints, where if the receiver cannot distinguish small differences in time or frequency, discrete elements may be interpreted as joined.

Figure 6. State transition diagram equivalent to a 2<sup>nd</sup> order Markov model and trigram model (N=3) for a sequence containing *A*'s and *B*'s.

Figure 7. State transition diagram of a two state (*X*, *Y*) hidden Markov model capable of producing sequences of acoustic units *A* and *B*. When in state *X*, acoustic units emission of signals *A* and *B* are equally likely  $Pr_e(A|X) = Pr_e(B|X) = 0.5$ , and when in state *Y*, acoustic unit *A* is much more likely  $Pr_e(A|Y) = 0.9$  than *B*  $Pr_e(B|Y) = 0.1$ . Transitioning from state *X* to state *Y* occurs with probability  $Pr_t(X \rightarrow Y) = 0.6$ , whereas from state *Y* to state *X* with probability  $Pr_t(Y \rightarrow X) = 0.3$ .

Figure 8. Simple networks constructed from the sequence of acoustic units *ABC*. The undirected binary network (left) simply indicates that *A*, *B*, and *C* are associated with one another without any information about transition direction. The directed binary network (centre) adds ordering information, for example that *C* cannot follow *A*. The weighted directed network (right) show the probabilities of the transitions between units based on a bigram model.

Figure 9. Grammar (rewrite rules) for approximating the sequence of acoustic units produced by Eastern Pacific blue whales *Balaenoptera musculus*. There are three acoustic units, *A*, *B*, and *D* (Oleson, Wiggins & Hildebrand, 2007), and the sequence begins with a start symbol *S*. Individual *B* or *D* calls may be produced, or song, which consists of repeated sequences of an *A* call followed by one or more *B* calls. The symbol | indicates a choice, and ε, the empty string, indicates that the rule is no longer used. A derivation is shown for the song *ABBAB*. Grammar produced with contributions from Ana Širović (Scripps Institution of Oceanography).



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1932 Figure 10. The classes of formal grammars known as the Chomsky hierarchy (Chomsky, 2002). Each  
1933 class is a generalisation of the class it encloses, and is more complex than the enclosed classes. Image  
1934 publicly available under the Creative Commons Attribution-Share Alike 3.0 Unported license.  
1935 [https://commons.wikimedia.org/wiki/File:Wiki\\_inf\\_chomskeho\\_hierarchia.jpg](https://commons.wikimedia.org/wiki/File:Wiki_inf_chomskeho_hierarchia.jpg)

1937 Figure 11. Network diagram for a sample of California thrasher song data. Nodes indicate distinct phrase  
1938 types (acoustic units), arrows indicate transitions, with the colour of the arrow indicating the strength of  
1939 probability of transition between two units (hotter colours represent more likely transitions).

For Review Only

Table 1. Examples of different approaches to unit definition, from different taxa and with different research aims.

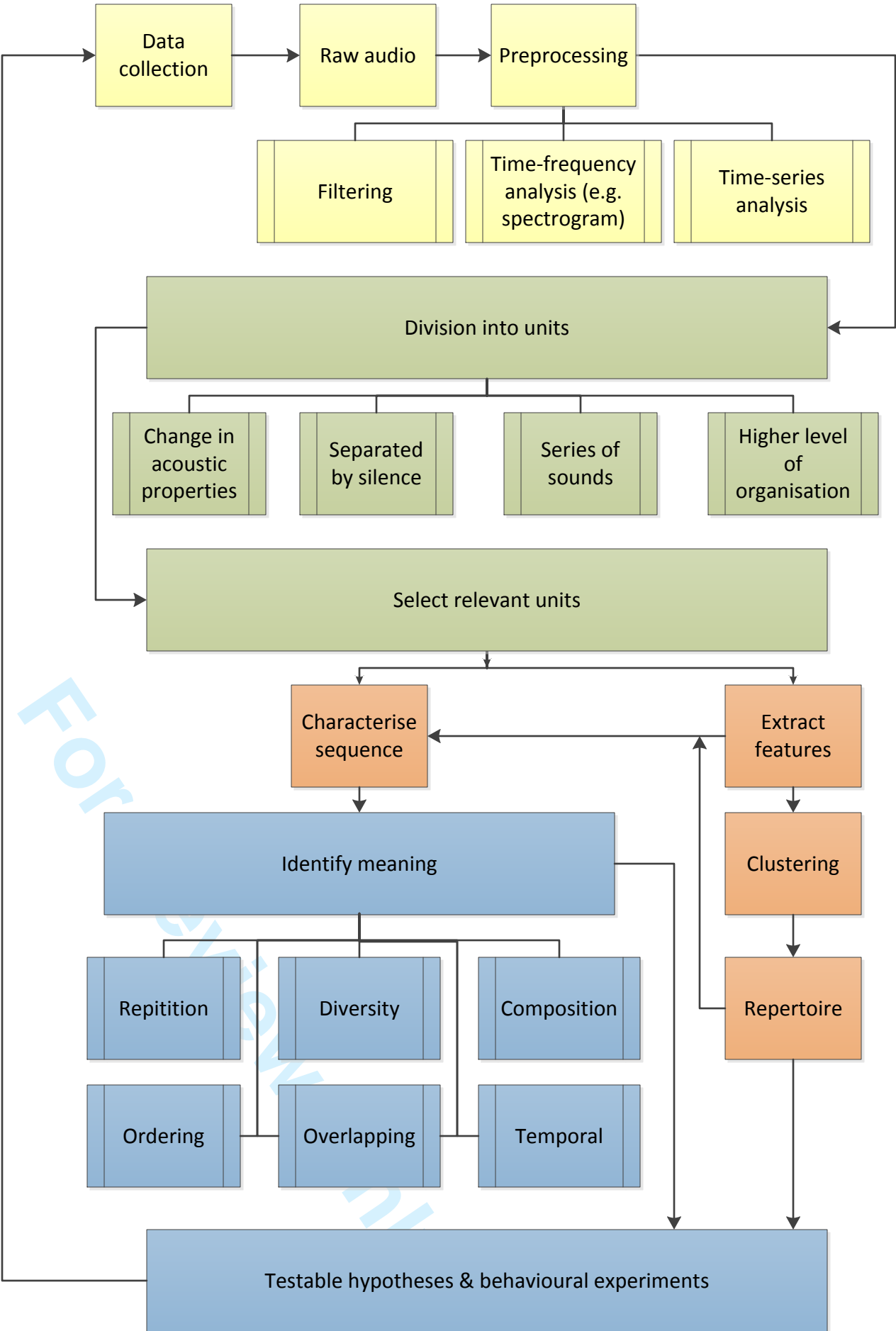
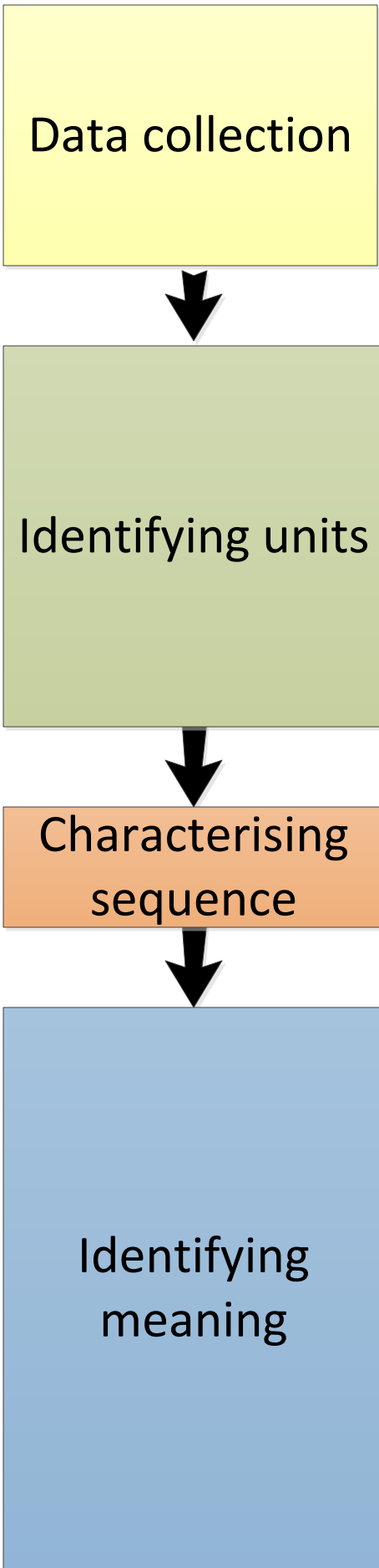
Unit criterion	Taxon	Goal of division into “units”			
		Descriptive	Production	Perception	Function
Separated by silence	Birds	Swamp sparrow <i>Melospiza georgiana</i> note: (Marler & Pickert, 1984)  Black capped chickadee <i>Poecile atricapillus</i> note: (Nowicki & Nelson, 1990)	Zebra finch <i>Taeniopygia guttata</i> syllable: (Cynx, 1990)	Swamp sparrow <i>Melospiza georgiana</i> note: (Nelson & Marler, 1989)  Black-capped chickadee <i>Poecile atricapillus</i> notes: (Sturdy, Phillmore & Weisman, 2000)	Carolina chickadee <i>Poecile carolinensis</i> and Black-capped chickadee <i>P. atricapillus</i> note composition → <b>predator, foraging activity, identity</b> : (Krams et al., 2012)
	Terrestrial mammals	Meerkat <i>Suricata suricatta</i> calls: (Manser, 2001)  Gibbon <i>Hylobates lar</i> phrase: (Raemaekers, Raemaekers & Haimoff, 1984)  Rock hyrax <i>Procavia capensis</i> songs: (Kershenbaum et al., 2012)  Free-tailed bat <i>Tadarida brasiliensis</i> syllable: (Bohn et al., 2008)  Mustached bat <i>Pteronotus parnellii</i> syllable: (Kanwal et al., 1994)		Meerkat <i>Suricata suricatta</i> calls: (Manser, 2001)	Meerkat <i>Suricata suricatta</i> calls → <b>predator type</b> : (Manser, 2001)  Rock hyrax <i>Procavia capensis</i> songs → <b>male quality</b> : (Koren & Geffen, 2009)  Free-tailed bat <i>Tadarida brasiliensis</i> syllable → <b>courtship</b> : (Bohn et al., 2008)

	Marine mammals	Humpback whale <i>Megaptera novaeangliae</i> unit: (Payne & McVay, 1971)  Killer whale <i>Orcinus orca</i> calls: (Ford, 1989)  Bottlenose dolphin <i>Tursiops truncatus</i> signature whistles: (Caldwell, 1965; McCowan & Reiss. 1995)		Bottlenose dolphin <i>Tursiops truncatus</i> signature whistles: (Janik, Sayigh & Wells, 2006)	Bottlenose dolphin <i>Tursiops truncatus</i> signature whistles → <b>individual identity</b> : (Sayigh et al., 1999; Harley, 2008)  Killer whale <i>Orcinus orca</i> calls: → <b>group identity</b> : (Ford, 1989)
Change in acoustic properties (regardless of silence)	Birds		Northern cardinal <i>Cardinalis cardinalis</i> : (Suthers, 1997)		
	Terrestrial mammals	Black-fronted titi monkey <i>Callicebus nigrifrons</i> : (Cäsar et al., 2012b)  Western gorilla <i>Gorilla gorilla</i> : (Salmi, Hammerschmidt & Doran-Sheehy, 2013)  Red titi monkey <i>Callicebus cupreus</i> : (Robinson, 1979)	Banded mongoose <i>Mungos mungo</i> : (Jansen, Cant & Manser, 2012)	Mustached bat <i>Pteronotus parnellii</i> : (Esser et al., 1997)	Black-fronted titi monkey <i>Callicebus nigrifrons</i> alarm calls → <b>predator type and behaviour</b> : (Cäsar et al., 2012a)  Western gorilla <i>Gorilla gorilla</i> vocalisations → <b>multiple functions</b> : (Salmi, Hammerschmidt & Doran-Sheehy, 2013)  Tufted capuchin monkeys <i>Sapajus nigritus</i> calls→ <b>predator type</b> : (Wheeler, 2010b)  Banded mongoose <i>Mungos mungo</i> close calls → <b>individual identity, group cohesion</b> : (Jansen, Cant and Manser, 2012)

	Marine mammals	<p>Bottlenose dolphin <i>Tursiops truncatus</i> whistle loops: (Caldwell, Caldwell &amp; Tyack, 1990)</p> <p>Killer whale <i>Orcinus orca</i>, subunit of calls: (Shapiro, Tyack &amp; Seneff, 2010)</p> <p>Humpback whale <i>Megaptera novaeangliae</i> subunit: (Payne &amp; McVay, 1971)</p> <p>Leopard seal <i>Hydrurga leptonyx</i> calls: (Klinck, Kindermann &amp; Boebel, 2008)</p>			
Series of sounds	Birds	Song sparrow song phrases (note clusters and trills): (Mulligan, 1966; Marler & Sherman, 1985)	Emberizid sparrow trills: (Podos, 1997)		Carolina chickadee <i>Poecile carolinensis</i> D-notes → <b>food availability</b> : (Mahurin & Freeberg, 2009)
	Terrestrial mammals	<p>Black-fronted titi monkey <i>Callicebus nigrifrons</i>: (Cäsar et al., 2012b; Cäsar et al., 2013)</p> <p>Mustached bat <i>Pteronotus parnellii</i> syllable: (Kanwal et al., 1994)</p> <p>Free-tailed bat <i>Tadarida brasiliensis</i> calls: (Bohn et al., 2008)</p>		<p>Black-fronted titi monkey <i>Callicebus nigrifrons</i>: (Cäsar et al., 2012a)</p> <p>Colobus <i>Colobus guereza</i> sequences: (Schel, Candiotti &amp; Zuberbühler, 2010)</p> <p>Tufted capuchin monkey <i>Sapajus nigritus</i> bouts:</p>	<p>Chimpanzee <i>Pan troglodytes</i> pant hoots → <b>foraging</b>: (Notman &amp; Rendall, 2005)</p> <p>Free-tailed bat <i>Tadarida brasiliensis</i> calls → <b>courtship</b>: (Bohn et al., 2008)</p>

		Hyrax <i>Procavia capensis</i> social calls: (Ilany et al., 2013)		(Wheeler, 2010b)	
		Chimpanzee <i>Pan troglodytes</i> pant hoots: (Notman & Rendall, 2005)			
	Marine mammals	Humpback whale <i>Megaptera novaeangliae</i> phrases: (Payne & McVay, 1971)			Bottlenose dolphin <i>Tursiops truncatus</i> signature whistles → <b>individual identity, group cohesion</b> : (Quick & Janik, 2012)
		Bottlenose dolphin <i>Tursiops truncatus</i> whistles: (Deecke & Janik, 2006)			Humpback whale <i>Megaptera novaeangliae</i> phrases → <b>unknown</b> : (Payne & McVay, 1971)
		Free-tailed bat <i>Tadarida brasiliensis</i> syllable: (Bohn et al., 2008)			
Higher levels of organisation	Birds		Swamp sparrow <i>Melospiza georgiana</i> trills: (Podos, 1997)	Song sparrow <i>Melospiza melodia</i> songs: (Searcy et al., 1995)	Skylark <i>Alauda arvensis</i> songs → <b>group identity</b> : (Briefer, Rybak & Aubin, 2013)
	Terrestrial mammals	Red titi monkey <i>Callicebus cupreus</i> syllable: (Robinson, 1979)	Rhesus-macaque <i>Macaca mulatta</i> vocalisations: (Fitch, 1997)	Putty-nosed monkey <i>Cercopithecus nictitans</i> sequences: (Arnold & Zuberbühler, 2006b)	Chimpanzee <i>Pan troglodytes</i> phrases → <b>group identity</b> : (Arcadi, 1996)
		Free-tailed bat <i>Tadarida brasiliensis</i> songs: (Bohn et al., 2008)	Black-fronted titi monkey <i>Callicebus nigrifrons</i> : (Cäsar et	Red titi monkey <i>Callicebus cupreus</i> syllable: (Robinson,	Putty-nosed monkey <i>Cercopithecus nictitans</i> sequences → <b>predators presence, group movement</b> : (Arnold & Zuberbühler, 2006b)

			al., 2013)	1979)	Tufted capuchin monkeys <i>Sapajus nigritus</i> calls→ <b>predator type:</b> (Wheeler, 2010b)
	Marine mammals	Humpback whale <i>Megaptera novaeangliae</i> theme and song: (Payne & McVay, 1971)			Humpback whale <i>Megaptera novaeangliae</i> song → <b>mating display - female attraction/male-male interactions</b> (Darling, Jones & Nicklin, 2006; Smith et al., 2008)



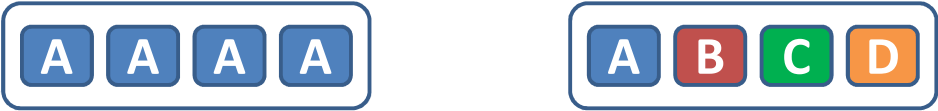


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

(a) Repetition



(b) Diversity



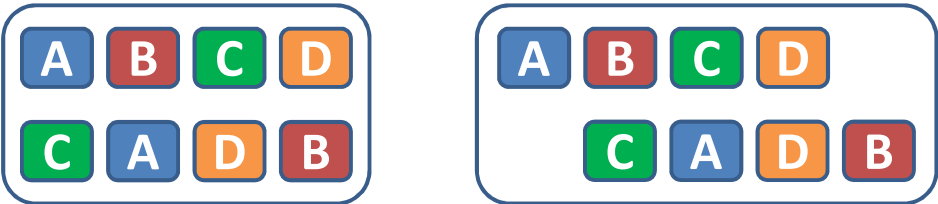
(c) Combination



(d) Ordering



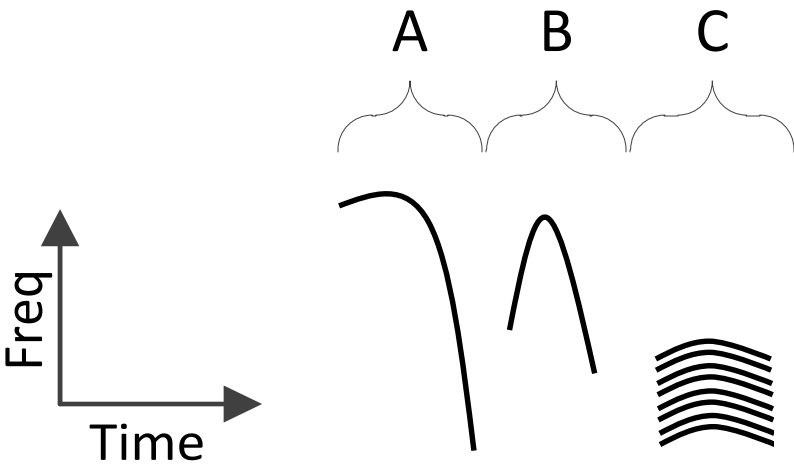
(e) Overlapping



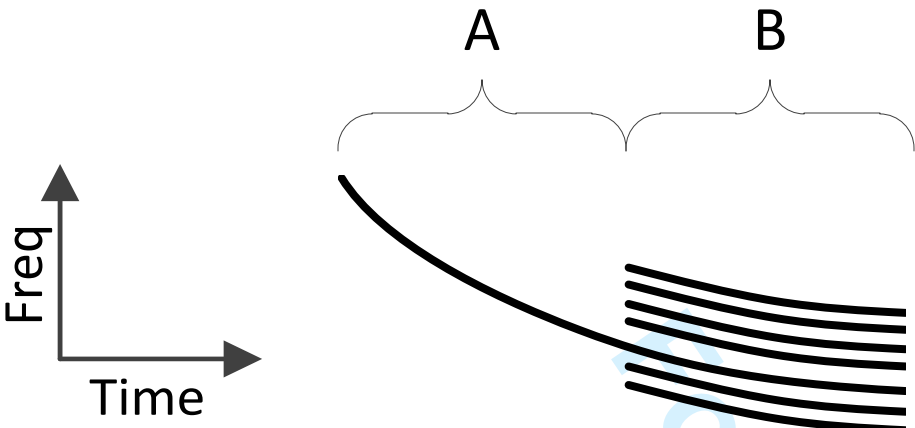
(f) Timing



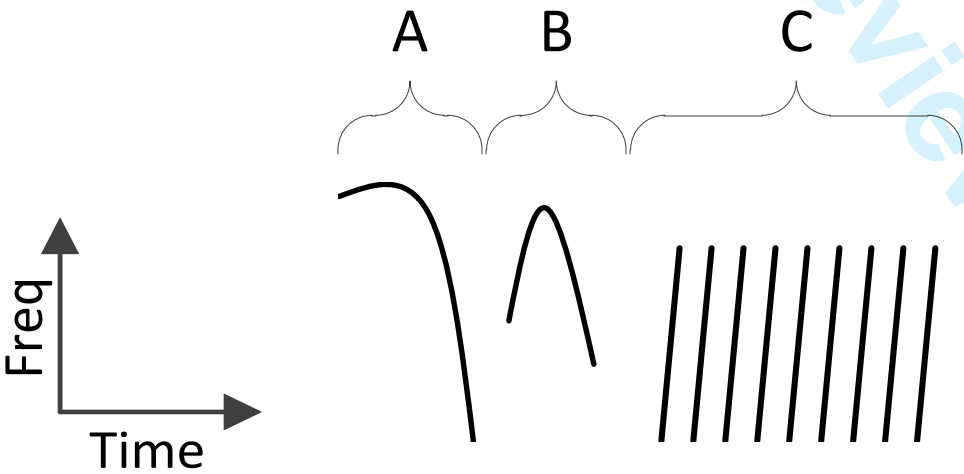
	Type	Criterion	Example
<b>a</b>	Repetition	Single unit repeated more than once	Chickadee D-note mobbing call (Baker & Becker. 2002)
<b>b</b>	Diversity	A number of distinct units are present. Order is unimportant.	Birdsong repertoire size (Searcy. 1992)
<b>c</b>	Combination	Set of units has different information from each unit individually. Order is unimportant.	Banded mongoose close calls (Jansen, Cant & Manser. 2012)
<b>d</b>	Ordering	Set of units has different information from each unit individually. Order is important	Human language, Humpback song (Payne & McVay. 1971)
<b>e</b>	Overlapping	Information conveyed in the relationship between sequences of two or more individuals	Sperm whale codas (Schulz et al. 2008)
<b>f</b>	Timing	Timing between units (often between different individuals) conveys information	Group alarm calling (Thompson & Hare. 2010)



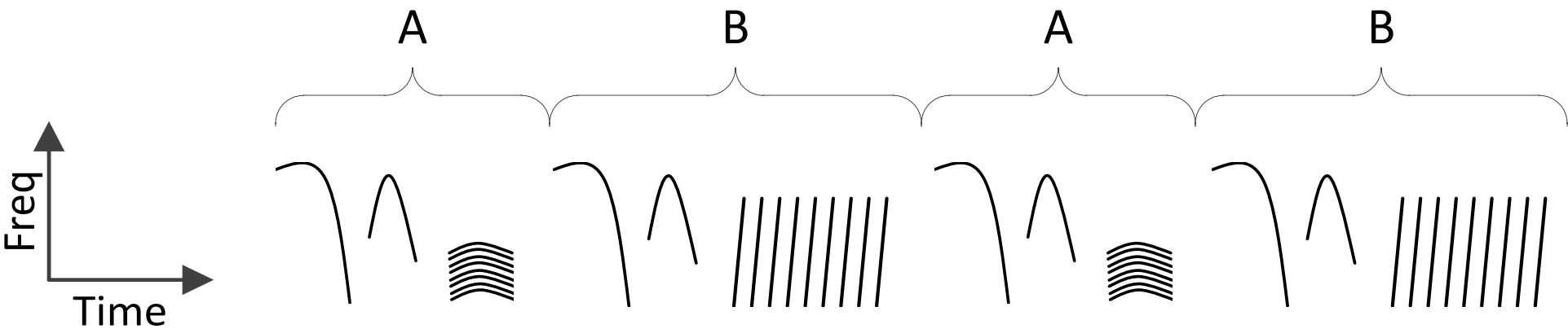
(a) Separated by silence



(b) Change in acoustic properties (regardless of silence)



(c) Series of sounds



(d) Higher levels of organization

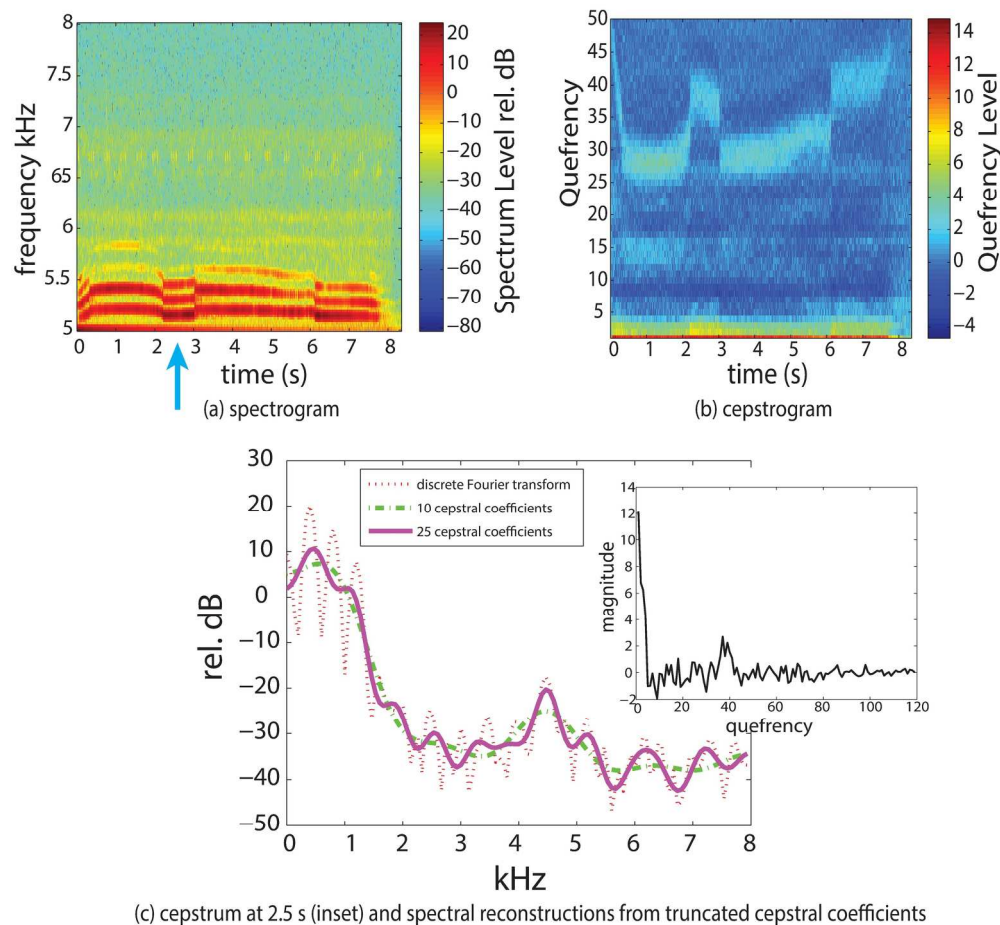
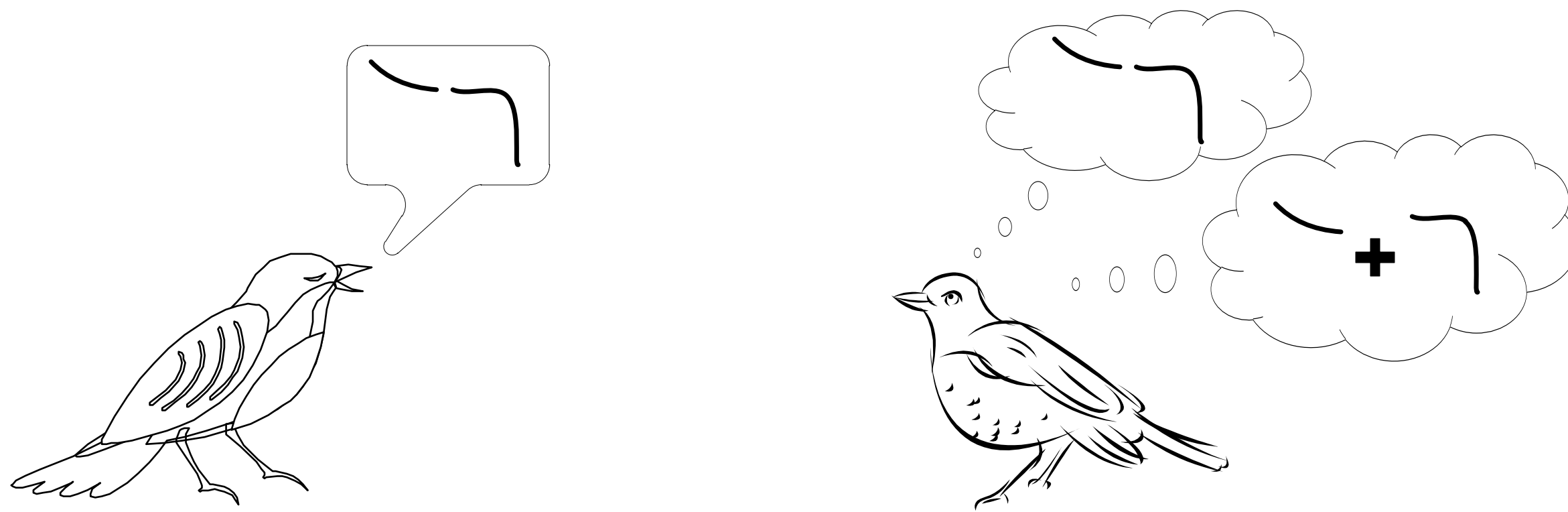


Figure 4. Example of cepstral processing of a grey wolf *Canis lupis* howl and crickets chirping. Recording was sampled at  $F_s = 16$  kHz, 8 bit quantization. (a) Standard spectrogram analyzed with a 15 ms Blackman-Harris window. (b) Plot of transform to cepstral domain. Lower quefrencies are related to vocal tract information. F0 can be determined from the "cepstral bump" apparent between quefrencies 25-45 and can be derived by  $F_s/\text{quefrency}$ . (c) Cepstrum (inset) of the frame indicated by an arrow (2.5 s) along with reconstructions of the spectrum created from truncated cepstral sequences. Fidelity improves as the number of cepstra are increased.

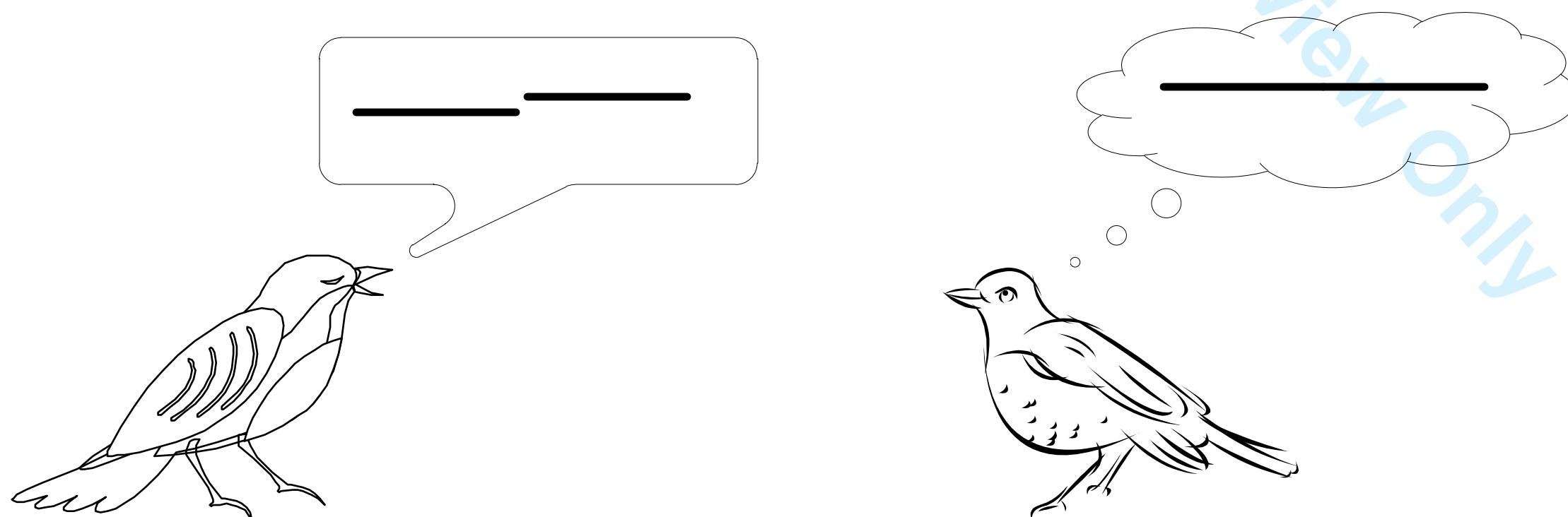
190x178mm (300 x 300 DPI)



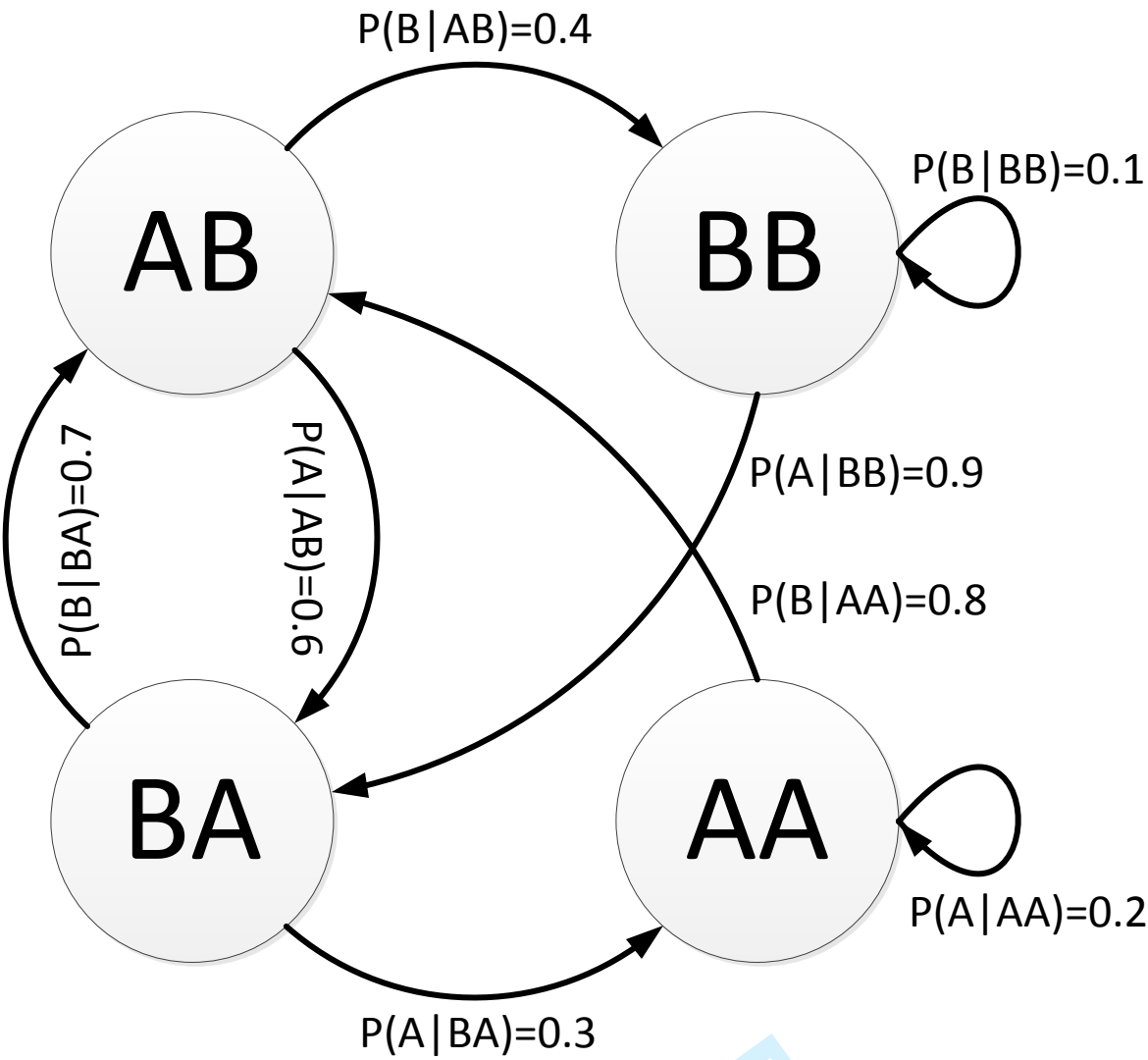
(a) Perceptual binding. Two discrete acoustic elements may be perceived by the receiver either as a single element, or as two separate ones



(b) Categorical perception. Continuous variation in acoustic signals may be interpreted by the receiver as discrete categories

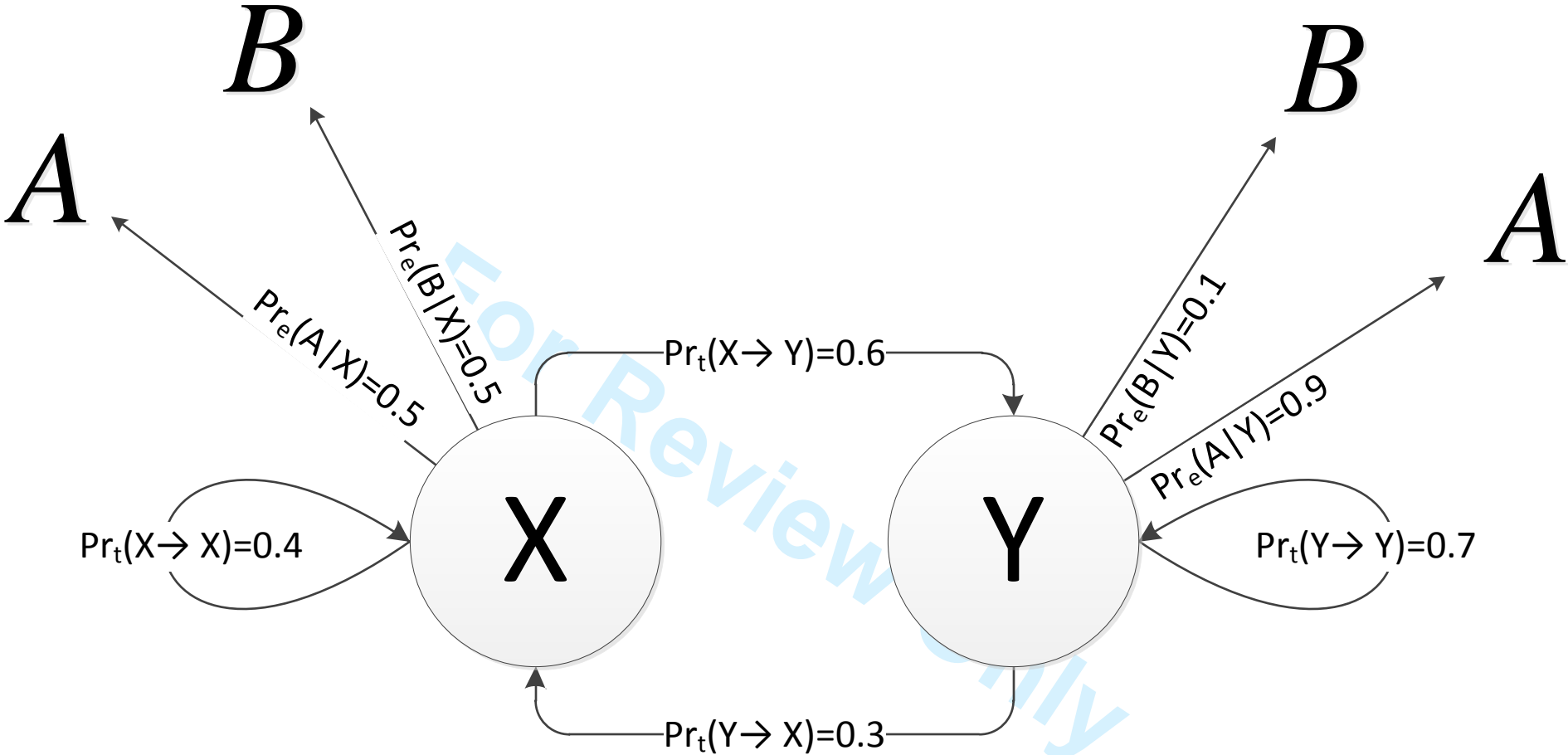


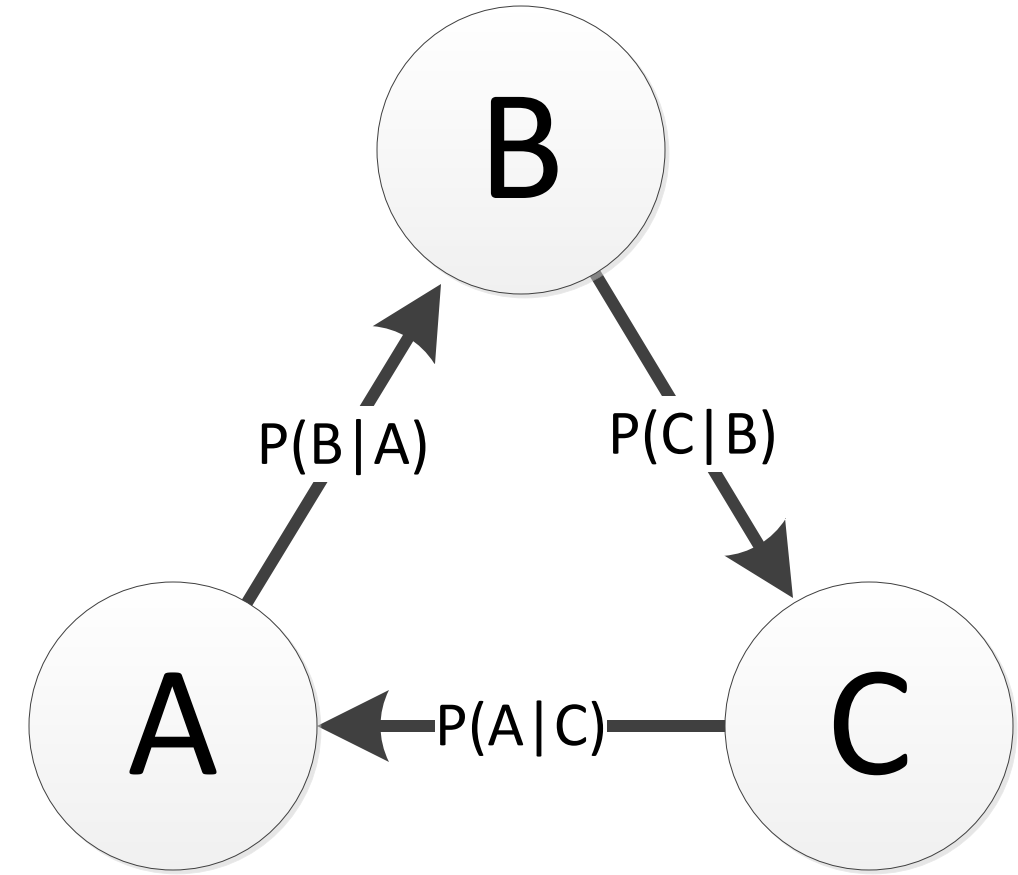
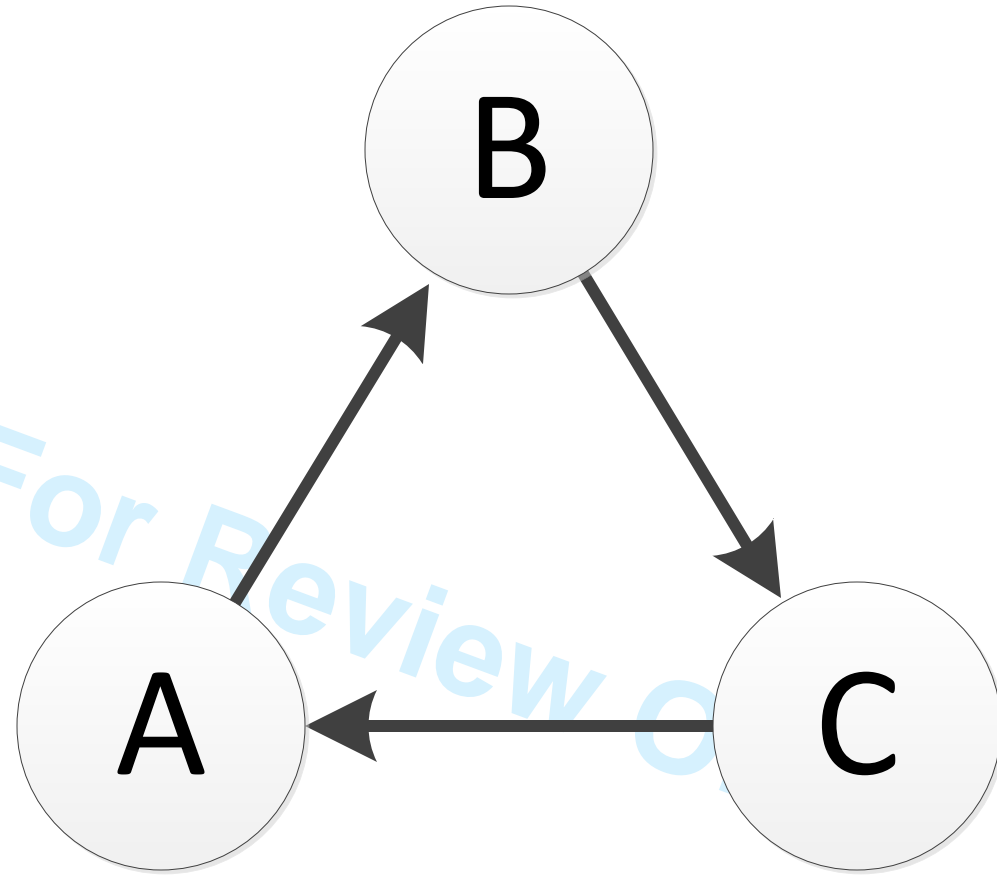
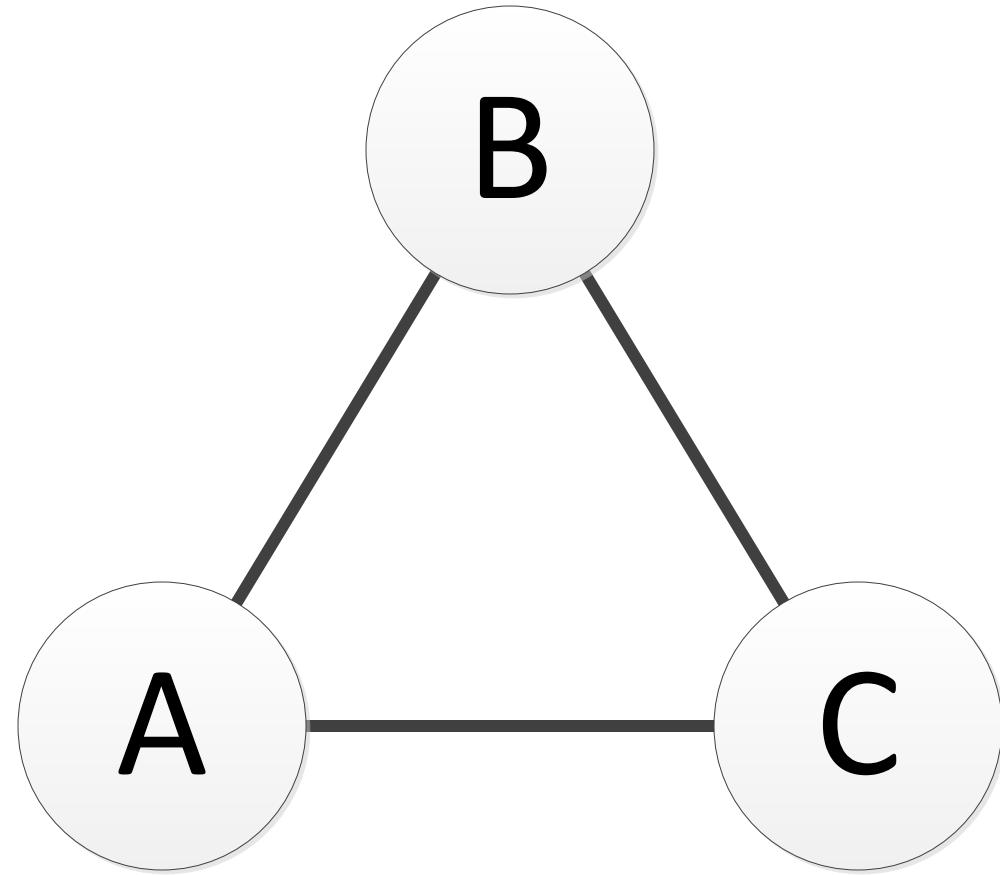
(c) Spectrotemporal constraints. If the receiver cannot distinguish small differences in time or frequency, discrete elements may be interpreted as joined



	A	B
$P(X AA)$	0.2	0.8
$P(X AB)$	0.6	0.4
$P(X BA)$	0.3	0.7
$P(X BB)$	0.9	0.1

For Review Only







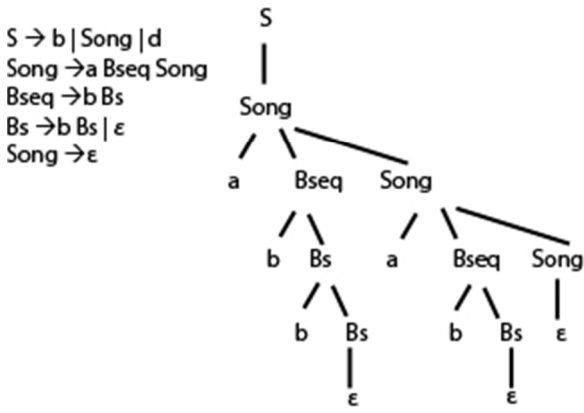


Figure 9. Grammar (rewrite rules) for approximating the sequence of acoustic units produced by Eastern Pacific blue whales *Balaenoptera musculus*. There are three acoustic units, A, B, and D (Oleson, Wiggins & Hildebrand. 2007), and the sequence begins with a start symbol S. Individual B or D calls may be produced, or song, which consists of repeated sequences of an A call followed by one or more B calls. The symbol | indicates a choice, and  $\epsilon$ , the empty string, indicates that the rule is no longer used. A derivation is shown for the song ABBAB. Grammar produced with contributions from Ana Širović (Scripps Institution of Oceanography).

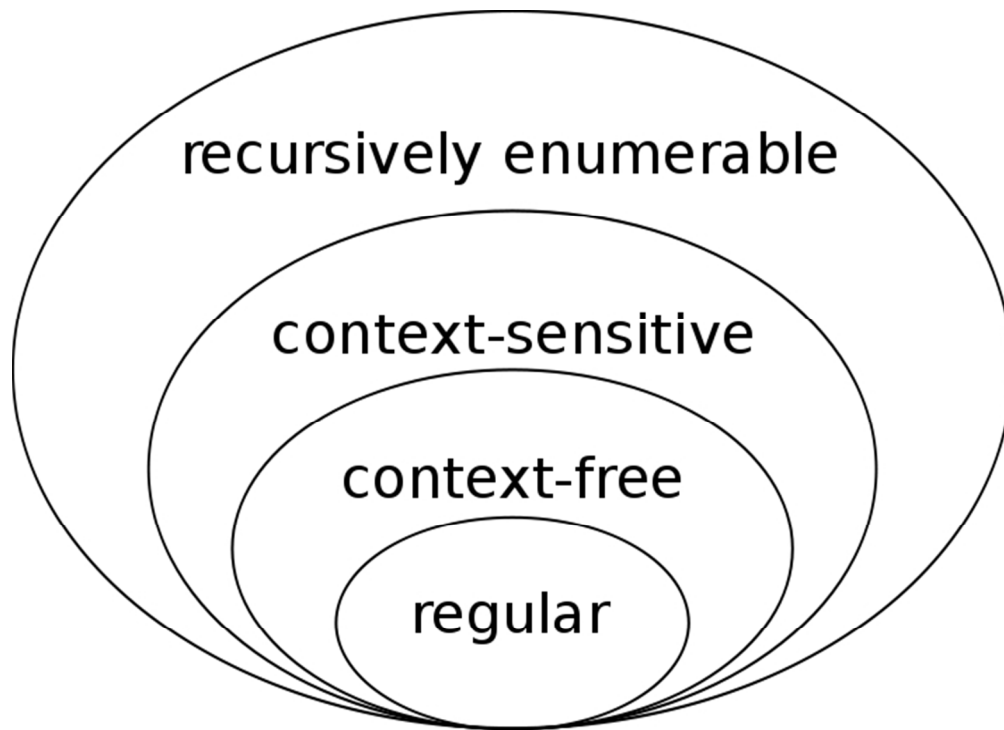


Figure 10. The classes of formal grammars known as the Chomsky hierarchy (Chomsky, 2002). Each class is a generalisation of the class it encloses, and is more complex than the enclosed classes. Image publicly available under the Creative Commons Attribution-Share Alike 3.0 Unported license.  
[https://commons.wikimedia.org/wiki/File:Wiki\\_inf\\_chomskeho\\_hierarchia.jpg](https://commons.wikimedia.org/wiki/File:Wiki_inf_chomskeho_hierarchia.jpg)  
251x181mm (72 x 72 DPI)

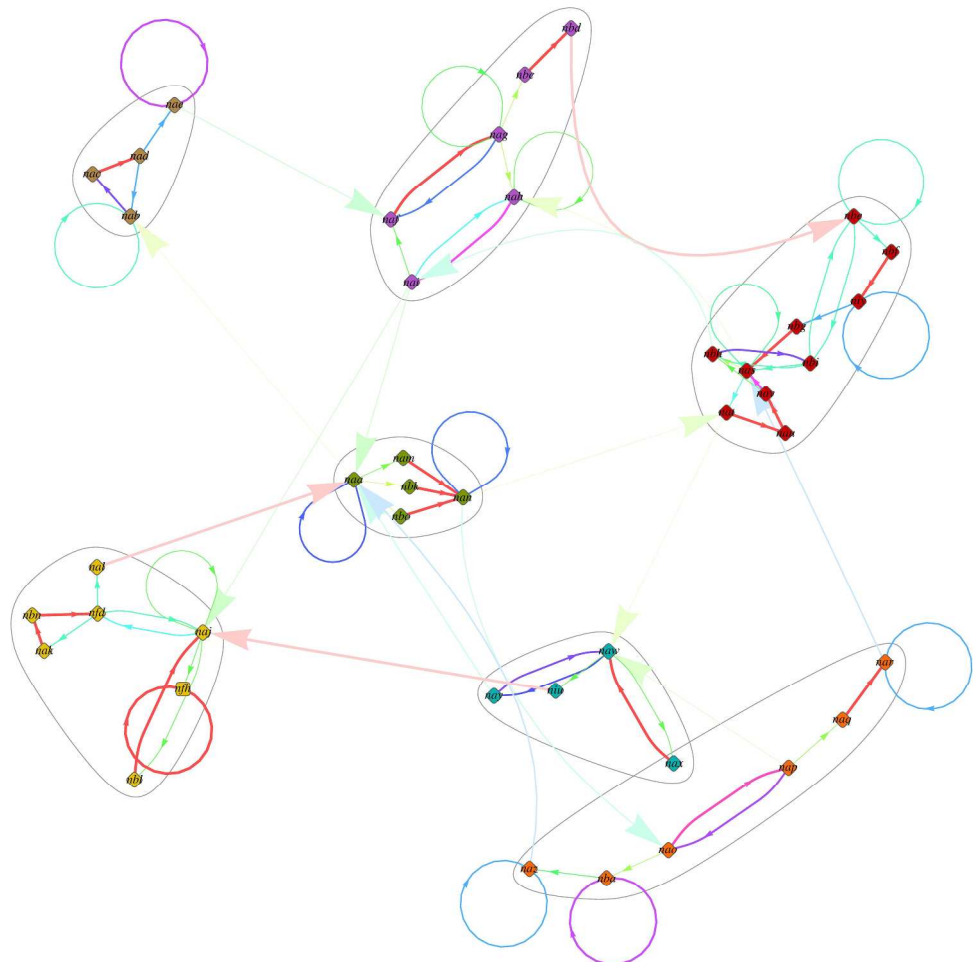


Figure 11. Network diagram for a sample of California thrasher song data. Nodes indicate distinct phrase types (acoustic units), arrows indicate transitions, with the colour of the arrow indicating the strength of probability of transition between two units (hotter colours represent more likely transitions).  
281x280mm (300 x 300 DPI)