

1 **Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*)**

2 José Z. Abramson<sup>1,2,3</sup>, M<sup>a</sup> Victoria Hernández-Lloreda<sup>3,4</sup>, Lino García<sup>5</sup>, Fernando Colmenares<sup>3,6</sup>, Francisco  
3 Aboitiz<sup>1</sup> & Josep Call<sup>7,8</sup>

4 <sup>1</sup>Departamento de Psiquiatría, Facultad de Medicina and Centro Interdisciplinario de Neurociencias,  
5 Pontificia Universidad Católica de Chile, Marcoleta 391, Santiago, Chile <sup>2</sup>Facultad de Ciencias, Escuela de  
6 Medicina Veterinaria, Universidad Mayor, Santiago de Chile <sup>3</sup>Grupo UCM de Psicobiología Social,  
7 Evolutiva y Comparada, Universidad Complutense de Madrid, 28223, Madrid, Spain <sup>4</sup>Departamento de  
8 Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Campus de Somosaguas,  
9 Universidad Complutense de Madrid, 28223, Madrid, Spain <sup>5</sup>Dpto. Teoría de la Señal y Comunicaciones  
10 ETSIST, UPM, Spain <sup>6</sup>Departamento de Psicobiología, Facultad de Psicología, Campus de Somosaguas,  
11 Universidad Complutense de Madrid, 28223, Madrid, Spain <sup>7</sup>School of Psychology and Neuroscience,  
12 University of St Andrews, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP, United Kingdom  
13 <sup>8</sup>Department of Developmental and Comparative Psychology, Max-Planck Institute for Evolutionary  
14 Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

15 **Abstract**

16 Vocal imitation is a hallmark of human spoken language, which, along with other advanced  
17 cognitive skills, has fuelled the evolution of human culture. Comparative evidence has revealed that  
18 although the ability to copy sounds from conspecifics is mostly uniquely human among primates, a  
19 few distantly related taxa of birds and mammals have also independently evolved this capacity.  
20 Remarkably, field observations of killer whales have documented the existence of group-  
21 differentiated vocal dialects that are often referred to as traditions or cultures and are hypothesized  
22 to be acquired non-genetically. Here we use a *-Do as I do-* paradigm to study the abilities of a killer  
23 whale to imitate novel sounds uttered by conspecific (vocal imitative learning) and human models  
24 (vocal mimicry). We found that the subject made recognizable copies of all familiar and novel  
25 conspecific and human sounds tested and did so relatively quickly (most during the first 10 trials  
26 and three in the first attempt). Our results lend support to the hypothesis that the vocal variants

27 observed in natural populations of this species can be socially learned by imitation. The capacity for  
28 vocal imitation shown in this study may scaffold the natural vocal traditions of killer whales in the  
29 wild.

30 Keywords: Vocal learning; imitation; mimicry; cetacean culture; do as I do; killer whale

31

## 32 **1. Introduction**

33 Learning a previously unknown behaviour by observation from another individual [1] enables the  
34 non-genetic transfer of information between individuals and constitutes a potential driver for the  
35 diffusion and consolidation of group-specific behavioural phenotypes (i.e., traditions and cultures)  
36 [2,3]. Imitation of novel sounds, also referred to as vocal production learning [4] and defined as the  
37 ability to learn to produce a novel sound just from hearing it, is a core property of human speech  
38 which, along with other cognitive skills, has fuelled the evolution of another adaptation uniquely  
39 evolved in our species, human culture [5]. Comparative evidence has revealed that although the  
40 ability to copy sounds from conspecifics is widespread in birds, it is strikingly rare in mammals  
41 [4,6], and among primates it is unique to humans [7,8, but see 9]. Cetaceans are one of the few  
42 mammalian taxa shown to have evolved the ability for vocal production learning. Studies of several  
43 cetacean species in the wild have revealed that they exhibit substantial behavioural diversity  
44 between sympatric groups in terms of the acoustic features of their vocal repertoires (songs, calls)  
45 [10]. Bottlenose dolphins (*Tursiops truncatus*) display individual recognition signature whistles,  
46 humpback whales (*Megaptera novaengliae*) produce songs where some elements persist while  
47 others evolve over generations, and sperm whales (*Physeter macrocephalus*) possess group-specific  
48 coda repertoires [4,10,11]. It has been suggested that imitative learning can underpin these  
49 behaviours with experimental evidence for the ability for sound imitation demonstrated mainly in  
50 the bottlenose dolphin [11,12,13] and recently in the beluga (*Delphinapterus leucas*) [14,15].

51 Among cetaceans, the killer whale (*Orcinus orca*) stands out regarding the study of vocal

52 dialects in the wild [16]. Each family or matrilineal unit or pod within a population has been  
53 documented to have a unique vocal dialect, including a combination of unique and shared call types  
54 [17-19]. These dialects are believed to be transmitted via social learning [16-18], not only from  
55 mother to offspring (vertical transmission), but also between matrilines (horizontal transmission)  
56 [18-21]. Moreover, the similar acoustic features found between different populations in the same  
57 area do not correlate with geographic distance [22]. Since many of these group-differentiated  
58 signatures are not accounted for by ecological factors or genetic inheritance, the hypothesis that  
59 they may have been acquired through social learning, particularly imitation, appears plausible [16-  
60 24].

61 Elucidating the precise mechanism of social learning involved is difficult, however, particularly  
62 for acoustic communication in wild populations. Although killer whales have been shown to be  
63 capable of learning novel motor actions from conspecifics through imitation [25], the experimental  
64 evidence for production imitation skills in the vocal domain is still scarce in this species. There are  
65 reports on killer whales in the field and in captive settings indicating that they can copy novel calls  
66 from conspecifics [26,27] and even from heterospecifics such as bottlenose dolphins [28] or sea  
67 lions [24]. One Icelandic female was found to match novel calls from a Northern Resident female  
68 with whom she had been housed together for several years [26]. Two juvenile killer whales,  
69 separated from their natal pods, were observed to mimic the barks of sea lions in a field study [24].  
70 Crance et al. [27] and Musser et al. [28] took advantage of two unplanned cross-socializing  
71 experimental situations to show that two juvenile males learned novel calls from an unrelated but  
72 socially close adult male, and three individuals learned novel whistles from a dolphin, respectively.

73 However, as suggestive as these reports of killer whales copying sounds from other individuals  
74 are, the lack of experimental controls curtails the interpretation about the underlying acquisition  
75 mechanisms. Experimental data are needed to ascertain whether vocal learning is a plausible  
76 mechanism underlying the complexity of vocal traditions in killer whales in the wild. However, to

77 the best of our knowledge, not even anecdotal reports exist about killer whales spontaneously  
78 mimicking human speech similar to those reported in some birds (e.g., parrots [29], mynahs [30])  
79 and mammals (elephants [31] seals [32], belugas [14]).

80 In most mammals, sound production occurs in the vocal folds within the larynx (the sound  
81 source) and the supralaryngeal vocal tract, consisting of pharyngeal, oral, and nasal cavities (the  
82 filter) [33]. In humans, this apparatus increases in complexity due to the unusual neurological and  
83 motor control that we can exert on these structures [33,34]. In contrast, toothed cetaceans (e.g.,  
84 killer whales, belugas and dolphins) have evolved a pneumatic sound production in the nasal  
85 complex passages (instead of the larynx) involving bilateral structures such as a pair of phonic lips,  
86 that can operate as two independent sound sources and filters [35,36]. This difference in the sound  
87 production system between toothed cetaceans and humans make the investigation of cetacean vocal  
88 production particularly valuable for comparative analyses of flexible vocal production.

89 Here we report an experimental study of sound learning and mimicry abilities of a killer whale  
90 listening to familiar or novel sounds uttered by a conspecific or a human model and requested to  
91 reproduce them on command ('Do this!'). The *Do-as-I-do* paradigm [37] involves the copying of  
92 another's untrained (familiar or novel) motor or vocal actions under a specific previously trained  
93 signal in the absence of results-based cues. The Do-as-I-do training method has been successfully  
94 used in studies of primates, birds, dogs, and two species of cetaceans [12, 25, 38]. In fact, this  
95 method was previously used in a study of production imitation of novel motor actions in the group  
96 of killer whales studied here [25]. Ultimately, we wanted to test the long-standing hypothesis that  
97 the group-specific vocal patterns documented in wild populations of killer whales can be learned  
98 socially and, more specifically, through production imitation learning.

## 99 **2. Methods**

### 100 **(a) Subjects**

101 We tested a 14 year-old female killer whale (*Orcinus orca*), named Wikie, housed at Marineland  
102 Aquarium in Antibes, France. The conspecific model was her own 3 years old calf, named Moana,  
103 born in Marineland. Wikie had been trained for a variety of examination and exercise behaviours  
104 with standard operant conditioning procedures and fish/tactile positive reinforcement. Also, she had  
105 participated in a previous experimental study of action imitation [25], so she was already trained  
106 with the ‘copy’ command.

## 107 **(b) Procedure**

108 The study comprised three phases. *Phase 1* involved retraining and reinforcing the subject to  
109 respond to the gesture-based command ‘copy’ (‘Do that!’) given by the trainer, that had been used 4  
110 years earlier in the previous study of action imitation aforementioned [25]. *Phase 2* involved testing  
111 the subject’s response to the trainer’s copy command when the model uttered familiar vocalizations  
112 (n = 3 different sounds), that is, vocalizations that the subject had already performed herself, either  
113 because she had been trained with them or because they were part of her natural repertoire (see  
114 Table 1). Finally, *Phase 3* involved testing the subject with novel sounds (n = 11 different sounds),  
115 that is, sounds that were unknown to the subject in terms of neither having heard them, nor having  
116 been uttered by her previously. To ensure that the unfamiliar sounds (conspecific and humans) were  
117 as different as possible from what they had produced before we compared them with 278 sound  
118 samples extracted from ‘Hodgins’ sound recording baseline of the vocal repertoire in this same  
119 group of killer whales [39], where she had identified up to eleven distinct discrete call types and we  
120 found no matching with our sample of novel conspecific or human sounds. In addition, before  
121 running the experiment we recorded 28 hours of in-air spontaneous sounds produced by the killer  
122 whales during their free time to see if the subject (or any other killer whale in the group) uttered  
123 sounds similar to the novel sounds in our sample. (Further details are given in the ESM.) *Phase 3*  
124 comprised two testing conditions: a conspecific model (condition 1) and a human model (condition  
125 2). In condition 1, the subject first listened to a conspecific model’s performance that included three

126 familiar sounds and five novel sounds (test trials), and then was signalled to copy them. The sounds  
127 were presented in two formats: 1) performed by a killer whale model live and 2) played through a  
128 speaker (e.g. conspecific sounds like airy atonal sounds as ‘Breathy’ and ‘Strong’ raspberries, or  
129 tonal whiny sirens sounds like ‘Wolf’). In condition 2, the subject also listened to three familiar and  
130 six other novel sounds (test trials), but now they were produced by a human model (e.g. human  
131 sounds like a human laugh ‘Ah Ah’ or human words like ‘One Two’ (Table S1 gives the complete  
132 description of each sound). In the two conditions, the sounds were presented with the constraint  
133 that no more than three consecutive test trials of the novel sound could occur in a row. In each  
134 session, a single novel sound was presented to the subject at a time. We also interspersed the three  
135 familiar sounds that had been used in the previous phases and control trials consisting of ‘non-copy’  
136 trials during which the subject’s trainer did not make the copy sign and asked for any other trained  
137 action that the subject regularly was requested to perform during the aquarium shows. Therefore,  
138 sessions consisted of several familiar sounds and control trials and from six to ten test trials of the  
139 novel sound. The subject was positively reinforced with fish and/or tactile and voice reinforcement  
140 signals whenever she yielded a correct response as judged in real time by two observers (Wikie’s  
141 trainer and one experimenter), but only when she was asked to copy familiar sounds or perform  
142 familiar actions (control trials). During the test trials (novel sounds from conspecific and human  
143 models), the subject received no rewards irrespective of whether she responded correctly or not, so  
144 that the experimenter did not provide any cues, thus making real time judgments unnecessary.  
145 Altogether, phase 1 lasted one session, phase 2 lasted seven sessions and phase 3 fifty- two  
146 sessions. All the sounds were asked and performed when the subject’s head was above the water  
147 surface with her blowhole exposed.

148 Three different set-ups were used. (a) *Conspecific live condition*: The two trainers ( $T_M$  and  $T_S$ ;  
149 M for model and S for subject) were positioned on different sides of a wooden panel 2m long x 1.90  
150 cm high placed in a position in which S and M could see each other and their own trainer, but could

151 not see the other trainer's commands.  $T_M$  was positioned on the right side of the panel, and  $T_S$  was  
152 on the left side; thus, the trainers were in a position from which they were not able to see each  
153 other's signals either (see figure 1). (b) *Conspecific speaker condition*: two trainers were also  
154 required, one trainer held the speaker and another ( $T_S$ ) gave the copy command to the subject; and  
155 (c) *Human live condition*; just one trainer was needed, as he both uttered the sound and gave the  
156 'copy' signal (see figure 1). Table 1 gives the complete list of sounds by phase examined in this  
157 study and Table S1 gives the description of sounds. Audio samples of each demonstrated sound and  
158 of the subject's copy are available in the electronic supplementary material.

159 All sessions were videotaped and were recorded with Fostex Fr2 and Zoom H-4N digital  
160 recorders and a Rode NTG-2 condenser shotgun microphone. To play the sounds in the speaker  
161 condition a sound launcher app for iOS 'SoundPad Live' was developed. The sounds were played  
162 through an Ipad to an Ik Multimedia 'I Loud' portable Bluetooth speaker.

### 163 **(c) Coding and data analysis**

164 The analysis comprised two steps. In the first step we used a traditional method of categorization  
165 that consisted of using acoustic inputs and making a selection of the sounds that looked more  
166 similar [23,26,39-41]. That is, one experimenter listened to each test trial, and scored whether the  
167 subject's vocal response correctly matched the sound uttered by the model. Then, for reliability  
168 analysis, six naïve judges, blind to the model's true sound were presented with pair of sounds  
169 (model and candidate copies) and were asked to judge if the copy matched the model sample  
170 (scoring Yes for correct matching or No for non-matching) across 6 samples (3 correct and 3  
171 incorrect, the latter chosen randomly from the pool of sounds emitted by the subject) for each  
172 demonstrated sounds.

173 Next, using a visual inspection of the wave form we analysed two time domain-related parameters,  
174 namely, the number and duration of bursts, of a random sample of 5 copies of each novel  
175 vocalization using Adobe Audition and then we calculated the intraclass correlation coefficient

176 (ICC) as a measure of concordance between model and copy sounds. The ICC for absolute  
177 agreement was estimated using a two-way random effects model.

178 We also run an objective detailed analysis in which the demonstrated and imitated sounds  
179 selected in the first step were subjected to an analysis of matching accuracy using algorithms  
180 implemented in Matlab version 2014a, using the signal processing toolbox Version 6.21  
181 (R2014a) and the additional code and scripts designed by Lersch 2011 [42] available at  
182 <http://www.audiocontentanalysis.org/code/>. These analyses went through several steps.

183 First, we selected and extracted a subset of acoustic features (e.g. statistics, timbre or quality of  
184 sound, intensity-related, tonal or temporal) of both model-copy time-variant sounds that allowed us  
185 to compare sounds produced with the remarkably different acoustic modes of production  
186 aforementioned. These features form a compact informative set with respect to the desired  
187 properties of the original data [42]. All of these features were implemented using a 20 ms time  
188 window, hamming windowing, with an overlap of 50 % (hop 10 ms).

189 The challenge was to select in an exploratory approach a subset of these features in time and  
190 frequency domains that a priori seemed suitable for comparing sounds made by two species that use  
191 totally different production mechanisms. The main features selected were as follows: 1) *Spectral*  
192 *Pitch Contour ACF* (Autocorrelation Function of the Magnitude Spectrum), that shows the  
193 evolution of the fundamental frequency over time; 2) *Time Energy Evolution*, that allows to  
194 compare the evolution of the energy pattern over time between the model's and the subject's  
195 acoustic signals (temporal regularity and rhythm); 3) *Pitch Class Profile*, a histogram-like 12-  
196 dimensional vector (corresponding to the 12 notes of the diatonic musical scale) with each  
197 dimension representing both the number of occurrences of the specific pitch class in a time frame  
198 and its energy or velocity throughout the analysis block [42]. Figure 2 presents an example of a  
199 *Wave form*, *Spectrogram* and *Pitch Class Profile* of the demonstrated and the copy of the human  
200 (tonal) novel sound 'Hello', and of the conspecific (atonal) novel sound 'Breathy Raspberry'



201 acoustic analyses. (See figures S2-S4 in the electronic supplementary material for one example for  
202 each spectral analysis for each of the main features selected and for a complete list of all features  
203 selected.)

204       Second, once these features were selected all the characteristics of each frame were compacted  
205 into a single vector. Finally, for the comparison it was necessary to then take into account that these  
206 signals were of different duration. We utilized a *Dynamic Time Warping (DTW)* method to deal  
207 with the alignment task, that is, with the operations of stretching and compressing audio parts  
208 allowing similar shapes to match even if they are out of phase in the time domain. DTW represents  
209 a family of algorithms developed for the automated recognition of human speech that allows for  
210 limited compression and expansion of the time axis of a signal to maximize frequency overlap with  
211 a reference signal [42]. DTW is a more robust distance measure for time series capable of  
212 quantifying similarity (or dissimilarity) in an optimal way [42] as, typically, dissimilarity function is  
213 a Euclidean distance measure that calculates and cumulates a cost according to a correspondence  
214 function (where a zero cost indicates a perfect match). That is, the higher the matching cost, the  
215 more dissimilar (less similar) the two sequences.

216       DTW has been widely documented and used in digital signal processing, artificial intelligence  
217 tasks such as pattern recognition (e.g., sign and gestural language), music information retrieval and  
218 signal processing, audio forensic or machine learning [42] and has recently proven to be an  
219 excellent technique for assessing matching accuracy between sounds produced by marine mammals  
220 and in particular for automatic classification of killer whale call types [43-,45]. In the present study,  
221 DTW was used to measure dissimilarity of the aforementioned acoustic subset of features that were  
222 previously selected between the audio signal of the demonstrated sound and that of the subject,  
223 revealing the extent of alignment or synchronization between both signals.

224       Finally, in order to establish relative comparisons between any model-copy sound pair a  
225 ‘dissimilarity index’ scale was constructed, which allowed us to calibrate the distance measures

226 obtained in the DTW analyses and thus establish how similar or dissimilar were the two sounds  
227 (demonstrated sound and that of the subject) in all the subsets of features selected. Since the  
228 dissimilarity index does not have a fixed upper limit, we rescaled the index into an interval from 0  
229 to 1 to quantitatively assess the degree of dissimilarity. As in the non-rescaled version, 0 in this  
230 scale represents a perfect copy (i.e., a sound compared with a copy of itself) and 1 represents  
231 maximum dissimilarity. To establish this ceiling value (the top of the scale) we chose a main  
232 benchmark value, technically referred to as ‘anchor’. Since the value depends on the particular  
233 vocalizations analysed, indices of dissimilarity were calculated between four randomly chosen  
234 demonstration sounds and copies uttered by the subject that corresponded to other different  
235 demonstrated sound. The benchmark value chosen was the round score closest to the maximum  
236 found (940378 score for ‘Amy’ paired with ‘One Two Three’), which accordingly in this case was  
237 rounded to 1000000 (See ESM for a complete list of DTW dissimilarity index scores.) The rescaled  
238 dissimilarity index represents the division of the accumulated distance in relation to the distance  
239 value of the anchor of dissimilarity. Among these same four pairs of different sounds we also took  
240 the lowest score (the more similar) as another benchmark for what could be considered bad and  
241 good copies. Finally, another benchmark was included to serve as a reference point for what could  
242 be considered a ‘high quality match’ (i.e. a human copying another human known word). For this  
243 we calculated the dissimilarity index between the sound ‘Hello’ produced by the trainer and the  
244 experimenter copy of the same sound (see figure 4).

### 245 **3. Results**

246 Inter-observer reliability of whether model and subject sounds matched was high (Fleiss’ weighted  
247 kappa: 0.8;  $p < 0.001$ ; Observed agreement = 0.90).

#### 248 (a) Familiar sounds

249 The subject correctly copied all of the trained sounds, either demonstrated by a conspecific or by a  
250 human. In *Phase 1* the subject recalled the copy command given by the trainer 4 years before as

251 indicated by her response in the first trial. *Phase 2* involved testing the subject's response to the  
252 trainer's copy command when the model uttered familiar sounds. With the copy signal alone the  
253 sound 'Song' was copied in the 1st trial, 'Blow' was copied in the 2nd trial (first session) and  
254 'Birdy' was accurately matched in the 34<sup>th</sup> trial (sixth session). The criterion required for moving to  
255 the final experimental phase, i.e., 90 % of correct trials, on these three intermixed familiar sounds  
256 was achieved by Wikie in the seventh session. In *Phase 3* the subject also copied correctly all of the  
257 trained conspecific sounds performed by a human model in the transfer sessions (n = 2) and in the  
258 first trial. In sum, the subject made recognizable copies of the demonstrated sound judged in real  
259 time by two observers, Wikie's trainer and one experimenter, and then confirmed by both of them  
260 listening to the recordings.

#### 261 (b) Novel sounds

262 The subject produced recognizable copies of all of the untrained sounds, either demonstrated by a  
263 conspecific or by a human (as judged by two experimenters that listened to the sound recordings  
264 after the test and then confirmed by 6 independent observers). In the *live conspecific condition* the  
265 novel sounds (n = 3) were copied before the 10<sup>th</sup> trial ('Strong Raspberry'), with one sound copied  
266 in the 2nd trial ('Creaking Door'), and the other in the 3rd trial ('Breathing Raspberry'). In the  
267 *conspecific through speaker condition*, the novel sounds (n = 2) were copied before the 17<sup>th</sup> trial  
268 ('Wolf'), with the other sound copied in the 6th trial ('Elephant'). In the *conspecific through human*  
269 *model condition* the novel sound tested (n = 1) was copied in the first trial ('Strong Raspberry').  
270 Finally, in the *human sound condition* the novel sounds (n = 6), although they weren't perfect  
271 copies, Wikie produced recognizable copies of the human model sounds before the 17<sup>th</sup> trial ('Ah  
272 Ah'), with two sounds copied in the first trial ('Hello ' and 'One, Two, Three').

273 Visual examination of spectral patterns revealed a good matching of the demonstrated sound  
274 and the subject's copy in several of the acoustic features analysed. For all sound parameters tested,  
275 no differences were observed between the model's sound and the subject's match in the *total*

276 *number of bursts* (Cohen's kappa = 1,  $p < .0005$ ). When tested with novel conspecific sounds, a  
277 high concordance was found between *burst duration* of the model's sound and the subject's copy  
278 (ICC: 0.79;  $p < 0.001$ , N=31 bursts). When tested with human sounds, a very high concordance  
279 between *burst duration* of model's sound and subject's copy was found (ICC: 0.89;  $p < 0.001$ ,  
280 N=65 bursts) showing better performance compared to killer whale sounds.

281 In the automated quantitative analysis, the DTW showed an optimal overlap represented by a  
282 diagonal line alignment between both sounds (demonstrated and copy) in all the examples for each  
283 sound judged by the experimenters as correct imitations in phase 1. This diagonal line alignment of  
284 the 'shortest line' between both signals indicated similarity in all features selected [42]. Figure 3  
285 presents an example of a DTW analysis in the matching of the subject's and the human model's for  
286 the sound 'Hello' (tonal); the conspecific's novel sound 'Breathy Raspberry' (atonal) and the  
287 familiar sounds 'Birdy' (tonal) and 'Blow' (atonal). (See figure S1 in the electronic supplementary  
288 material for one DTW example of all the others novel sounds tested). Although the fundamental  
289 frequency of copies made by human and killer whale models was remarkably dissimilar, the outline  
290 F0 contours turned out to be very similar. Figure 4 shows a representation of a DTW distance  
291 dissimilarity index between the demonstrated sound and the best match (the lowest DTW value)  
292 among the random sample of 5 copies of each vocalization type of the subject for each and every  
293 sound tested plus four 'incorrect' reference control points (corresponding to randomly chosen  
294 demonstrated sounds paired with copies that corresponded to other different subject's sound and  
295 another 'high quality copy' reference control point (human copying another human known word),  
296 (see ESM for a complete list of DTW dissimilarity index scores). Overall, expected matches (when  
297 demonstration and copy were of the same sound type) did match, while expected non-matches  
298 (when demonstration and copy were of different sound types) did not. Specifically, we found that  
299 copies of *familiar* conspecific sounds fell below a dissimilarity index threshold (horizontal red  
300 dotted line below the lowest incorrect random pair copy) that divided our results in good or bad  
301 copies and most of them were close to the 'high quality match' score (human imitating human

302 anchor), with one score being below this value ('Blow'). Copies of *novel* conspecific sounds were  
303 located very close to this 'high quality match' score and *novel* speech sounds demonstrated by  
304 humans were distributed across the whole range of good copies with one even below this 'high  
305 quality match' benchmark. If we take as a criterion of matching accuracy the values obtained with  
306 familiar sounds from conspecifics, we observe that except for the sound 'Blow', which is the  
307 simplest untrained sound consisting only of a single burst of atonal voiceless breath (see electronic  
308 supplementary material second example on Sound File N°1), the copies of *novel* conspecific sounds  
309 and three of *novel* speech sounds ('Amy', 'Hello' and 'Ah Ah') were even more closely matched  
310 than were tonal *familiar* conspecific sounds.

311 Finally, analysing the features selected for the DTW analysis separately, the spectrogram  
312 analysis revealed that the subject produced harmonics when exposed to tonal sounds, but not when  
313 exposed to atonal or noisy sounds (see figure 2 and electronic supplementary material figures S2-  
314 S4). This pattern held even for the human tonal sounds.

#### 315 **4. Discussion**

316 Although the subject did not make perfect copies of all novel conspecific and human sounds,  
317 nonetheless, they were recognizable copies as assessed by both external independent blind  
318 observers and the acoustic analysis. There was great variability in the number of good copies  
319 produced after a sound was copied for the first time (Table 1). Possible factors that could explain  
320 this variability are the difficulty in producing novel sounds and some uncontrolled factors such as  
321 variation on motivational levels and social dynamics across sessions. Additionally, our non-  
322 differential reinforcement regime (good copies of novel sounds were not reinforced to avoid  
323 shaping) may have also contributed to this variability. Consequently, it is conceivable that our data  
324 represent a conservative estimate of the killer whale's capacity for vocal imitation.

325 According to the DTW dissimilarity scale (figure 4), all the copies of novel conspecific  
326 utterances fell below the dissimilarity index threshold for good and bad copies (pairs of different

327 demonstrated and copied sounds randomly chosen) and most of them were close or even fell below  
328 the ‘high quality match’ score, as represented by the human-copying-human anchor. Similarly,  
329 although three of the copies of human sounds were only close to the dissimilarity index threshold  
330 for good and bad copies (incorrect randomly paired copies), the other three fell close to the ‘high  
331 quality match’ score (human imitating human anchor); that is, they were very accurate copies, with  
332 one falling even below this benchmark. This level of accuracy is particularly remarkable given that  
333 the subject possessed a very different sound production system compared to humans. Some  
334 parameters such as the fundamental frequency were sometimes drastically different between the  
335 human model and the killer whale copies, but the outline F0 contours were nonetheless quite similar  
336 (figure 4).

337 Overall, the DTW analyses revealed that the accuracy of copies was much higher when these  
338 were of the same sound than when they involved a different sound, which strongly suggests that the  
339 copies were specific to the demonstrated sound. We believe that the subject’s responses represent a  
340 case of *vocal imitation* rather than response facilitation, as the latter form of social learning does not  
341 apply to individuals reproducing a model’s novel sound [46]. Moreover, the subject’s perfect  
342 performance in the control ‘non-copy’ trials in which she was requested to perform a trained action  
343 or sound different from that of the model, ruled out automatic response facilitation (i.e., copying the  
344 model’s sound spontaneously) [46] because she only copied what she was requested to do so.

345 DTW analyses also revealed that the subject’s copies of *novel* conspecific and human sounds  
346 were in most cases even more accurate than were the copies of *familiar* sounds. Thus, in three of  
347 the novel speech sounds (‘Hello’, ‘Amy’ and ‘Ah Ah’), the accuracy of the copies was even greater  
348 than the matching accuracy of some of the familiar sounds uttered by the conspecific model.  
349 Moreover, four copies of novel sounds were found to be high quality matches, as they were close to  
350 the benchmark score of a human copy of the human sound, and one was even a better match (see  
351 ‘Breathy Raspberry’ in figure 4). A greater copying accuracy for novel compared to familiar sounds

352 might suggest that the cognitive mechanisms responsible for producing familiar and novel sounds  
353 do not fully overlap. It is possible that the matching of familiar sounds relies more heavily on  
354 response facilitation than imitation where the subject's copy is mainly shaped by the *general*  
355 characteristics of the stored representation than by the sound's *specific* individual components. In  
356 contrast, learning to match a novel action or sound might require the subject to carefully process the  
357 individual components of the auditory experience, which might generate a better match. The  
358 subject's matching accuracy is all the more remarkable as she was able to accomplish it (a) in the  
359 absence of extensive trial-and-error across all the experimental conditions, (b) in response to sounds  
360 presented in-air and not in-water (the species' usual medium for acoustic communication), and (c)  
361 in the case of her matching of speech sounds, through the use of a sound production system that  
362 greatly differs from that of the model's [35, 36]. Note that the subject readily matched the harmonic  
363 quality of human tonal sounds (see figure 2 and electronic supplementary material figures S2-S4).  
364 The anatomical structures involved in sound production of cetaceans differ from those used by  
365 terrestrial mammals and birds in that cetaceans are adapted to an aquatic lifestyle where the sound  
366 producing organs compress while diving because of water pressure related changes [35]. This has  
367 been hypothesized to have favoured the development of vocal learning in marine mammals as they  
368 need to have a substantial voluntary control over sound production in order to successfully meet the  
369 demands of reliably generating the same sounds at different depths [47].

370 Our experimental findings lend support to the hypothesis that the group-differentiated acoustic  
371 dialects that have been documented in many field studies of killer whales [16-23] and other  
372 cetaceans [10] can be acquired and maintained through social learning and, more specifically,  
373 through imitation. These results add to the growing database of socially learned sounds reported in  
374 previous non-experimental and experimental studies of killer whales and other cetaceans (dolphins  
375 [11-13]; belugas [14,15]). As a mammalian order, cetaceans stand out for their complex sociality,  
376 elevated encephalization, and advanced cognitive skills [48]. Compared to the fission-fusion  
377 societies of bottlenose dolphins, however, the social systems of killer whales are reported to be

378 more strongly structured and closed [10,16]. Thus, the well-developed propensity of killer whales to  
379 copy what others are doing, that is, to translate visual or auditory input into motor responses that  
380 conform to the group's norm would be consistent with the body of observations on group-specific  
381 acoustic dialects, synchronized behaviour, and sophisticated cooperative strategies documented in  
382 this species [10].

383 The results reported here show that killer whales have evolved the ability to control sound  
384 production and qualify as open-ended vocal learners. It can be argued that since our experimental  
385 design included in-air (rather than in-water) sounds, the positive results obtained cannot directly  
386 reflect the killer whales' capacity for learning to copy underwater sounds in their natural  
387 environment. However, our main objective was to test whether the killer whales were capable of  
388 learning novel sounds through imitative learning, regardless of the type of sound (in-air vs. in-  
389 water) and the model (conspecifics vs. heterospecifics). The atypical nature of the sounds that we  
390 used represents a strength rather than a weakness in relation to our main question because it  
391 demonstrates flexibility not just on *what* is copied but on *how* is copied. With regard to what is  
392 copied, our data demonstrate that killer whales can copy sounds outside their usual repertoire –  
393 which is an important piece of information if one wants to know not only know what a species does,  
394 but also what it can do, under a variable set of circumstances. With regard to the issue of how it is  
395 copied, our data might indicate that the sensory-perceptual and cognitive skills recruited in  
396 imitating in-air sounds are ancestral traits, dating back to the terrestrial ancestors of cetaceans.  
397 Moreover, given the highly derived state of the sound producing apparatus uniquely evolved by  
398 cetaceans, the imitative capacities found in this study also underscore the fine-tuned ability of this  
399 species to flexibly produce accurate matches of heterospecific in-air sounds.

400 Future experimental studies of imitation of in-water sounds demonstrated by conspecifics are  
401 needed to firmly establish the role of social learning in the killer whale's vocal dialects documented  
402 in the wild. Another challenge for future research is to ascertain whether the neural and cognitive



403 scaffolding for vocal learning in cetaceans and humans (and other taxa) are homologous or  
404 analogous, and whether they are adaptations or have been co-opted for new fitness-enhancing  
405 functions in the unique suite of environmental challenges they encounter in the seascape they  
406 inhabit [6,10,11,47,48]. Finally, we extended DTW analysis used in previous studies [39,44,45] by  
407 incorporating several additional features of killer whales' demonstrated and imitated sounds into the  
408 algorithm. However, these results must be taken with caution because the choice of features was  
409 exploratory. Further studies are thus needed to standardize the assessment of the matching accuracy  
410 of different sound features as well as the validation of the dissimilarity index. Although we see  
411 great potential in this analytical approach for comparative studies of vocal learning, its applicability  
412 may vary depending on the study's objectives, the sounds investigated, and the species' vocal  
413 production system.

414

## 415 **References**

- 416 1. Zentall TR. 2012 Perspectives on observational learning in animals. *J. Comp. Psychol.* **126**, 114. (doi: 10.1037/a0025381)
- 417 2. Boyd R, Richerson PJ, Henrich J. 2011 The cultural niche: why social learning is essential for human adaptation. *Proc. Natl.*  
418 *Acad. Sci. USA* 108 (Suppl 2), 10918–10925. (doi: 10.1073/pnas.1100290108)
- 419 3. Whiten A, Van Schaik CP. 2007 The evolution of animal ‘cultures’ and social intelligence. *Philos. Trans. R. Soc. Lond. B, Biol.*  
420 *Sci.* **362**, 603-620. (doi: 10.1098/rstb.2006.1998)
- 421 4. Janik VM, Slater PJB. 2000 The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1-11.  
422 (doi:10.1006/anbe.2000.1410)
- 423 5. Tomasello M. 2011 Human culture in evolutionary perspective. *Adv. Cult. Psychol.* **1**, 5-51.
- 424 6. Colbert-White EN, Corballis MC, Frigaszy DM. 2014. Where apes and songbirds are left behind: A comparative assessment of  
425 the requisites for speech. *Comp. Cog. Behav. Rev.* **9**, 1-28. (doi: 10.3819/ccbr.2014.90004)
- 426 7. Hauser M, Chomsky N, Fitch WT. 2002 The Language Faculty: What is it, who has it, and how did it evolve? *Science* **298**,  
427 1569-1579. (doi: 10.1126/science.298.5598.1569)
- 428 8. Seyfarth RM, Cheney DL. 2014 The evolution of language from social cognition. *Curr. Opin. Neurobiol.* **28**, 5-9. (doi:  
429 10.1016/j.conb.2014.04.003)
- 430 9. Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. 2016. Vocal fold control beyond the species-specific repertoire in  
431 an orang-utan. *Sci. Rep.*, **6**.
- 432 10. Whitehead H, Rendell L. 2014 *The cultural lives of whales and dolphins*. University of Chicago Press.
- 433 11. Janik VM. 2014 Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* **28**, 60-65. (doi:  
434 10.1016/j.conb.2014.06.010)
- 435 12. Herman LM. 2002 Vocal, social, and self-imitation by bottlenosed dolphins. In *Imitation in animals and artifacts* (eds K.  
436 Dautenhahn C. Nehaniv), pp. 63-108. Cambridge, UK: MIT Press.
- 437 13. Richards, D. G., Wolz, J. P., & Herman, L. M. 1984 Vocal mimicry of computer-generated sounds and vocal labeling of objects  
438 by a bottlenosed dolphin, (*Tursiops truncatus*). *J. Comp. Psychol.* **98**, 10
- 439 14. Ridgway S, Carder D, Jeffries M, Todd M. 2012 Spontaneous human speech mimicry by a cetacean. *Curr. Biol.* **22**, R860-R861.  
440 (doi: 10.1016/j.cub.2012.08.044)
- 441 15. Murayama T, Iijima S, Katsumata H, Arai K. 2014 Vocal imitation of human speech, synthetic sounds and beluga sounds, by a  
442 beluga (*Delphinapterus leucas*). *Int. J. Comp. Psychol.* **27**.
- 443 16. Filatova OA, Samarra FI, Deecke VB, Ford JK, Miller PJ, Yurk, H. 2015 Cultural evolution of killer whale calls: background,  
444 mechanisms and consequences. *Behaviour* **152**, 2001-2038. (doi:10.1163/1568539X-00003317)
- 445 17. Ford JK. 1991 Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.*  
446 **69**, 1454-1483. (doi:10.1139/z91-206)
- 447 18. Deecke VB, Ford JK, Spong P. 2000 Dialect change in resident killer whales: implications for vocal learning and cultural  
448 transmission. *Anim. Behav.* **60**, 629-638. (doi: 10.1006/anbe.2000.1454)
- 449 19. Miller PJ, Bain DE. 2000 Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca*. *Anim. Behav.* **60**,  
450 617-628.
- 451 20. Filatova OA, Miller PJO. 2015 An agent-based model of dialect evolution in killer whales. *J. Theor. Biol.* **373**, 82-91.  
452 (doi: 10.1016/j.jtbi.2015.03.020)

- 453 21. Weiß BM, Symonds H, Spong P, Ladich F. 2011 Call sharing across vocal clans of killer whales: Evidence for vocal imitation?  
454 *Mar. Mam. Sci.* **27**, E1-E13. (doi:10.1111/j.1748-7692.2010. 00397)
- 455 22. Filatova OA, Deecke VB, Ford JK, Matkin CO, Barrett-Lennard LG, Guzev MA., ... Hoyt, E. 2012 Call diversity in the North  
456 Pacific killer whale populations: implications for dialect evolution and population history. *Anim. Behav.* **83**, 595-603.
- 457 23. Yurk H, Barrett-Lennard L, Ford JKB, Matkin CO. 2002 Cultural transmission within maternal lineages: vocal clans in resident  
458 killer whales in southern Alaska. *Anim. Behav.* **63**, 1103–1119. (doi:10.1006/ anbe.2002.3012)
- 459 24. Foote AD, Griffin RM, Howitt D, Larsson L, Miller PJ, Hoelzel AR. 2006 Killer whales are capable of vocal learning. *Biol. Lett.*  
460 **2**, 509-512.
- 461 25. Abramson JZ, Hernández-Lloreda MV, Call J, Colmenares F. 2013 Experimental evidence for action imitation in killer whales  
462 (*Orcinus orca*). *Anim. Cogn.* **16**, 11-22.
- 463 26. Bain DE. 1986 Acoustic behavior of *Orcinus*: sequences, periodicity, behavioral correlates and an automated technique for call  
464 classification. *Behavioral biology of killer whales*, 335-371.
- 465 27. Crance JL, Bowles AE, Garver A. 2014 Evidence for vocal learning in juvenile male killer whales, (*Orcinus orca*), from an  
466 adventitious cross-socializing experiment. *J. Exp. Biol.* **217**, 1229-1237.
- 467 28. Musser WB, Bowles AE, Grebner DM, Crance JL. 2014 Differences in acoustic features of vocalizations produced by killer  
468 whales cross-socialized with bottlenose dolphins. *The Journal of the Acoustical Society of America*, **136**, 1990-2002.
- 469 29. Pepperberg IM. 2010 Vocal learning in Grey parrots: A brief review of perception, production, and cross-species comparisons.  
470 *Brain. Lang.* **115**, 81-91.
- 471 30. Klatt DH, Stefanski RA. 1974. How does a mynah bird imitate human speech? *J. Acoust. Soc. Am.* **55**, 822–832.
- 472 31. Stoeger AS, Mietchen D, Oh S de Silva S, Herbst CT, Kwon S, Fitch WT. 2012 An Asian elephant imitates human speech. *Curr.*  
473 *Biol.* **22**, 2144-2148.
- 474 32. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Can. J. Zool.* **63**, 1050–  
475 1056.
- 476 33. Fitch WT. 2006 Production of vocalizations in mammals. *Vis. Commun.* **3**,145.
- 477 34. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016. Monkey vocal tracts are speech-ready. *Sci. Adv.* **2**, e1600723.
- 478 35. Cranford TW, Amundin M, Norris KS. 1996 Functional morphology and homology in the odontocete nasal complex:  
479 implications for sound generation. *J. Morph.* **228**, 223-285.
- 480 36. Tyack PL, Miller EH. 2002. Vocal anatomy, acoustic communication and echolocation. *Marine mammal biology: An*  
481 *evolutionary approach*, 142-184.
- 482 37. Hayes KJ, Hayes C. 1952. Imitation in a home-raised chimpanzee. *-J. Comp. Physiol. Psychol.* **45**, 450-459.
- 483 38. Huber L, Range F, Voelkl B, Szucsich A, Viranyi Z, Miklosi A. 2009 The evolution of imitation: what do the capacities of non-  
484 human animals tell us about the mechanisms of imitation? *Phil. Trans. R. Soc. B.* **364**, 2299-2309. (doi:10.1098/rstb.2009.0060)
- 485 39. Hodgins-Davis A. 2004 ‘An Analysis of the Vocal Repertoire of the Captive Killer Whale Population at Marineland of Antibes,  
486 France,’ Thesis, Wellesley College,
- 487 40. Ford JK. 1989 Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. *Can. J.*  
488 *Zool.*, **67**, 727-745.
- 489 41. Deecke VB, Ford JK, Spong P. 1999 Quantifying complex patterns of bioacoustic variation: use of a neural network to compare  
490 killer whale (*Orcinus orca*) dialects. *J. Acoust. Soc. Am.* **105**, 2499-2507.

- 491 42. Lerch, A. 2012 Temporal Analysis. An Introduction to Audio Content Analysis: Applications in Signal Processing and Music  
492 Informatics, 119-137.
- 493 43. Deecke VB, Janik VM. 2006 Automated categorization of bioacoustic signals: avoiding perceptual pitfalls. *J. Acoust. Soc. Am.*  
494 **119**, 645-653.
- 495 44. Brown JC, Hodgins-Davis A, Miller PJ. 2006 Classification of vocalizations of killer whales using dynamic time warping. *J.*  
496 *Acoust. Soc. Am.* **119**, EL34-EL40. (doi:10.1121/1.2166949)
- 497 45. Brown JC, Miller PJ. 2007 Automatic classification of killer whale vocalizations using dynamic time warping. *J. Acoust. Soc.*  
498 *Am.* **122**, 1201-1207. (doi:10.1121/1.2747198)Byrne RW. 2002 Imitation of novel complex actions: What does the evidence  
499 from animals mean? *Adv. Study Behav.* **31**, 77-105. (doi: 10.1016/S0065-3454(02)80006-7)
- 500 46. Hoppitt W, Laland KN. 2008. Social processes influencing learning in animals: a review of the evidence. *Adv. Study Behav.* **38**,  
501 105-165.
- 502 47. Tyack PL, Sayigh LS. 1997. Vocal learning in cetaceans. Social influences on vocal development. **11**, 208.
- 503 48. Marino L, Connor RC, Fordyce RE, Herman LM, Hof PR et al (2007) Cetaceans have complex brains for complex cognition.  
504 PLoS Biol **5** 966–972. ( doi:10.1371/journal.pbio.0050139)

505 **Electronic Supplementary Material** is available in the online content of the paper and at  
506 <https://figshare.com/s/2991d28752ca0690e843>. This includes methods details, raw data, figures S1-S5, and 12 audio files examples  
507 (ESM Audio File S1: 3 conspecific familiar sounds; ESM Audio File S2.1-S2.5: 5 conspecific novel sounds; ESM Audio File S3.1 –  
508 S3.6: 6 human novel sounds).

510 **Ethics.** The Ethics and Animal Welfare Committee (CEBA-MEDUC) of the School of Medicine, Pontifical Catholic University of  
511 Chile, have approved this research. This research adhered to the legal requirements of the country (France) in which the work was  
512 carried out and Marineland institutional guidelines.

513 **Data accessibility.** The datasets supporting this article have been uploaded as part of the supplementary material and at  
514 <https://figshare.com/s/2991d28752ca0690e843>.

515 **Author Contributions** J. Z.A. conceived the study. J.Z.A., M.V.H.LL and J.C. designed the experiment, which was conducted by  
516 J.Z.A. and M.V.H.LL. M.V.H.LL designed and carried out the data analyses and interpretation. J.Z.A. and L.G. performed the sound  
517 analyses and interpretation. J.Z.A. and M.V.H.LL drafted the paper. J.Z.A. and F.C. co-wrote the paper. J.C. and F.A. helped to write  
518 and provided critical revisions and some of the ideas in the paper. J.Z.A., F.C., F.A. and J.C. provided financial support. All authors  
519 have read and approved the final manuscript.

520 **Acknowledgements** Special thanks go to Andrea Hodgins for sharing her data and making her thesis available to us and  
521 to Manuel Castellote for his advice on sound analyses. We are grateful to the directors of Parques Reunidos and the Marineland  
522 Aquarium of Antibes, Jesús Fernández and Jon Kershaw for allowing us to conduct this research. Furthermore, we appreciate the  
523 work of head coaches Lindsay Rubinacam and Duncan Versteegh, and all orca trainers team for their help and support. Special  
524 thanks go to Amy Walton for her dedication and training of the killer whale model Moana. Finally we want to thank Francisco  
525 Serradilla from Universidad Politécnica de Madrid for developing the sound launcher app for iOS.

526 **Funding** This project was conducted at the Marineland Aquarium Antibes, France and supported by a Postdoctoral Scholarship  
527 FONDECYT N° 3140580 to J.Z. Abramson. This study was partly funded by project grants PSI2011-29016-C02-01, PSI2014-  
528 51890-C2-1-P (MINECO, Spain) and UCM-BSCH GR3/14-940813 (Universidad Complutense de Madrid y Banco Santander  
529 Central Hispano) to F. Colmenares.

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

531 **Author Information** Correspondence and requests for materials should be addressed to J.F.Z.A (zabramson@psi.ucm.es)

532

533

534

535 **Figure 1. Experimental set up.** (a) Conspecific live condition: The two trainers ( $T_M$  and  $T_S$ ; M for model and S for  
536 subject) were positioned on different sides of a wooden panel 2m long x 1.90cm high placed in a position in which S  
537 and M could see each other and their own trainer, but could not see the other trainer's commands; (b) conspecific  
538 speaker condition: One trainer holds the speaker and another ( $T_S$ ) gave the copy command to the subject; and (c) human  
539 live condition; Just one trainer was needed, as he both uttered the sound and gave the 'copy' signal.

540 **Figure 2. Wave form and spectrogram** of the model (a1) and the copy (a2) of the human (tonal) novel sound 'HE'. Note  
541 the harmonic pattern in both signals. 'HE' *Pitch Class profile* of the model (b1) and the copy (b2) *Wave form and*  
542 *spectrogram* of the model (c1) and the copy (c2) of the conspecific (atonal) novel sound 'BR'. Note the in harmonic  
543 pattern in both signals. 'BR' *Pitch Class profile* of the model (d1) and the copy (d2)

544 **Figure 3. Dynamic Time Warping Familiar and Novel Conspecific and Human Sounds (Tonal and Atonal).** In  
545 both axes all the characteristic features of the signals are aligned and the black line shows the. shortest path (minimum  
546 distance) between the model and the observer sounds streams.. (a); DTW familiar sound 'BL' (atonal) of the model and  
547 the copy (b); DTW familiar sound 'BI' (tonal) of the model and the copy (c); DTW novel sound 'HE' (tonal) of the  
548 model and the copy (d); DTW novel sound 'BR' (atonal) of the model and the copy.

549 **Figure 4 Dynamic Time Warping dissimilarity index distribution.** Distribution of the DTW dissimilarity index  
550 between the model and the copy for each vocalization; familiar (blue dots), killer whale novel (green dots) and human  
551 novel (turquoise dots). Five control benchmarks (red dots) separated by a red vertical dotted line are also  
552 represented, where the first one correspond to the 'high quality match' score (human imitating human benchmark) and  
553 the others correspond to the four randomly chosen incorrect copies (model sounds paired with copies that corresponded  
554 to other different models). The horizontal red dotted line below the lowest incorrect random pair copy serves as a  
555 benchmark for dividing the results between good and bad copies.

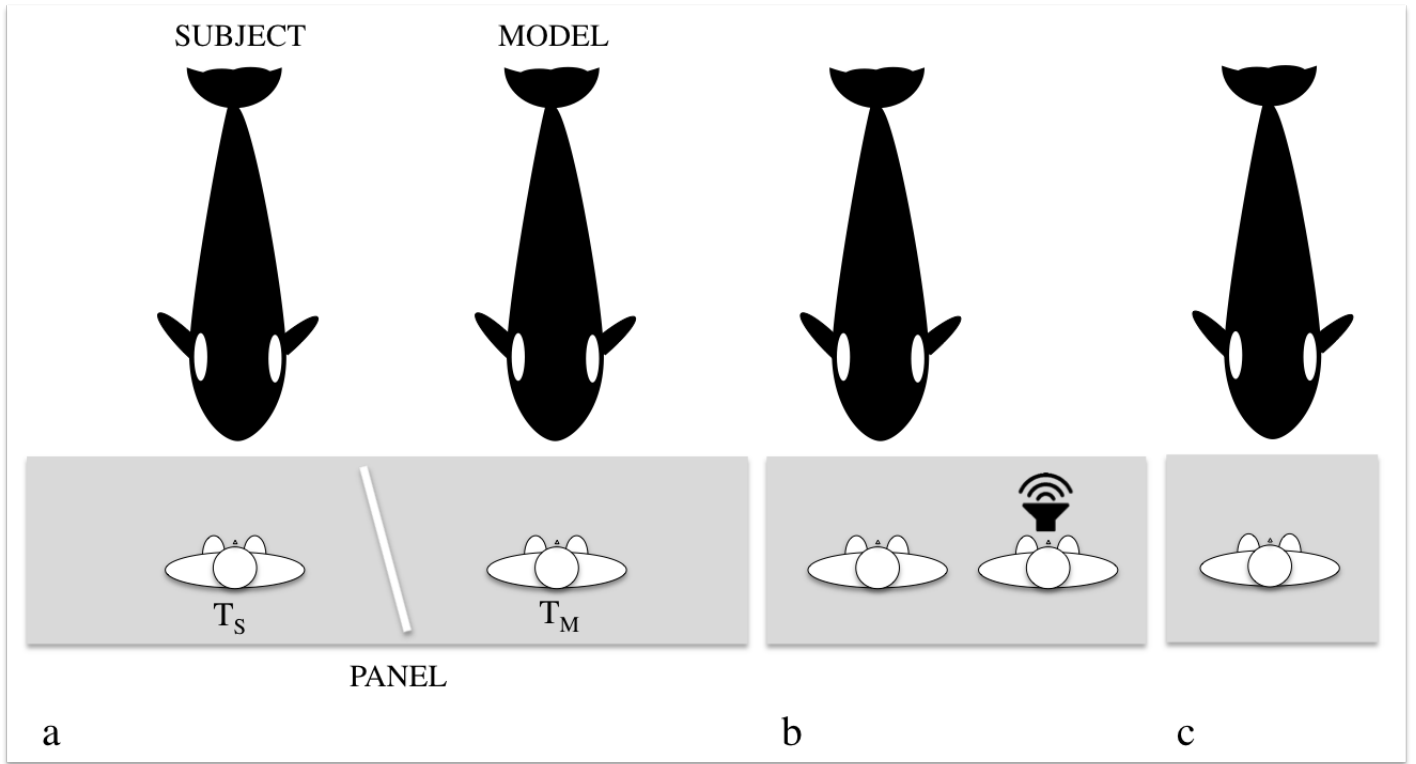
556

557 **Table 1.** Total number of trials for each sound tested, number of trials until the model's sound was judged to be copied  
 558 by the subject (according to two experimenters that listened to the sound recordings after the test and then confirmed by  
 559 6 independent observers), and percentage of correct trials since the first full copy

	No. of trials	First trial copied	% correct since the 1st copy
<b>FAMILIAR SOUNDS</b>			
Song (SO)	394	1	100
Birdy (BI)	316	34	98
Blow (BL)	371	2	99
<i><b>Through human model (transfer sessions)</b></i>			
SO	30	1	100
BL	30	1	100
<b>NOVEL SOUNDS</b>			
<i><b>Conspecific Alive Model</b></i>			
Strong Raspberry (SR)	30	10	19
Creaking Door (CD)	30	2	100
Breathy Raspberry (BR)	30	3	30
<i><b>Conspecific through speaker</b></i>			
SR	30	1	100
CD	30	4	44
BR	30	1	57
Wolf (WO)	30	17	36
Elephant (EL)	30	6	28
<i><b>Conspecific through human model (transfer sessions)</b></i>			
SR	30	1	100
<i><b>Human</b></i>			
Ah Ah (AA)	30	17	14
Hello (HE)	30	1	55
Bye Bye (BB)	30	12	21
Amy (AM)	30	8	26
One Two (OT)	30	3	36
One Two Three (OTT)	30	1	23

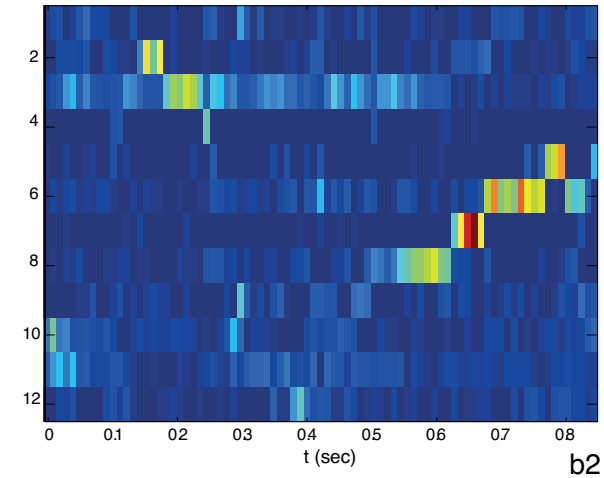
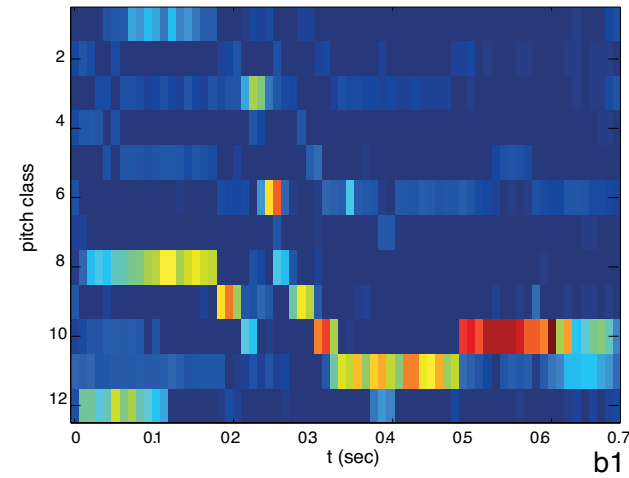
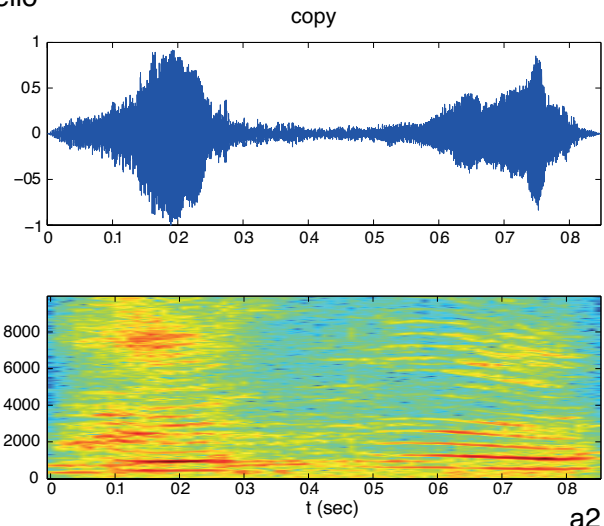
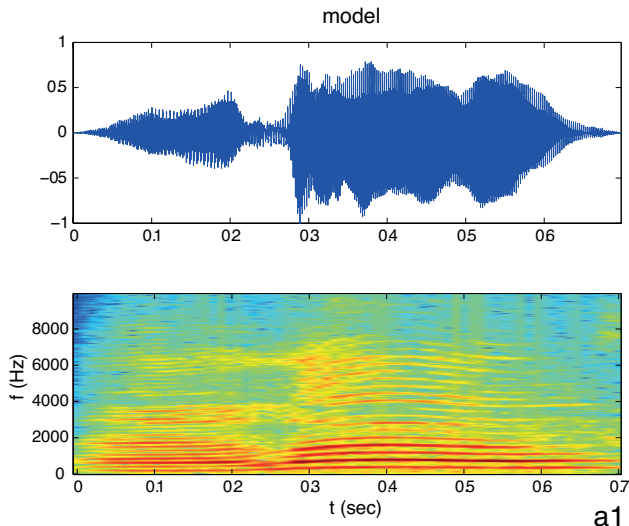
560

561

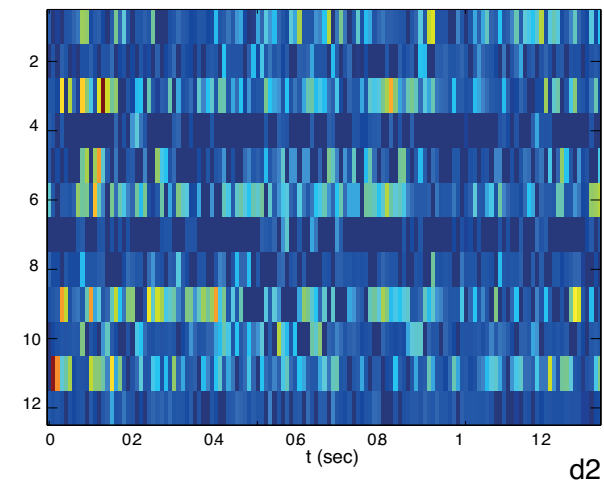
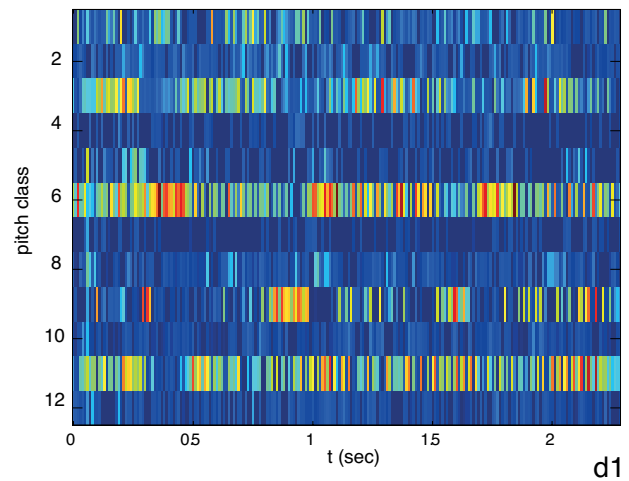
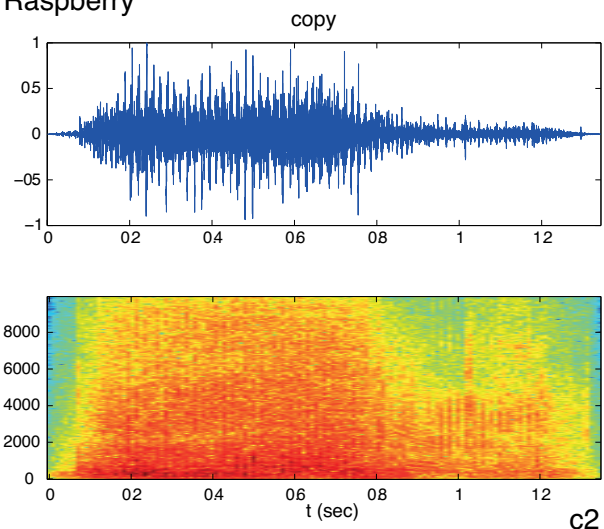
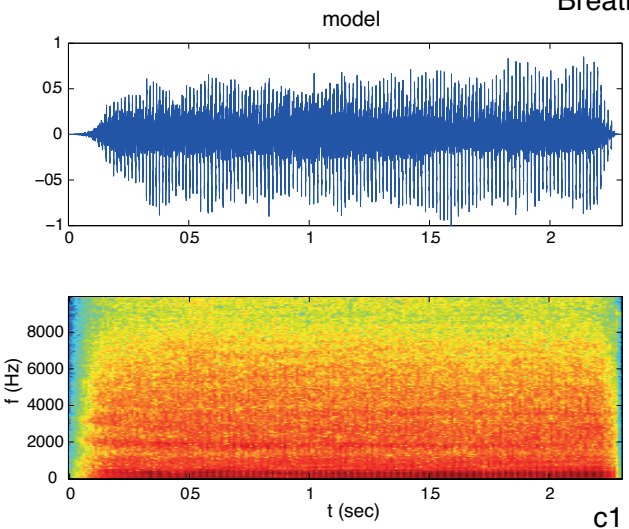




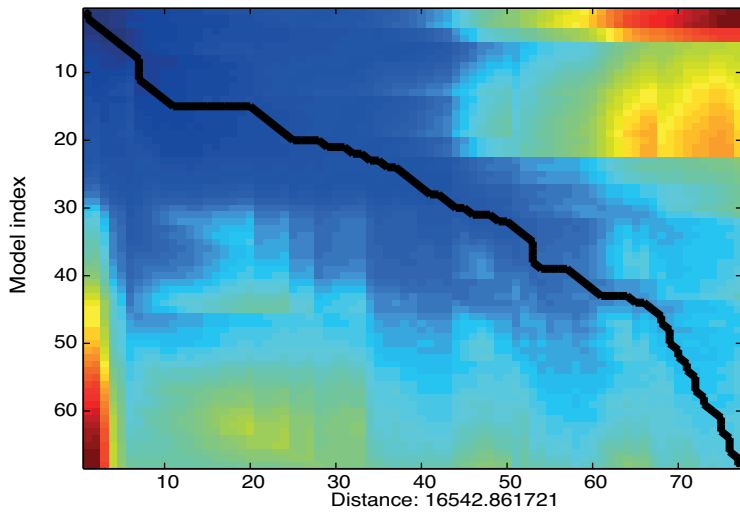
# Hello



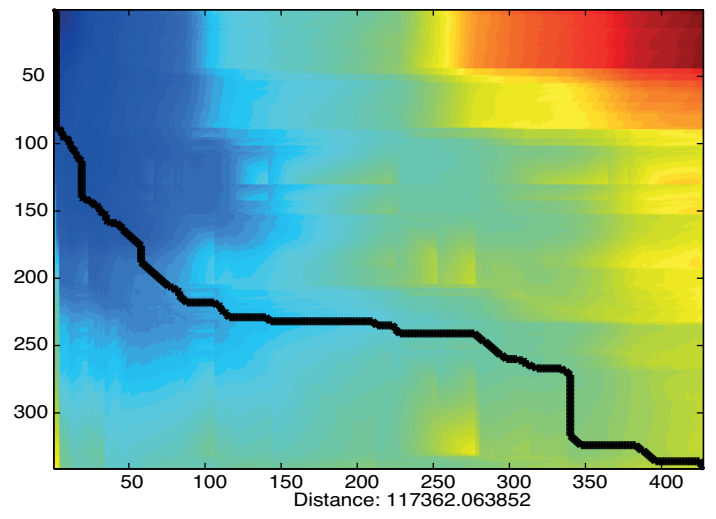
# Breathy Raspberry



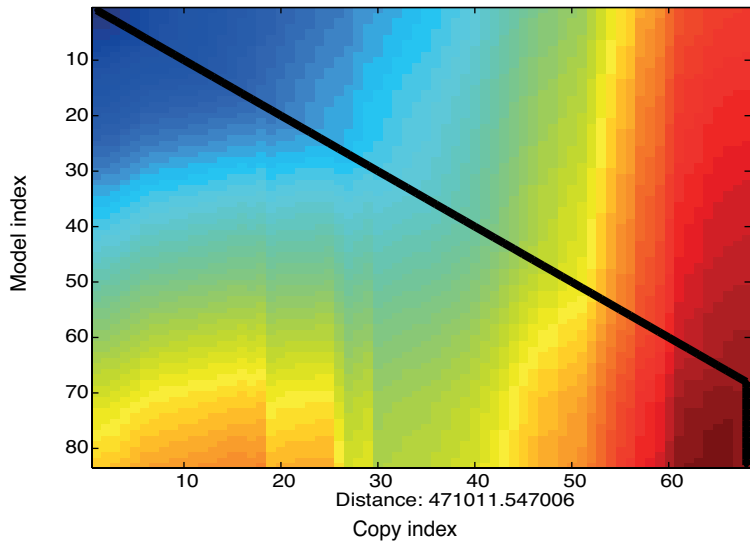
Blow



Birdy



Hello



Breathy Raspberry

