

1 Comparing call-based vs. subunit-based methods for categorizing
2 Norwegian killer whale (*Orcinus orca*) vocalisations

3

4 Ari Daniel Shapiro*

5 Peter L. Tyack

6 Biology Department, Woods Hole Oceanographic Institution, MS #50, Woods Hole,

7 Massachusetts 02543, USA

8

9 Stephanie Seneff

10 Massachusetts Institute of Technology Computer Science and Artificial Intelligence Laboratory,

11 32 Vassar Street, Cambridge, MA 02139, USA

12

13 Submitted August 13, 2010

14

15

16

17

18

19

20 * Corresponding author; Electronic mail: ashapiro@whoi.edu; Telephone: 617.417.6996; Fax:

21 508.457.2041.

22

23 Word count: 7085

24 ABSTRACT

25 Students of animal communication face significant challenges when deciding how to
26 categorise calls into subunits, calls, and call series. Here, we use algorithms designed to parse
27 human speech to test different approaches for categorising calls of killer whales. Killer whale
28 vocalisations have traditionally been categorised by humans into discrete call types. These calls
29 often contain internal spectral shifts, periods of silence, and synchronously produced low and
30 high frequency components, suggesting that they may be composed of subunits. We describe
31 and compare three different approaches for modelling Norwegian killer whale calls. The first
32 method considered the whole call as the basic unit of analysis. Inspired by human speech
33 processing techniques, the second and third methods represented the calls in terms of subunits.
34 Subunits may provide a more parsimonious approach to modelling the vocal stream since (1)
35 there were fewer subunits than call types; (2) nearly 75% of all call types shared at least one
36 subunit. We show that contour traces from stereotyped Norwegian killer whale calls yielded
37 similar automatic classification performance using either whole calls or subunits. We also
38 demonstrate that subunits derived from Norwegian stereotyped calls were detected in some
39 Norwegian variable (non-stereotyped) calls as well as the stereotyped calls of other killer whale
40 populations. Further work is required to test whether killer whales use subunits to generate and
41 categorize their vocal repertoire.

42

43

44 **KEYWORDS:** Killer whale, subunit, vocalisations, call type

45

46

47 An important question in animal behaviour and communication concerns how the brain
48 encodes and decodes a sequence of acoustic signals. Most studies of animal communication
49 categorize vocalizations in terms of complete calls. An alternative view hypothesizes that many
50 vocalizations may be made up of smaller subunits. The vocal production system might generate
51 complete vocalizations by sequencing these subunits (Fee et al. 2004; Glaze & Troyer 2006,
52 2007). The auditory system may parse the acoustic stream into subunits, which are then
53 combined to yield categorization of complete calls. For example, humans first process speech
54 into subunits called phonemes, which are then combined into meaningful words. Songbirds may
55 similarly construct songs from subunits called syllables, which are thought to be the basic unit of
56 song production (Glaze & Troyer 2006, 2007). Humans and songbirds can recombine a small
57 number of subunits to form a large number of signals.

58
59 Vocal production learning is the process by which vocal signals are modified due to
60 experience with the signals of other individuals (Janik & Slater 1997, 2000). There is evidence
61 for vocal production learning in humans, many species of songbirds and certain species of
62 cetaceans (see Janik & Slater 1997, 2000 for reviews). For species capable of vocal learning,
63 neural mechanisms for learning subunits and sequences of those subunits may enable production
64 of novel calls. Zebra finches (*Taeniopygia guttata*), for example, learn serial strings of song
65 syllables and intervening periods of silence from a variety of adult tutors and can then reorder
66 these syllables to produce their own song (Williams & Staples 1992). In this paper, we
67 investigate whether killer whale calls can be categorized in terms of subunits, and we explore the
68 implications of the subunit level of analysis.

69

70 Certain features help define the boundaries of subunits. In birdsong, for example,
71 syllables are generally defined as continuous and discrete elements separated by silence (Lemon
72 & Chatfield 1971; Eales 1985; Eens et al. 1989). Zebra finches whose songs were
73 experimentally interrupted during a song syllable usually stopped singing at the end of the
74 syllable, suggesting that the syllable was a unit of sound production (Cynx 1990; Franz and
75 Goller 2002; and see Riebel & Todt 1997 for a similar result for nightingales (*Luscinia*
76 *meigarhynchos*)). In human speech, positions of maximum spectral transition are important for
77 consonant and vowel perception (Furui 1986).

78

79 The classic mode for categorizing the calls of killer whales (*Orcinus orca*) is at the level
80 of complete stereotyped calls (see Ford 1987; Strager 1993; Filatova et al. 2004). By contrast,
81 Yurk (2005) proposed that the calls of killer whales from Alaska and British Columbia may be
82 composed of subunits. He used abrupt and discontinuous spectral shifts to define boundaries
83 between subunits in killer whale vocalisations. Two features of Norwegian killer whale calls
84 suggest that they may also be comprised of smaller subunits. First, Norwegian killer whales
85 produce compound calls, which are concatenations of multiple discrete calls, each of which can
86 be produced individually or within other compound calls (Strager 1993, 1995). This might
87 suggest that Norwegian killer whales take advantage of a similar mechanism to fashion the call
88 types themselves from a set of even smaller subunits. Second, call subtypes have been described
89 (Ford 1991; Strager 1995), which can be formed by rearranging the subunits. Moreover, Yurk
90 (2005) extracted subunits (as contour traces from spectrograms) from the calls of resident and
91 transient killer whales living in British Columbia, categorised these traces by eye (using their
92 “gestalt” to aid the differentiation) and found that human classifiers agreed with these divisions.

93 Examples of the Norwegian call types motivating this work are depicted in Figure 1, and are
94 consistent with the notion that a set of subunits might be rearranged to generate a larger
95 repertoire of call types.

96
97 Killer whale call types have been classified into similar categories by human observers
98 sorting calls aurally and spectrographically and by neural networks relying on temporal and
99 spectral features of the fundamental frequency (Deecke et al. 1999; Deecke & Janik 2006). A
100 portion of the killer whale repertoire has been labelled “variable,” a miscellaneous class of vocal
101 behaviour containing the calls that have not sorted neatly into one of the stereotyped categories
102 (Rehn et al. 2007). The stereotyped calls of killer whales can contain synchronously produced
103 low and high frequency components (LFCs (80 – 2400 Hz) and HFCs (2 – 12 kHz), respectively,
104 Hoelzel & Osborne 1986; Ford 1987; Miller & Bain 2000) that are presumably generated by two
105 independent sources (see Miller et al. 2007 for a discussion). Such independent control would
106 allow these animals to increase their repertoire simply by varying the LFC and HFC pairings.

107
108 Here, we take advantage of techniques that have been developed in the field of human
109 speech recognition to compare different approaches for modelling the stereotyped calls of
110 Norwegian killer whales. One of the hallmarks of human language includes an ability to use
111 vocal production learning to generate a vast array of words from a set of a few dozen phonemic
112 units (Nowak et al. 2000; Hauser et al. 2002). Early attempts at using words to drive automatic
113 speech recognition gave way to phonemic representations that improved performance on large
114 vocabularies considerably. Vocal learning by killer whales has been suggested (though not
115 demonstrated) by several studies (Bowles et al., 1988; Foote et al., 2006; Nousek et al., 2006;

116 Riesch et al., 2006), offering additional incentive for considering the parallels of vocal
117 production in killer whales, songbirds and humans. Traditionally, killer whale call types have
118 been regarded as the fundamental units of vocal production in this species. Fashioned after the
119 approaches developed for human speech research, in this study killer whale vocalisations were
120 decomposed to test whether a smaller set of acoustic segments may define subunits that can be
121 assembled to form the call repertoire. This research does not propose that early categorizations
122 were incorrect. Rather, we evaluated whether a subunit approach could yield similar call
123 categories.

124

125 In this paper, we first evaluated and compared three approaches for classifying calls of
126 Norwegian killer whales. One approach was based on traditional classification techniques,
127 which used the whole call type as the basic unit of analysis. The other two approaches modelled
128 call types using subunits, which could be shared or not shared across the call types. Second, we
129 tested the possibility of modelling certain variable (non-stereotyped) calls using subunits derived
130 from stereotyped calls. Finally, inspired by the small yet universal set of phonemes that are sub-
131 sampled to form each human language (see Zhu et al. 2005), we quantified how well the
132 inventory of subunits derived from Norwegian stereotyped calls could be used to characterise the
133 repertoire of resident and transient killer whales from the waters off British Columbia.

134

135 METHODS

136

137 *Data Collection*

138

139 Fieldwork was conducted in November 2005 and 2006 in the northern Norwegian fjords
140 of Tysfjord and Vestfjord near the Lofoten islands (~68°15' N, ~16°E). Fourteen free-ranging
141 killer whales were outfitted with digital archival tags (DTAGs) that sampled audio at 96 kHz
142 using 16-bit resolution and individual movement at 50 Hz (Johnson & Tyack 2003). The
143 dimensions of the tag in its housing were 20 cm x 10 cm x 4 cm. The tag including its housing
144 and attachment weighed 330 g in air (<0.0003% of weight of whale), and it was slightly buoyant
145 in the water. (Since the identity of the social group in which the killer whales spent much of
146 their time was not known reliably for all of the study animals, categorising calls at the group
147 level was not feasible.) The animals were approached in a rigid hull inflatable boat and the tags
148 were attached with suction cups using a 7m carbon fibre hand pole. These approaches were
149 conducted following procedures to minimise disturbance of the tagged whale and other whales in
150 the group. The protocol called for leaving any whale bothered by the approach (e.g., showing
151 strong avoidance behaviour), but no negative responses were observed. If tagging attempts were
152 repeatedly unsuccessful (over the course of ~2 hours and corresponding to ~10-15 approaches)
153 with a particular animal, a different individual or group was selected for tagging to minimise
154 disturbance. If a whale was disturbed by the tag, it could have removed the tag by a sudden
155 motion or burst of acceleration. There was no evidence that the tags impeded normal behaviour.

156 The tags were programmed to have the suction cups release at a predetermined time with
157 the longest attachment lasting 4.7 hours. A VHF (very high frequency) beacon was used to track
158 the tagged animal from an observation platform (i.e., the sailboat *Iolaire* or the Norwegian
159 research vessel *Sverdrup*) and to recover the tag once it released after a pre-programmed period
160 of time. The data were offloaded and then cleared from memory to allow re-deployment of the
161 tag. A total of 31.8 hours of recordings were made from 13 animals engaged in feeding,

162 travelling and resting behaviours (one of the attachments did not contain any vocalisations).

163 In addition, calls produced by resident and transient killer whales off of British Columbia
164 were recorded with towed hydrophone arrays and single hydrophones, respectively (see Miller &
165 Tyack 1998; Deecke et al. 2005 for data collection and processing details), and were kindly
166 provided for the analysis here. The fieldwork was approved by the Norwegian Animal Care
167 Committee (Forsøksdyrutvalget ref 2004/20607-4) and followed the WHOI Institutional Animal
168 Care and Use Committee approved protocol for field work using DTAGs.

169

170 *Call Type Assignments*

171

172 All recordings were audited manually and calls with a sufficient signal-to-noise ratio
173 (SNR) were used for classification purposes. Calls were manually labelled by visual inspection
174 of spectrograms by three observers and the lead author. Earlier studies using this approach have
175 demonstrated high inter-observer reliability scores and compare favourably to automated
176 approaches involving neural networks (Bain 1986; Ford 1991; Deecke et al. 1999; Yurk et al.
177 2002; Deecke & Janik 2006). Call types were matched whenever possible to pre-existing call
178 type labels for Norwegian killer whales provided in the literature. Strager (1993) identified the
179 first 34 call types (i.e., N1 to N34) and Van Opzeeland et al. (2005) added call types N35 to N63.
180 New call types identified in this research were assigned new numbers (N64 to N103; see Shapiro
181 2008 for spectrograms). A total of 487 calls were considered variable because they could not be
182 classified to call type. Calls from the resident and transient killer whales off British Columbia
183 had already been sorted to type according to Ford (1987) and Deecke (2003).

184

185 *Contour Tracing*

186

187 A pitch tracking algorithm developed for human telephone speech that relies on the
188 harmonic structure of a vocal signal (Seneff 1978; Wang 2001) was used to trace the
189 fundamental frequency of both the low and high frequency components (LFC and HFC,
190 respectively) of killer whale calls. (For details on the algorithm, see Shapiro & Wang (2009).)
191 All pitch contours were checked manually against the original spectrograms and if necessary,
192 portions were smoothed and re-traced.

193

194 Research on birdsong, killer whale calls and human speech has used silent periods and
195 abrupt spectral shifts to define subunit boundaries. Similarly, in this study we designated subunit
196 boundaries whenever there was at least a 0.1s span of silence or a 500 Hz spectral jump
197 occurring within 0.25s (Figure 2; see Shapiro 2008 for the contour traces of all calls). The LFC
198 and HFC of a call were often divided into subunits at different time points. For several calls, the
199 frequency changes were gradual instead of abrupt (i.e., lasting longer than 0.25s). These calls
200 were divided into two or three subunits according to call type using a segmentation algorithm.
201 This algorithm fit the call contour to a polynomial and designated the subunit boundaries at time
202 points of maximum deviation (see Shapiro 2008 for a full explanation of this algorithm and
203 Adam 2008 for another approach).

204

205 *Call Classification Experiments*

206

207 In the rest of this paper, we propose and evaluate three different approaches towards

208 modelling killer whale calls. These approaches used the whole call type, unshared subunits and
209 shared subunits as the basic units of analysis, respectively. For evaluation purposes, 3530 calls
210 were collected and manually labelled as described earlier. These calls were then randomly split
211 into a training set of 90% of the data, and a test set of 10%. Only call types with at least 10
212 exemplars were considered. Acoustic features based on Legendre polynomial coefficients were
213 extracted and used to train Gaussian mixture models (GMMs) for each call type. GMMs have
214 been shown to achieve robust performance when modelling the acoustic subunits of human
215 speech, or phonemes (e.g., Bonafonte et al. 1996). Figure 3 illustrates how each of the
216 approaches modelled the contours of three sample call types.

217

218 *Whole contour experiment (WCE)*

219

220 In the first experiment, the basic unit of killer whale vocal production was the whole call
221 type. A 4th-order Legendre polynomial was fit to each LFC and HFC, linearly interpolating
222 between subunits separated by silent intervals. Although any polynomial family might have
223 performed similarly, this class and order of Legendre polynomials were selected because of their
224 precedence in human speech research (i.e., Legendre polynomials were used to characterise the
225 tones of Mandarin Chinese: Chen & Wang 1990; Wang 2001). This interpolation permitted the
226 contour of the entire call to be represented continuously. Six parameters that characterised the
227 contour properties were calculated: (1) the duration, (2 – 5) the first 4 Legendre coefficients,
228 which represented the basic spectral shape, and (6) the root mean square (RMS) error between
229 the polynomial curve and the actual trace, capturing the extent of frequency modulation. This
230 feature space was designed to provide a simple rendering of the dataset, but lacked information

231 about the energy, harmonics and other details of the original signal.

232

233 A GMM was trained for each call type by computing the mean vector and covariance
234 matrix of the 6-dimensional acoustic features extracted over the corresponding training set.
235 Classification was performed by assigning each call in the test set to the call type whose GMM
236 yielded the maximum likelihood. For the test set calls containing both an LFC and an HFC
237 contour, the likelihood of the LFC evaluation was added to that of the HFC evaluation. The call
238 type yielding the maximum sum constituted the matching class.

239

240 For all experiments, classification was aided by only considering as possible matches
241 those call types containing the same components (i.e., LFC alone, HFC alone, LFC and HFC
242 together) and number of subunits (see USE below) as the test contour. All three experiments
243 were assessed based on call type error rate (i.e., how often an incorrect call type was chosen for
244 each call type in the test set). Results were reported as error rates \pm the standard error (s.e.),
245 which allowed comparison between the three experiments.

246

247 *Unshared subunit experiment (USE)*

248

249 In the second experiment, call types were assumed to be composed of subunits, which
250 where were not shared across different call types. Each call type was characterized by a unique
251 subunit or sequence of these subunits. For example, call type N9.2 consisted of two subunits,
252 which represented the sections before and after the spectral shift (Figure 2). A separate GMM
253 was trained for each of the subunits. To classify a new incoming call, we computed the

254 likelihood(s) of each of the contour subunit(s) matching all of the call type subunits. For calls
255 with multiple subunits, the likelihood was summed over all the subunits; the sequence of
256 subunits providing the maximum sum constituted the matching class. Suppose, for example, we
257 had an incoming call, *C*, and the segmentation algorithm determined that it consisted of segments
258 *A* and *B*. For each of these two segments, the 6-dimensional acoustic features were extracted and
259 scored against all the trained GMMs. Each segment was then labelled with the subunit
260 corresponding to the highest likelihood. If *A* and *B* were matched to the first and second subunits
261 of call type N9.2, call *C* would be classified as N9.2. In a second scoring for this experiment, a
262 call type was considered a correct match if the call contained a sequence of the same shared
263 subunits (see SSE) as the correct call type. This ensured that the SSE was not given any unfair
264 advantage over the USE because of differences in how the subunits were labelled. Rather, the
265 difference was whether the subunits were unshared or shared across call types.

266

267 *Shared subunit experiment (SSE)*

268

269 In the third experiment, call types were assumed to be composed of subunits, where
270 subunits could be shared across different call types. The subunits presented in the previous
271 section were collapsed into classes using the following technique: 1) a GMM was trained for
272 each subunit using all of its corresponding data; 2) each subunit was then classified using
273 maximum likelihood; 3) the subunits that were the most confusable were collapsed into the same
274 class. Certain call types had exemplars containing both an LFC and HFC, but had entries in the
275 test set that contained traces of only the LFC or HFC because the SNR of the other component
276 was too low for accurate pitch tracking. Due to the absence of one of the frequency components,

277 the call type could not be resolved appropriately and these test set entries were considered to
278 match incorrectly. Using the same sample across the three experiments allowed us to compare
279 the results equivalently.

280

281 *Variable Norwegian Killer Whale Calls and Stereotyped British Columbia Killer Whale Calls*

282

283 In this section, variable (non-stereotyped) Norwegian calls and stereotyped calls from
284 resident and transient killer whales in British Columbia will be referred to collectively as
285 “auxiliary calls.” The segmentation algorithm presented earlier was applied to variable
286 Norwegian call traces whose RMS error (between the contour and the 4th-order Legendre
287 polynomial) exceeded that of the stereotyped subunits. The resulting divisions were inspected
288 visually and only segmentation decisions corresponding to abrupt (within 0.25s) frequency
289 changes were retained. For each of the 26 Norwegian stereotyped shared subunits described
290 earlier, a set of likelihood values was collected as follows: 1) each subunit from the test set was
291 evaluated against the 26 trained GMMs; 2) the best likelihood value was retained if the subunit
292 matched correctly. These will be referred to as correct-match likelihood values. Each auxiliary
293 call subunit was classified to the Norwegian stereotyped shared subunit that returned the largest
294 likelihood, L_{max} . Success was evaluated by comparing L_{max} to the correct-match likelihood
295 values of the matching subunit. For example, if L_{max} was greater than 25% of the correct-match
296 likelihood values, this would be considered a match at the 25% correct-match threshold.

297

298 RESULTS

299

300 The 3530 calls used in this analysis belonged to 31 call types (16 with LFC only, 3 with
301 HFC only and 12 with both LFC and HFC). A total of 62 different subunits (39 LFC subunits
302 and 23 HFC subunits) was considered by the unshared subunit experiment (USE), and 26
303 different subunits (19 LFC subunits and 7 HFC subunits) for the shared subunit experiment
304 (SSE). The categorization efficiency of all three experiments was equivalent (error rates \pm
305 standard error for WCE: 0.074 ± 0.014 ; USE first scoring: 0.088 ± 0.015 ; USE second scoring:
306 0.077 ± 0.014 ; SSE: 0.071 ± 0.014). The distributions of the error rates formed from 100 runs of
307 these experiments overlapped (Figure 4). In other words, representing stereotyped calling
308 behaviour in terms of whole calls, unshared subunits and shared subunits all provided equally
309 strong categorization results.

310

311 Results for the Norwegian variable calls ($N = 576$ calls; 675 subunits) and British
312 Columbia Northwest resident ($N = 192$ calls and subunits, since each of the call types used here
313 contained a single subunit) and transient ($N = 162$ calls; 207 subunits) stereotyped calls are
314 presented in Table 1. The values in this table correspond to the percentage and number of
315 auxiliary calls whose L_{\max} exceeded the correct-match threshold for that particular subunit (see
316 Methods). That is, a portion of the auxiliary calls and subunits matched successfully with the
317 Norwegian stereotyped subunits. The Norwegian variable calls matched more often than British
318 Columbia calls with the Norwegian stereotyped calls, and the transient calls matched more
319 frequently than resident calls. Figure 5 plots a sampling of the best matches for each auxiliary
320 call category. The first column corresponds to a 90% correct-match threshold and the second
321 and third columns correspond to 25% correct-match thresholds.

322

323 Although each call type was defined by a unique combination of subunits, certain LFC or
324 HFC subunit sequences were occasionally shared across types. We examined how many call
325 types drew on the pool of shared subunits and what their patterns of occurrence were. Seven
326 LFC and 4 HFC subunits were formed from subunits shared by at least two different call types
327 (Figure 6). Table 2 counts the number of call types composed entirely of call-specific subunits,
328 entirely of shared subunits, or a mixture of shared and call-specific subunits.

329

330 Eighteen of the 31 call types were comprised entirely of shared subunits and another five
331 contained at least one shared subunit (together constituting 53% of the calls, Table 2). Within
332 our recordings, some segments were used much more frequently than others, including LFC
333 subunit 4 and HFC subunits 1 and 2 (see Figure 6 for subunit labels). Distinct call types resulted
334 when the same arrangements of HFC subunits were paired with different combinations of LFC
335 subunits (and *vice versa*). Three shared LFC and two shared HFC subunits were produced at the
336 same position within a sequence, consistently beginning or ending multiple call types. For
337 example, when paired, HFC subunit 2 always preceded HFC subunit 1. Subunits could be
338 ordered more flexibly as well: two shared LFC and one shared HFC subunits were produced at
339 different positions within a sequence.

340

341 DISCUSSION

342

343 Human speech processing methods (i.e., a pitch tracking algorithm, segmentation
344 algorithm, and Gaussian mixture models (GMMs)) were successfully adapted and applied here to
345 analyze Norwegian killer whale vocalisations. We are not aware of any similar applications of

346 these human speech processing algorithms to animal calls, and our success suggests that these
347 methods may be useful for other species. Though the call type has historically been viewed as
348 the basic unit of killer whale stereotyped vocal production, we used a classification test to
349 explore whether the vocal repertoire could be parsed and represented just as accurately using
350 smaller subunits. Compared to the approach of categorizing whole call types, a few pieces of
351 evidence from this study support the notion that killer whales may use an inventory of call
352 subunits to assemble at least some of their calls. In particular, 1) similar classification was
353 achieved when the sorting was based on either the whole call type or the inventory of subunits,
354 and the set of shared subunits may well be superior because it is simpler, requiring less
355 information, computation and memory load (see below), 2) nearly 75% of all stereotyped calls
356 contained at least one subunit shared across calls, and 3) the set of stereotyped shared subunits
357 provided reasonable matches for many of the variable calls. Each of these points is considered in
358 turn.

359

360 The whole contour experiment (WCE) and shared subunit experiment (SSE) suggested
361 that a killer whale could construct its entire stereotyped vocal repertoire either by storing a large
362 number of whole call types or by rearranging a smaller set of shared subunits, respectively.
363 Since the subunit set tripled in size when the subunits were not shared across calls, one may
364 conclude that the unshared subunit-based representation was less likely from the perspective of
365 reducing memory requirements. The SSE demands only a third of the number of subunits used
366 by the unshared subunit experiment (USE), permitting a more condensed and efficient
367 representation. The success of the SSE approach supports the viewpoint that killer whale calls
368 can be segmented into fewer and simpler vocal units, which can generate the same repertoire

369 defined by the larger and more complex set of whole call type contours used in the WCE. The
370 reduction in memory load afforded by the smaller shared subunit count may be offset by the need
371 to retain the rules used to generate and decode the repertoire from these subunits.

372
373 The 11 shared LFC and HFC subunits identified here generated a variety of different call
374 types (Figure 6). Most call types were built from at least one of the shared subunits, supporting
375 the notion that many of the calls were constructed from a set of common subunits. One basic
376 pattern witnessed here indicated that the LFCs could be formed by linking longer strings of
377 subunits together successively, whereas the dominant HFC subunit combination was conserved
378 across multiple call types (see Figure 1). The data presented here suggest a rule-based system in
379 which subunits are only arranged in certain orders and combinations, but further work is required
380 to test the patterns by which killer whales recombine subunits to form calls.

381
382 The results suggested a system in which new call types could be generated by
383 concatenating additional subunits and interspersing them with periods of silence or stringing
384 them together as a continuous vocalisation. Such a system could flexibly yield the size and kind
385 of repertoire produced by these animals. In addition, new call types could be fashioned from
386 existing call types simply by adding, deleting or reordering subunits. This study does not offer
387 proof, however, that the killer whales were actually creating their calls in this manner. One
388 promising research direction in captivity would involve training killer whales to synthesise novel
389 calls by serially producing components heard from a loudspeaker and/or to decompose playback
390 calls by producing the set of constituent subunits with longer gaps of silence than usual. This
391 could offer important supporting evidence that killer whales are capable of generating and

392 deconstructing signals by combining and isolating subunits, respectively.

393

394 Nearly half of the variable calls matched a stereotyped subunit with a likelihood that
395 rivalled at least 10% of the correct-match likelihood values (Table 1, Figure 5). In other words,
396 many of the variable calls, which generally have been considered to be distinct from stereotyped
397 calls, sorted into the subunit categories generated from the stereotyped repertoire. This suggests
398 that variable calls may differ less from stereotyped calls than previously thought and that at least
399 some variable calls may represent different arrangements of the same subunits found in
400 stereotyped calls.

401

402 The prospect that killer whales build their calls from smaller subunits is reinforced by the
403 observation that compound calls can be constructed from whole stereotyped calls (Strager 1993,
404 1995; Shapiro 2008). This suggests a nested system of vocal production in which similar rules
405 of flexible sequencing assemble subunits into call types, which can then be assembled into
406 compound calls. A large portion of the killer whale vocal repertoire can be defined by a system
407 that flexibly generates new call types from a few dozen subunits but employs only a subset of the
408 possible combination of these subunits. This kind of vocal structure of smaller subunits building
409 the repertoire is consistent with analyses of birdsong (Glaze & Troyer 2007), analyses conducted
410 on Alaskan resident and transient killer whale stereotyped calls (Yurk 2005), and analyses of
411 humpback whale (*Megaptera novaeangliae*) song (Payne & McVay 1971; Payne et al. 1984;
412 Suzuki et al. 2006).

413

414 A portion of the stereotyped calls from British Columbia resident and transient killer

415 whale matched successfully with the Norwegian subunits identified here (Table 1, Figure 5).
416 There are two explanations for this result. First, it is to be expected that a certain proportion of
417 signals will overlap between populations by chance because the sound production apparatus of
418 killer whales operates within a finite temporal and spectral range. Second, similar to humans,
419 each population of killer whales may use a portion of a common universal inventory of subunits,
420 stringing these together in specific sequences to form the calls in its own vocal repertoire. Such
421 an explanation is consistent with the possibility of a repertoire of subunits shared across
422 populations. The lower success rates for matching the stereotyped British Columbia calls to
423 stereotyped Norwegian calls compared with matching the variable Norwegian calls to
424 stereotyped Norwegian calls (Table 1) do indicate important divergent properties between the
425 populations that need to be considered.

426

427 Supposing that these animals have separate production control over each individual
428 subunit, further work should test the rates and kinds of subunit modification that occur over time.
429 If shared subunits change (see Ford 1991 for an early discussion of the issue and Deecke et al.,
430 2000 for a demonstration; Miller & Bain 2000; Yurk 2005) similarly across call types, this would
431 support the view that calls are composed of discrete subunits whose acoustic features may slowly
432 alter independent of the call of which they are a part. This idea is analogous to the manner in
433 human language in which the drift in production of certain vowels across words can lead to
434 regional dialects and accents.

435

436 This paper suggests that killer whales may assemble their calls from subunits that appear
437 to be shared across calls and possibly across populations. This suggestion has implications both

438 for the auditory processing of call categorization and for the motor systems that produce vocal
439 behaviour. Most previous research on killer whale calls has treated the call as the basic unit and
440 has focused on an interpretation that each whale learns the distinctive repertoire of calls for its
441 group (e.g., Ford 1987, 1989; Strager 1993; Filatova et al. 2004). An alternative interpretation
442 developed by our results suggests that all killer whales may share a repertoire of subunits, and
443 that each individual learns to create calls from different combinations of a subset of these
444 subunits.

445

446 This view suggests that killer whales may devote auditory categorization processes to
447 detecting a common set of subunits before detecting and categorising the entire call as a unit. On
448 the motor side, this view predicts that killer whales might develop pattern generators for subunits
449 and generate calls by organizing sequences of subunits. Several kinds of evidence from
450 songbirds suggest that syllables are a basic unit of song production. For example, song syllables
451 are produced with a single expiratory pulse of air, and most gaps correspond to short inhalations
452 (Wild et al. 1998; Franz and Goller 2002). Glaze & Troyer (2006) showed that the timing of
453 syllables is more stable than that of the gaps. Cynx (1990) and Franz & Goller (2002)
454 demonstrated that when a songbird is disturbed by a flash of light, it completes the subunit
455 before silencing. All of these studies (which could be conducted in killer whales) suggest that
456 subunits are produced by relatively modular motor programs. Hahnloser et al (2002), Fee et al.
457 (2004), and Glaze & Troyer (2007) suggest that the fundamental unit of birdsong may be even
458 smaller than syllables, with song timed by 5-10ms bursts of neural activity that act like a
459 timekeeper for the sequence of sounds comprising song.

460

461 Compared with the call-oriented view, the subunit-oriented view presented here suggests
462 a different perspective on vocal production learning in killer whales. Future work on vocal
463 development in captivity could examine whether the vocal subunits emerge earlier than the
464 complete call types in a manner similar to language development in humans (see Kuhl 2000).
465 Ultimately, the ideas presented here will have to be tested at the neurobiological level to
466 determine whether killer whales construct their calls through sequences of motor programs that
467 generate subunits, and whether they perceive their calls by categorizing sequences of subunits.

468

469 ACKNOWLEDGEMENTS

470

471 Special thanks to Ghinwa Choueiter for her thoughtful and thorough revision of this
472 manuscript. Chao Wang generously allowed the use of her pitch-tracking algorithm on the killer
473 whale calls, supplying her time and technical support to optimise its performance. Andy Solow
474 offered comments on the evaluation component of the analysis. The analysis benefited
475 enormously from the dedicated hours of correcting of contour traces and discerning of
476 overlapping calls provided by Sara Kim, Gary Matthias, Rebecca McGowan, Levi Schmidt,
477 Katie Smyth, Clara Stefanov-Wagner, Marlene Brito, Deirdre Connolly, Maitagorri Schade and
478 Ivan Dimitrov. Diane Rak, Katie Stanchak and Jia Xin Wu assisted with call type classification.
479 Dr. Robert Gross offered a detailed explanation and programming help for fitting the Legendre
480 polynomials to the contour traces. Ghinwa Choueiter and Scott Cyphers helped with accessing
481 and programming the CSAIL SLS computers. The comments of several anonymous reviewers
482 strengthened the quality and clarity of this manuscript. The undergraduate students were
483 supported by the Massachusetts Institute of Technology Undergraduate Research Opportunities

484 Program office and the Ocean Life Institute (OLI) at the Woods Hole Oceanographic Institution
485 (WHOI). Field work was financed by the OLI, National Geographic Society and WWF Sweden.
486 A. D. Shapiro was funded by a National Defense Science and Engineering Graduate Fellowship
487 and the WHOI Academic Programs Office.
488

489 REFERENCES

- 490 **Adam, O.** 2008. Segmentation of killer whale vocalizations using the Hilbert-Huang transform.
491 *EURASIP Journal on Advances in Signal Processing*, 2008, 1-11.
- 492 **Bain, D. E.** 1986. Acoustic behavior of *Orcinus*: sequences, periodicity, behavioral correlates
493 and an automated method for call classification. In: *Behavioral Biology of Killer Whales* (Ed. by
494 Kirkeveld, B. C. & Lockard, J. S.), pp. 335-371. New York: A. R. Liss.
- 495 **Bonafonte, A., Nogueiras, A. & Rodriguez-Garrido, A.** 1996. Explicit segmentation of speech
496 using Gaussian models. In: *Proceedings of the International Conference on Spoken Language*
497 *Processing*, pp. 1269-1272.
- 498 **Bowles, A. E., Young, W. G. & Asper, E. D.** 1988. Ontogeny of stereotyped calling of a killer
499 whale calf, *Orcinus orca*, during her first year. *Rit Fiskideildar*, 11, 251-275.
- 500 **Chen, S.-H. & Wang, Y.-R.** 1990. Vector quantization of pitch information in Mandarin speech.
501 *IEEE Transactions on Communications*, 38, 1317-1320.
- 502 **Cynx, J.** 1990. Experimental determination of a unit of song production in the zebra finch
503 (*Taeniopygia guttata*). *Journal of Comparative Psychology*, 104, 3-10.
- 504 **Deecke, V. B.** 2003. The vocal behaviour of transient killer whales (*Orcinus orca*):
505 Communicating with costly calls. Ph.D. thesis, University of St. Andrews.
- 506 **Deecke, V. B., Ford, J. K. B. & Slater, P. J. B.** 2005. The vocal behaviour of mammal-eating
507 killer whales: communicating with costly calls. *Animal Behaviour*, 69, 395-405.
- 508 **Deecke, V. B., Ford, J. K. B. & Spong, P.** 1999. Quantifying complex patterns of bioacoustic
509 variation: Use of a neural network to compare killer whale (*Orcinus orca*) dialects. *Journal of*
510 *the Acoustical Society of America*, 105, 2499-2507.
- 511 **Deecke, V. B. & Janik, V. M.** 2006. Automated categorization of bioacoustic signals: Avoiding
512 perceptual pitfalls. *Journal of the Acoustical Society of America*, 119, 645-653.
- 513 **Eales, L. A.** 1985. Song learning in zebra finches: Some effects of song model availability on
514 what is learnt and when. *Animal Behaviour*, 33, 1293-1300.
- 515 **Eens, M., Pinxten, R. & Verheyen, R. F.** 1989. Temporal and sequential organisation of song
516 bouts in the starling. *Ardea*, 77, 75-86.

- 517 **Fee, M. S., Kozhevnikov, A. A. & Hahnloser, R. H. R.** 2004. Neural mechanisms of vocal
518 sequence generation in the songbird. *Annals of the New York Academy of Sciences*, 1016, 153-
519 170.
- 520 **Filatova, O. A., Burdin, A. M., Hoyt, E. & Sato, H.** 2004. A catalogue of discrete calls of
521 resident killer whales (*Orcinus orca*) from the Avacha Gulf of Kamchatka Peninsula.
522 *Zoologicheskyy Journal*, 83, 1169-1180.
- 523 **Foote, A. D., Griffin, R. M., Howitt, D., Larsson, L., Miller, P. J. O. & Hoelzel, A. R.** 2006.
524 Killer whales are capable of vocal learning. *Biology Letters*, 2, 509-512.
- 525 **Ford, J. K. B.** 1987. A catalogue of underwater calls produced by killer whales (*Orcinus orca*)
526 in British Columbia. *Canadian Data Report of Fisheries and Aquatic Sciences*, 633, 1-165.
- 527 **Ford, J. K. B.** 1989. Acoustic behavior of resident killer whales (*Orcinus orca*) off Vancouver
528 Island, British Columbia. *Canadian Journal of Zoology*, 67, 727-745.
- 529 **Ford, J. K. B.** 1991. Vocal traditions among resident killer whales (*Orcinus orca*) in coastal
530 waters of British Columbia. *Canadian Journal of Zoology*, 69, 1454-1483.
- 531 **Franz, M. & Goller F.** 2002. Respiratory units of motor production and song imitation in the
532 zebra finch. *Journal of Neurobiology*, 51, 129-141.
- 533 **Furui, S.** 1986. On the role of spectral transition for speech perception. *Journal of the Acoustical*
534 *Society of America*, 80, 1016-1025.
- 535 **Glaze, C. M. & Troyer, T. W.** 2006. Temporal structure in zebra finch song: implications for
536 motor coding. *Journal of Neuroscience*, 26, 991-1005.
- 537 **Glaze, C. M. & Troyer, T. W.** 2007. Behavioral measurements of a temporally precise motor
538 code for birdsong. *Journal of Neuroscience*, 27, 7631-7639.
- 539 **Hahnloser, R. H. R., Kozhevnikov, A. A., Fee, M. S.** 2002. An ultra-sparse code underlies the
540 generation of neural sequences in a songbird. *Nature*, 419, 65-70.
- 541 **Hauser, M. D., Chomsky, N. & Fitch, W. T.** 2002. The faculty of language: What is it, who
542 has it, and how did it evolve? *Science*, 298, 1569-1579.
- 543 **Hoelzel, A. R. & Osborne, R. W.** 1986. Killer whale call characteristics: Implications for
544 cooperative foraging strategies. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevold, B. C.
545 & Lockard, J. S.), pp. 373-403. New York: Alan R. Liss, Inc.

- 546 **Hultsch, H. & Todt, D.** 1989. Memorization and reproduction of songs in nightingales (*Luscinia*
547 *megarhynchos*): Evidence for package formation. *Journal of Comparative Physiology A*, 165,
548 197-203.
- 549 **Janik, V. M. & Slater, P. J. B.** 1997. Vocal learning in mammals. *Advances in the Study of*
550 *Behavior*, 26, 59-99.
- 551 **Janik, V. M. & Slater, P. J. B.** 2000. The different roles of social learning in vocal
552 communication. *Animal Behaviour*, 60, 1-11.
- 553 **Johnson, M. P. & Tyack, P. L.** 2003. A digital acoustic recording tag for measuring the
554 response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering*, 28, 3-12.
- 555 **Kuhl, P. K.** 2000. A new view of language acquisition. *Proceedings of the National Academy of*
556 *Sciences*, 97, 11850–11857.
- 557 **Lemon, R. E. & Chatfield, C.** 1971. Organization of song in cardinals. *Animal Behaviour*, 19,
558 1-17.
- 559 **Miller, P. J. & Tyack, P. L.** 1998. A small towed beamforming array to identify vocalizing
560 resident killer whales (*Orcinus orca*) concurrent with focal behavioral observations. *Deep Sea*
561 *Research II*, 45, 1389-1405.
- 562 **Miller, P. J. O. & Bain, D. E.** 2000. Within-pod variation in the sound production of a pod of
563 killer whales, *Orcinus orca*. *Animal Behaviour*, 60, 617-628.
- 564 **Miller, P. J. O., Samarra, F. I. P. & Perthuison, A. D.** 2007. Caller sex and orientation
565 influence spectral characteristics of "two-voice" stereotyped calls produced by free-ranging killer
566 whales. *Journal of the Acoustical Society of America*, 121, 3932-3937.
- 567 **Nousek, A. E., Slater, P. J. B., Wang, C. & Miller, P. J. O.** 2006. The influence of social
568 affiliation on individual vocal signatures of northern resident killer whales (*Orcinus orca*).
569 *Biology Letters*, 2, 481-484.
- 570 **Nowak, M. A., Plotkin, J. B. & Jansen, V. A. A.** 2000. The evolution of syntactic
571 communication. *Nature*, 404, 495-498.
- 572 **Payne, K., Tyack, P. & Payne, R.** 1984. Progressive changes in the songs of humpback whales
573 (*Megaptera novaeangliae*): A detailed analysis of two seasons in Hawaii. In: *Communication*
574 *and Behavior of Whales* (Ed. by Payne, R.), pp. 9-57. Boulder, Colorado: Westview Press, Inc.
- 575 **Payne, R. S. & McVay, S.** 1971. Songs of humpback whales. *Science*, 173, 585-597.

- 576 **Rehn, N., Teichert, S. & Thomsen, F.** 2007. Structural and temporal emission patterns of
577 variable pulsed calls in free-ranging killer whales (*Orcinus orca*). *Behaviour*, 144, 307-329.
- 578 **Riebel, K. & Todt, D.** 1997. Light flash stimulation alters the nightingale's singing style:
579 Implications for song control mechanisms. *Behaviour*, 134, 789-808.
- 580 **Riesch, R., Ford, J. K. B. & Thomsen, F.** 2006. Stability and group specificity of stereotyped
581 whistles in resident killer whales, *Orcinus orca*, off British Columbia. *Animal Behaviour*: 71, 79-
582 91.
- 583 **Seneff, S.** 1978. Real-time harmonic pitch detector. *IEEE Transactions on Acoustics, Speech and*
584 *Signal Processing*, 26, 358-365.
- 585 **Shapiro, A. D.** 2008. Orchestration: The movement and vocal behavior of free-ranging
586 Norwegian killer whales (*Orcinus orca*). Ph.D. thesis, MIT/WHOI.
- 587 **Shapiro, A. D. & Wang, C.** 2009. A versatile pitch tracking algorithm: From human speech to
588 killer whale vocalizations. *Journal of the Acoustical Society of America*, 126, 451-459.
- 589 **Strager, H.** 1993. Catalogue of underwater calls from killer whales (*Orcinus orca*) in northern
590 Norway. M.Sc. thesis, University of Århus.
- 591 **Strager, H.** 1995. Pod-specific call repertoires and compound calls of killer whales, *Orcinus*
592 *orca* Linnaeus, 1758, in the waters of northern Norway. *Canadian Journal of Zoology*, 73, 1037-
593 1047.
- 594 **Suzuki, R., Buck, J. R. & Tyack, P. L.** 2006. Information entropy of humpback whale songs.
595 *Journal of the Acoustical Society of America*, 119, 1849-1866.
- 596 **Van Opzeeland, I. C., Corkeron, P. J., Leyssen, T., Similä, T. & Van Parijs, S. M.** 2005.
597 Acoustic behaviour of Norwegian killer whales, *Orcinus orca*, during carousel and seiner
598 foraging on spring-spawning herring. *Aquatic Mammals*, 31, 110-119.
- 599 **Wang, C.** 2001. Prosodic Modeling for Improved Speech Recognition and Understanding. Ph.D.
600 thesis, MIT.
- 601 **Wild, J. M., Goller, F., & Suthers, R. A.** 1998. Inspiratory muscle activity during bird song.
602 *Journal of Neurobiology*, 36, 441-453.
- 603 **Williams, H. & Staples, K.** 1992. Syllable chunking in zebra finch (*Taeniopygia guttata*) song.
604 *Journal of Comparative Psychology*, 106, 278-286.

- 605 **Yurk, H.** 2005. Vocal culture and social stability in resident killer whales (*Orcinus orca*). Ph.D.
606 thesis, University of British Columbia.
- 607 **Yurk, H., Barrett-Lennard, L., Ford, J. K. B. & Matkin, C. O.** 2002. Cultural transmission
608 within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Animal*
609 *Behaviour*, 63, 1103-1119.
- 610 **Zhu, D., Adda-Decker, M. & Antoine, F.** 2005. Different size multilingual phone inventories
611 and context-dependent acoustic models for language identification. In: *Interspeech*. Lisbon,
612 Portugal.
- 613
- 614

615 FIGURE CAPTIONS

616

617 Figure 1. Examples of the call types motivating this study. The three spectrogram panels on the
618 left are variations of the basic N72 call. N72.2 contains a short additional subunit at the end of
619 the call and N72.3 contains two such additional subunits. The four spectrogram panels on the
620 right can be assembled from top to bottom by sequentially adding both low (LFC) and high
621 frequency components (HFC) to the N16.1 base.

622

623 Figure 2. The abrupt, non-continuous spectral shifts (left: N9.2) and intervals of silence (right:
624 N72.3) of certain call types were considered boundaries defining subunit edges (indicated here
625 with arrowheads).

626

627 Figure 3. Schematic illustration of the three experiments described in the text. The top row
628 depicts the original traces for the low frequency components of 3 different call types: N12.1, a
629 continuous and descending vocalisation; N16.1, a continuous call that was segmented due to its
630 abrupt frequency change (marked by the thin dotted line); N72.2, a call containing two subunits
631 separated by a brief period of silence. The second row plots the 4th-order Legendre polynomial
632 fits used in the first whole contour experiment (WCE) in light grey. Each contour was
633 considered continuous and silent intervals were interpolated over (indicated by the thick dotted
634 line connecting the two N72.2 subunits). The contours were labelled by adding zeros until two
635 places to the right of the decimal point were filled. The unshared subunit experiment (USE) is
636 shown in the third row. Here, the subunits of each call were represented with a polynomial
637 (N16.1 and N72.2 were divided into two subunits using the segmentation algorithm and silent

638 interval, respectively), but were labelled distinctly. The hundredths place in the label was used
639 to count each successive subunit for a particular call type. The final row demonstrates treatment
640 according to the subunit segment experiment (SSE). The segmentation decisions and polynomial
641 fits were the same as in the USE but the labelling allowed call types to share subunits. For
642 example, the second subunits of both N16.1 and N72.2 belonged to shared subunit 4. Subunits
643 that appeared only in a single call type retained their USE label (e.g., 12.11).

644

645 Figure 4. Categorization error rates (i.e., how often the members of the test set were assigned to
646 different call types than human expert judgments) for all experiments. Each experiment was run
647 100 times to compute the mean error rate \pm standard error.

648

649 Figure 5. Sampling of matches of variable (non-stereotyped) calls from the Norwegian
650 population (first column) and stereotyped calls from British Columbia killer whale resident
651 (second column) and transient (third column) populations (dark black traces) with stereotyped
652 calls from the Norwegian population (gray traces). Each sub-panel is titled with the Norwegian
653 stereotyped subunit match. The titles of the British Columbia call panels also contain the call
654 type of the resident or transient call. The number following the hyphen in the transient call label
655 refers to the subunit number. The frequency and time axes for each subpanel vary.

656

657 Figure 6. a: Subunits found in at least two call types. Note different axis scales. b: Colour code
658 for subunit traces from different call types (legend locations in (b) match panel locations in (a)).
659 The numbers refer to the call type (digits up to and including the tenths place) and subunit
660 number (digit in the hundredths place). For example, 8.01 refers to call type N8 and subunit

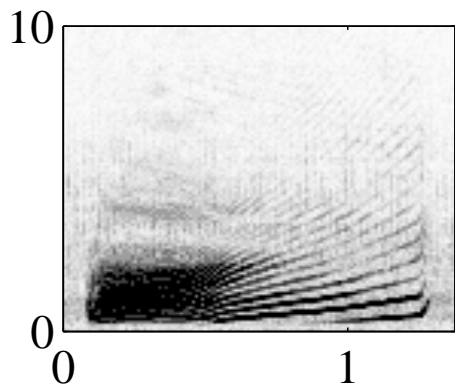
661 number 1 and 72.32 refers to call type N72.3 and subunit number 3.

ACKNOWLEDGEMENTS

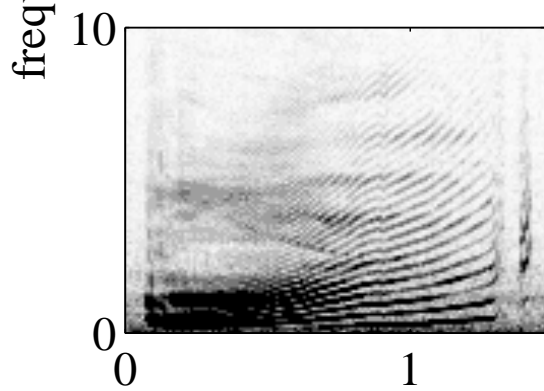
Special thanks to Ghinwa Choueiter for her thoughtful and thorough revision of this manuscript. Chao Wang generously allowed the use of her pitch-tracking algorithm on the killer whale calls, supplying her time and technical support to optimise its performance. Andy Solow offered comments on the evaluation component of the analysis. The analysis benefited enormously from the dedicated hours of correcting of contour traces and discerning of overlapping calls provided by Sara Kim, Gary Matthias, Rebecca McGowan, Levi Schmidt, Katie Smyth, Clara Stefanov-Wagner, Marlene Brito, Deirdre Connolly, Maitagorri Schade and Ivan Dimitrov. Diane Rak, Katie Stanchak and Jia Xin Wu assisted with call type classification. Dr. Robert Gross offered a detailed explanation and programming help for fitting the Legendre polynomials to the contour traces. Ghinwa Choueiter and Scott Cyphers helped with accessing and programming the CSAIL SLS computers. The comments of several anonymous reviewers strengthened the quality and clarity of this manuscript. The undergraduate students were supported by the Massachusetts Institute of Technology Undergraduate Research Opportunities Program office and the Ocean Life Institute (OLI) at the Woods Hole Oceanographic Institution (WHOI). Field work was financed by the OLI, National Geographic Society and WWF Sweden. A. D. Shapiro was funded by a National Defense Science and Engineering Graduate Fellowship and the WHOI Academic Programs Office.

Figure 1a

N72



N72.2



N72.3

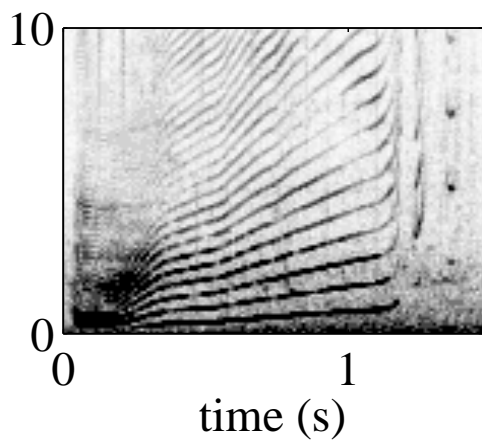
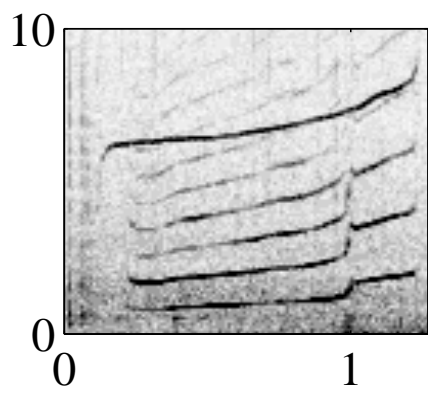
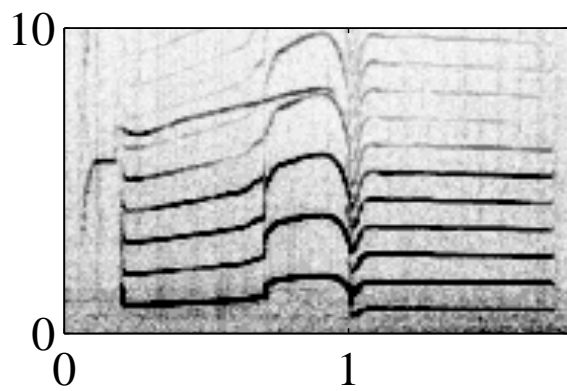


Figure 1b

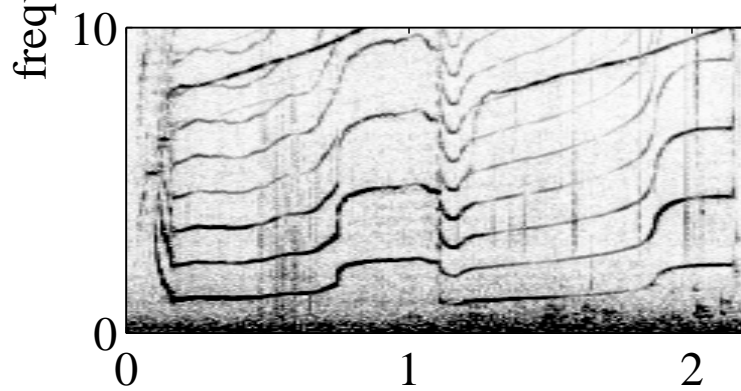
N16.1



N91



N91.2



N91.3

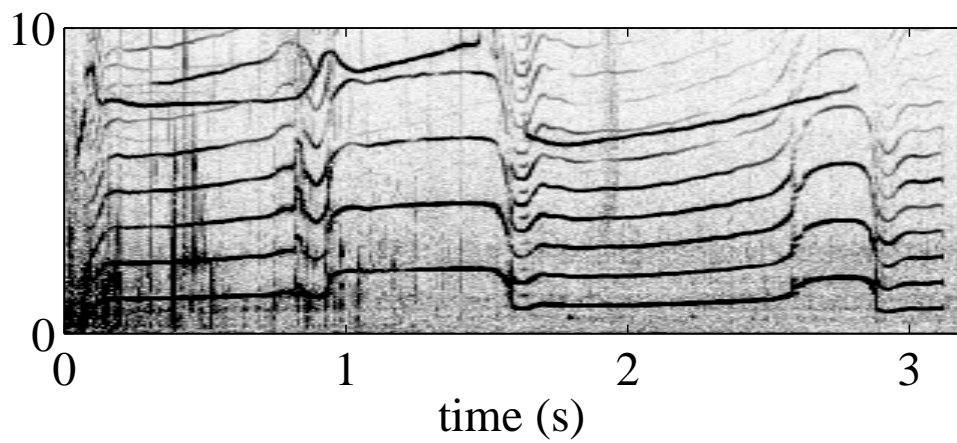


Figure 2

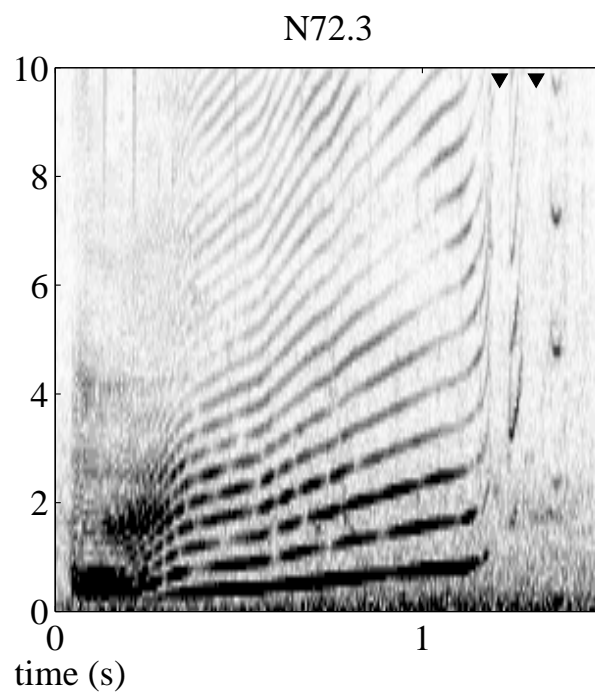
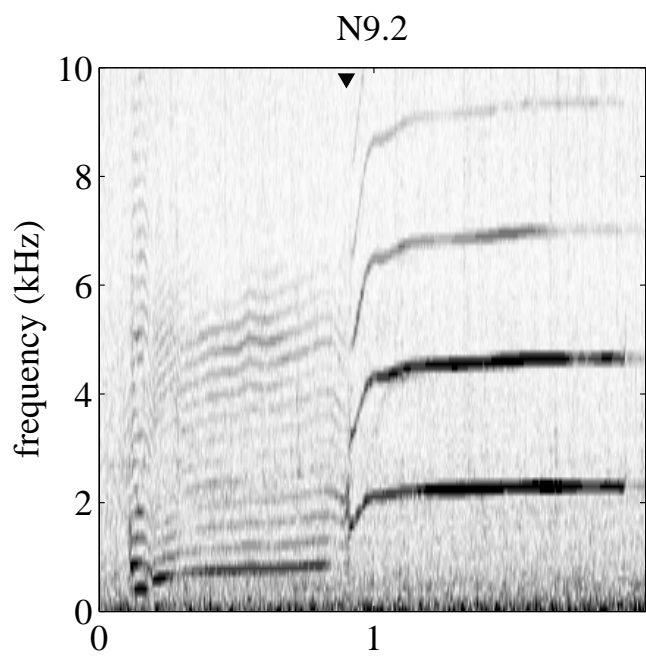


Figure 3

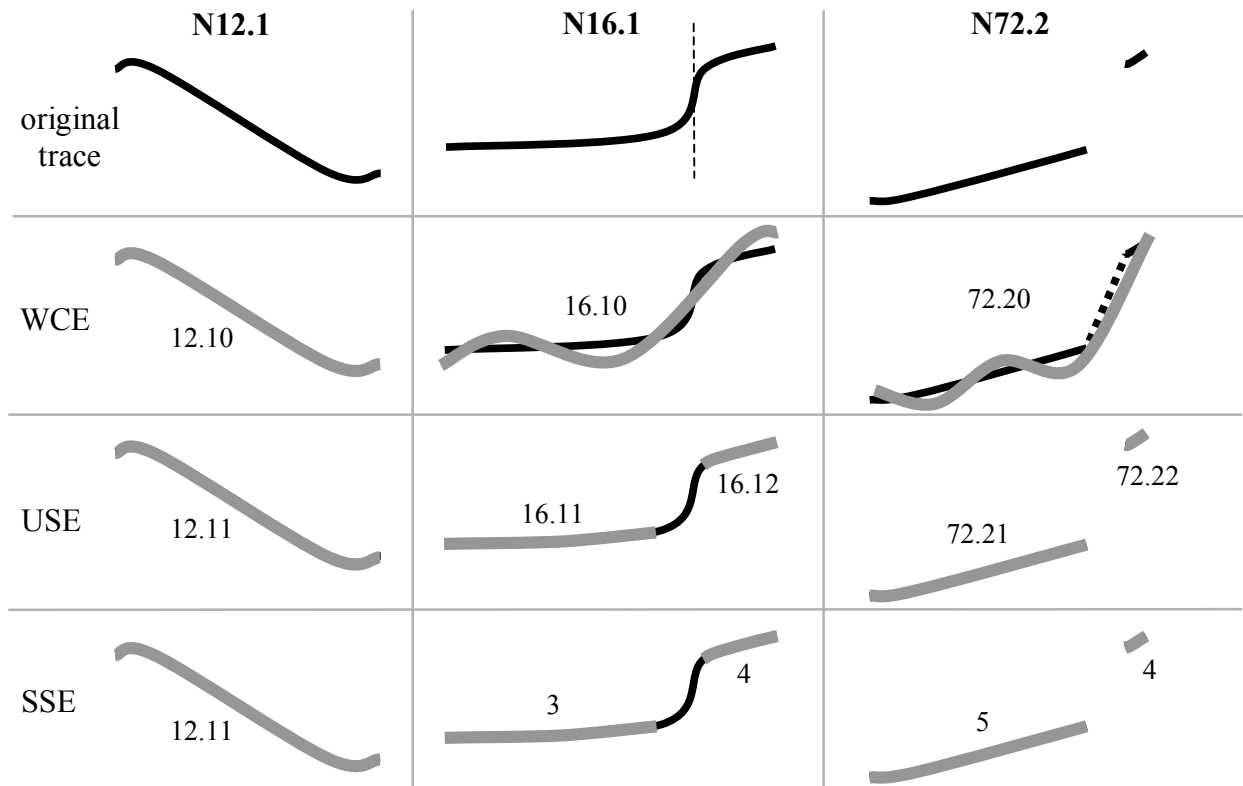


Figure 4

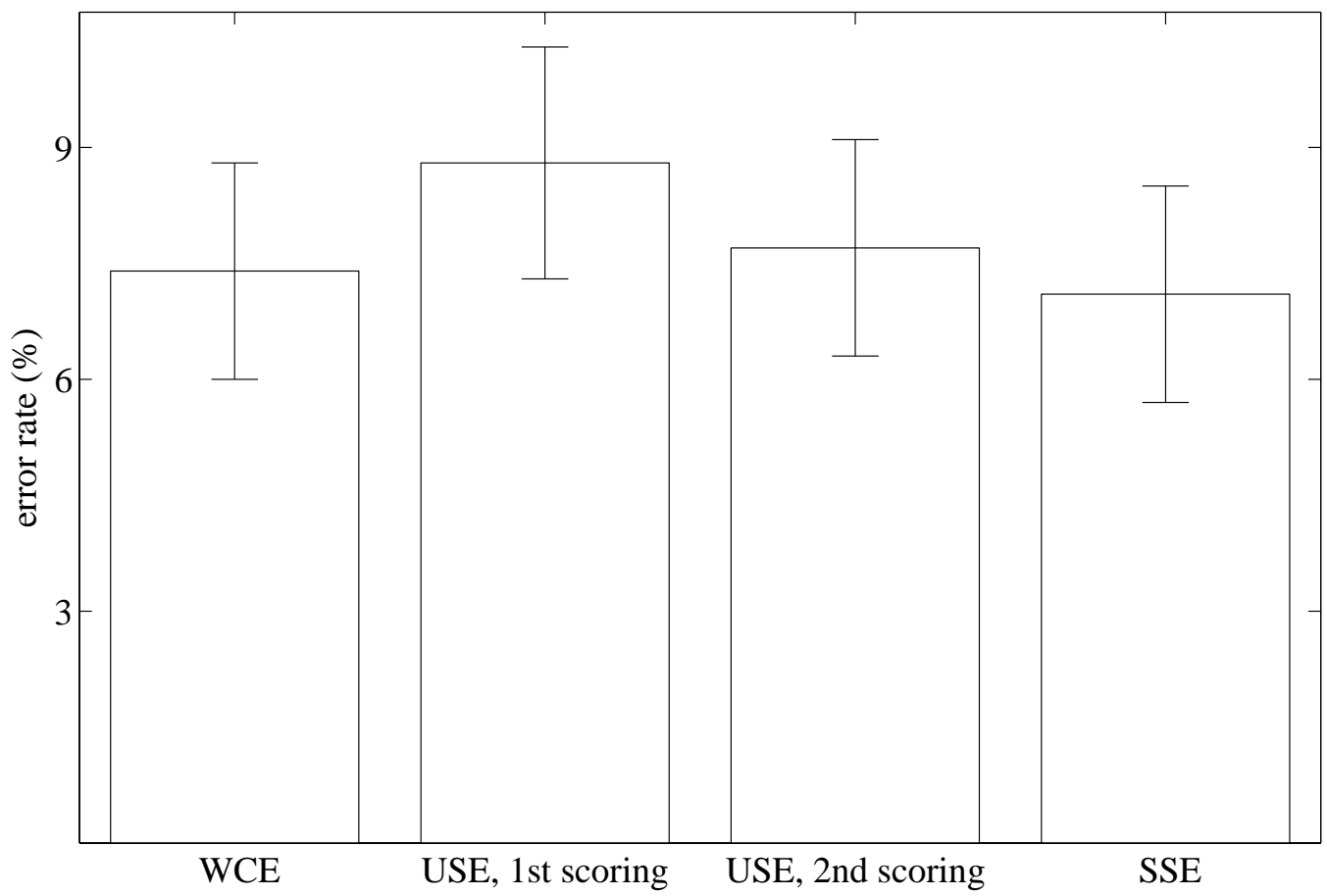


Figure 5

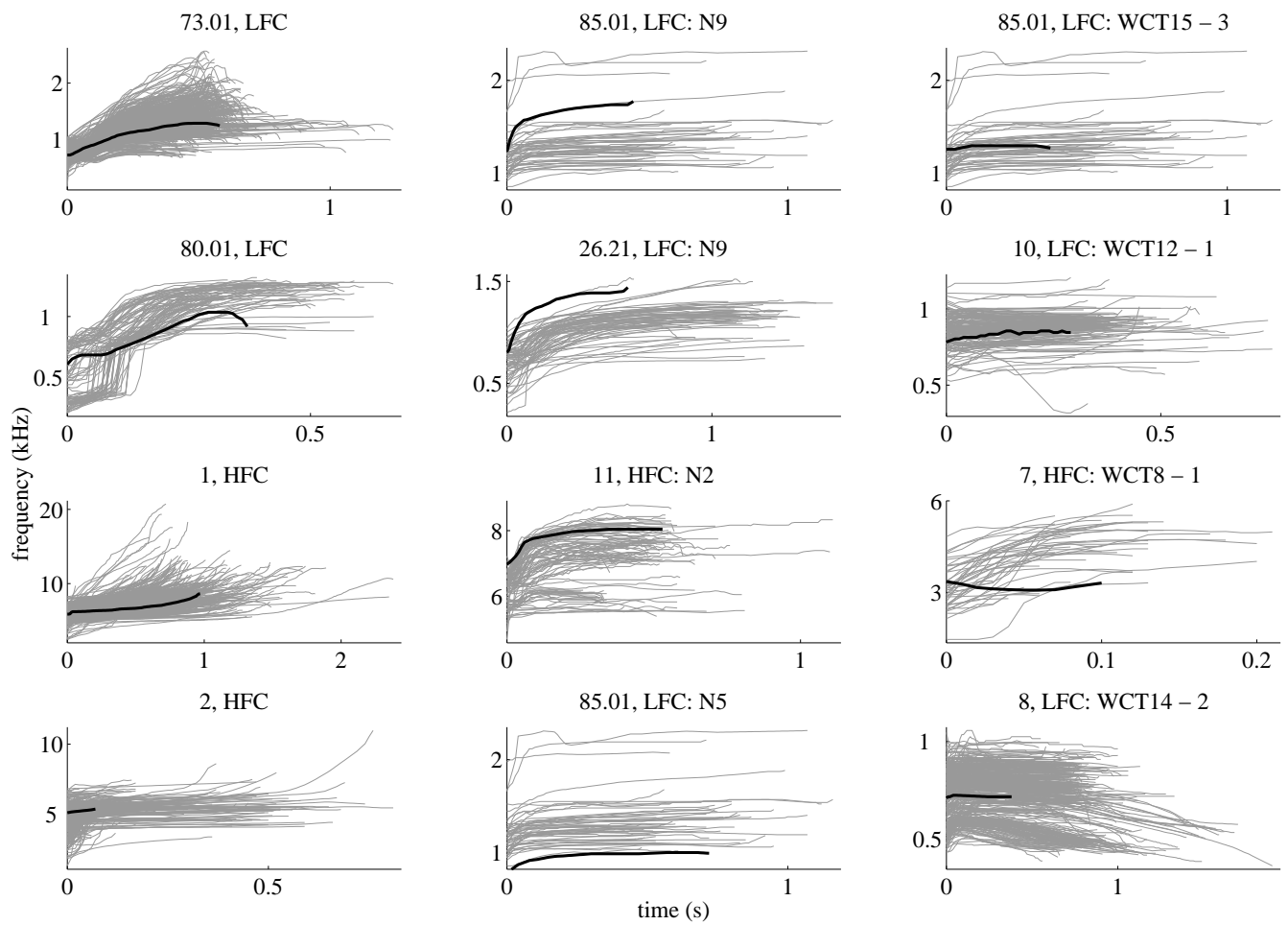


Figure 6a

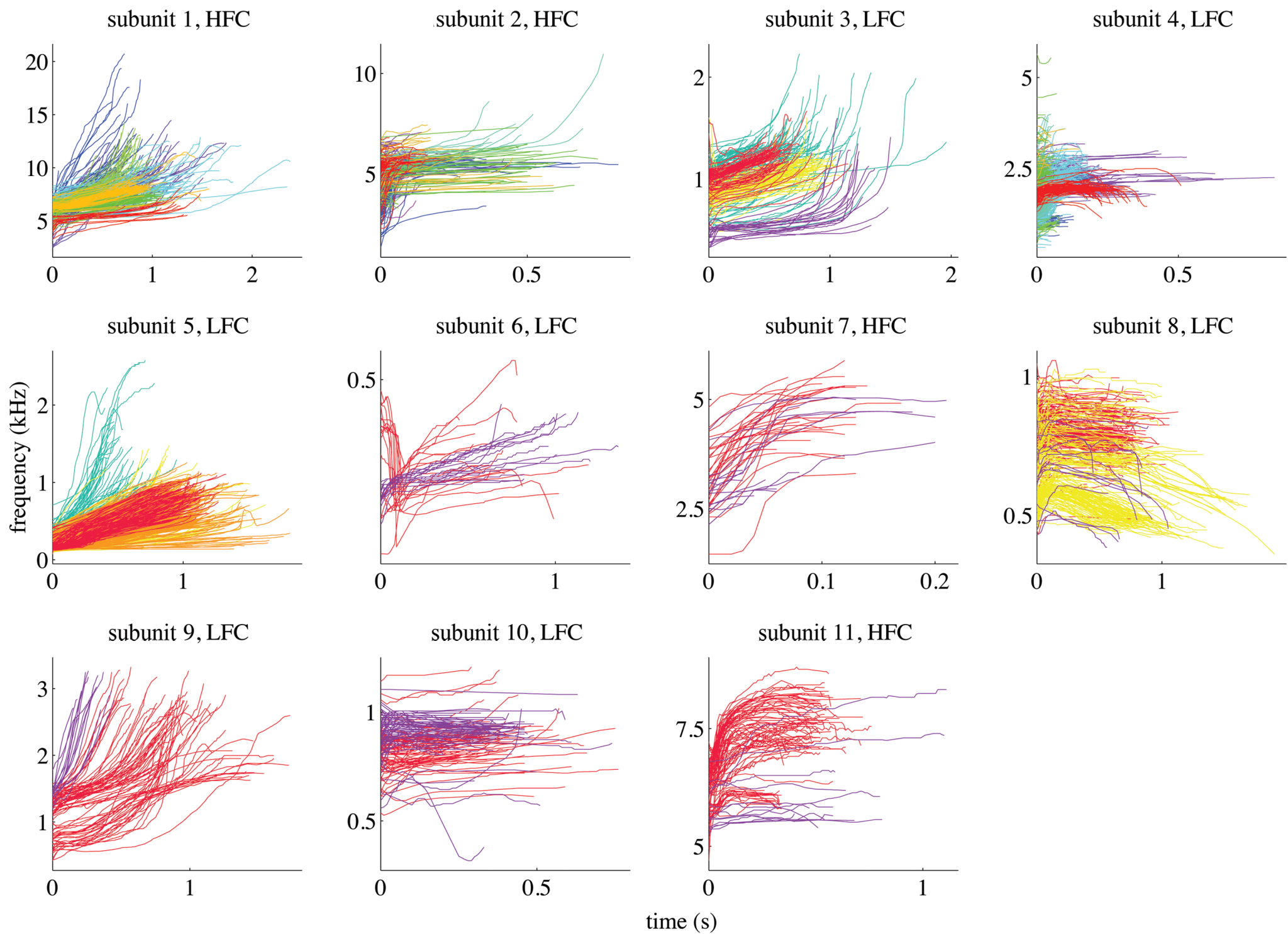
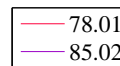
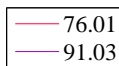
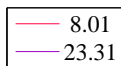
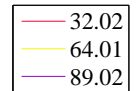
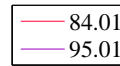
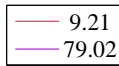
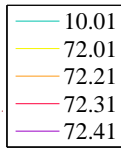
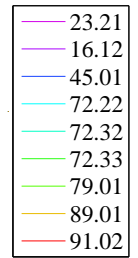
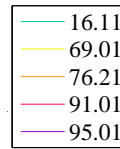
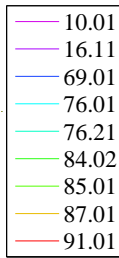
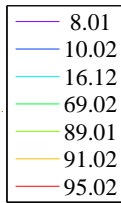


Figure 6b



Publicity Photograph (Optional)
[Click here to download high resolution image](#)



Lay summary for manuscript “Comparing call-based vs. subunit-based methods for categorizing Norwegian killer whale (*Orcinus orca*) vocalisations” by Shapiro, Tyack & Seneff.

Words are the fundamental unit of meaning for human speech, but are made up of shorter, shared acoustic elements called phonemes. However, many studies of animal communication treat each call as an independent unit. Here we test whether killer whale calls can be represented as effectively by sequences of subunits as by the entire call. We compared three models for how these calls might be represented: (1) whole calls, (2) assembled from subunits unique to each call, or (3) subunits shared across calls. Performance of the models was equivalent. Still, calls composed of shared subunits may provide a simpler approach to understand killer whale vocal behaviour since (1) there were fewer subunits than call types, (2) nearly 75% of all call types shared at least one subunit. Subunits derived from Norwegian stereotyped calls were also detected in variable calls from Norway, and even from two Canadian populations of killer whales.

TABLE 1. How well do the Norwegian variable calls and the stereotyped calls of British Columbia (BC) resident and transient killer whales (i.e., auxiliary calls) match to call subunits from the Norwegian population? For each bold heading, the first and second columns contain the percent and count of auxiliary calls, respectively, that met various correct-match thresholds (see Methods).

Percentage of auxiliary calls matching Norwegian stereotyped subunits at least as well as those subunits matched themselves	Norwegian variable calls		BC resident stereotyped calls		BC transient stereotyped calls	
	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>
5	72.4	489	30.7	59	56.0	116
10	53.5	361	20.3	39	39.6	82
25	26.2	177	3.6	7	11.6	24
50	10.8	73	0	0	2.9	6
75	4.0	27	0	0	0.0	0
90	1.9	13	0	0	0.0	0
95	1.1	7	0	0	0.0	0

TABLE 2. Counts of all call types (first row), call types containing both a LFC and HFC (second row), a LFC alone (third row) and a HFC alone (fourth row) that were comprised of call-specific subunits only (first column), subunits shared across different call types (second column), and a mixture of call-specific and shared subunits (third column).

	call types formed from call-specific subunits	call types formed from subunits shared across different call types	call types formed from a mixture of call-specific and shared subunits
all call types	8	18	5
LFC & HFC	2	8	2
LFC	6	7	3
HFC	0	3	0