Vocal matching of naval sonar signals by long-finned pilot whales (Globicephala melas)

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Marine Mammal Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>MMSCI-3632.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Note</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>11-Sep-2013</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Alves, Ana; Scottish Oceans Institute, Sea Mammal Research Unit Antunes, Ricardo; Scottish Oceans Institute, Sea Mammal Research Unit Bird, Anna; Scottish Oceans Institute, Sea Mammal Research Unit Tyack, Peter; Scottish Oceans Institute, Sea Mammal Research Unit Miller, Patrick; Scottish Oceans Institute, Sea Mammal Research Unit Lam, Frans-Peter; Netherlands Organisation for Applied Scientific Research, Acoustics &amp; Sonar Kvadsheim, Petter; Norwegian Defence Research Establishment,</td>
</tr>
<tr>
<td>Keywords:</td>
<td>vocal matching, naval sonar, long finned pilot whale, noise, contour extraction, Globicephala melas, Dtags, Norway</td>
</tr>
</tbody>
</table>
Concern over the impact of naval sonar on cetaceans has grown dramatically since mass strandings were linked to sonar exercises (Frantzis 1998). Tyack et al. (2011) reported that beaked whales stopped echolocating and broke off deep foraging dives when exposed to anthropogenic sounds, avoiding the sound source with a prolonged slow ascent. Avoidance reactions are a common response to anthropogenic sound (Southall et al. 2007, Miller et al. 2012). By contrast, long-finned pilot whales (Globicephala melas) were reported to respond to naval sonar by increasing their rate of vocalization while they were huddling in close ranks, as opposed to swimming away (Rendell and Gordon 1999), which resembles the response seen when long-finned pilot whales were exposed to playbacks of killer whale (Orcinus orca) sounds (Curé et al. 2012).

Little is known of the function of calling in pilot whales. One captive short-finned pilot whale (Globicephala macrorhynchus) repeated a stereotyped call made up of two independently modulated tonal frequencies followed by a pulsed component (Caldwell and Caldwell 1969); a similar structure was defined for wild long-finned pilot whales (Nemiroff and Whitehead 2009). Vocal repertoires of groups of long-finned pilot whales recorded in the North Atlantic have been analyzed by extraction of frequency contours (Taruski 1979). Different contours were grouped into seven broad classes with relatively unmodulated calls occurring in low arousal settings and more complex modulation patterns produced in higher arousal states (Weilgart and Whitehead 1990).

In order to quantify responses to sonar, behavior, and sound production of long-finned pilot whale groups were recorded before, during and after 30-60 min experimental sonar exposures conducted in Vestfjord, Northern Norway, in 2008 and 2009. The recordings revealed calls with similar frequency-modulation patterns as
the transmitted sonar signals. Here, we use quantitative comparisons of the contours
of calls and sonar signals to test whether or not the pilot whales changed their calls to
vocally match the sonar signals used during the experiments. Five whales tagged
with DTags (Johnson and Tyack 2003) were followed and tracked visually throughout
each experiment. Each experiment consisted of the period from tag on until the tag
came off, and included multiple exposures to different sonar signal types (Fig. 1)
(Low Frequency Active Sonar - LFAS\textsubscript{UP}: 1-2 kHz hyperbolic upsweep, LFAS\textsubscript{DN}: 1-2
kHz hyperbolic downsweep, and Medium Frequency Active Sonar - MFAS\textsubscript{UP}: 6-7
kHz hyperbolic upsweep). During each exposure the source vessel approached the
tagged whale from a distance of 6-7 km, transmitting sonar pings of 1 s duration with
a 20 s interval between pings, starting with a ramp-up procedure in which the source
level was increased from 152 dB re 1\(\mu\)Pa\textsubscript{m} to maximum level (214 dB re 1\(\mu\)Pa\textsubscript{m} for
LFAS and 199 dB re 1\(\mu\)Pa\textsubscript{m} for MFAS) over a period of 10 min. When the source
vessel was 1 km away from the tagged whale, its course was fixed and sonar
transmission was stopped 5 min after the whale was passed, marking the end of the
exposure. Each exposure had a total duration of approximately 40 min. Details of the
experimental procedures are described by Miller et al (2012). The tags recorded sound
at a sampling rate of 192 kHz and 16-bit resolution. The sound record was inspected
in Adobe Audition (Adobe Systems Incorporated, California, USA) using
spectrograms with 4096 FFT points, and those containing calls or series of
echolocation clicks were marked. Calls included tonal sounds and pulsed calls that are
perceived as tonal due to the high pulse repetition rate. The rate of production of
echolocation clicks occasionally increased steadily into a series perceived as a buzz,
which were not included in the analysis because they were considered to function in
echolocation (Aguilar de Soto \textit{et al}. 2008). Calls were classified into three categories
of perceived signal-to-noise ratio: 1) for faint, barely detectable calls, 3) for calls that were loud and clear in the recording and spectrogram, and 2) for sounds that were intermediate between those two extremes. Only calls in category 3 were analyzed further as these were most likely to have been produced by either the tagged whale or by whales in close proximity to the tagged individual. Also the restriction to category 3 calls ensured that frequency contours could be adequately extracted (see below).

The frequency contour of each call was extracted from spectrograms (FFT size 4096, 87% overlap) using a semi-automatic custom program in MATLAB (The MathWorks, Inc., Massachusetts, USA). Initially we ran an automatic contour extraction procedure that linked the dominant frequencies of adjacent FFT windows (Gillespie et al. 2013). These contours were checked by an operator who manually changed them for a better match to the underlying spectrogram. As detailed below, a dissimilarity metric \( m \) was calculated between the contour shape of calls and the three different sonar signals (Figs. 2 and 3).

All frequency contours were linearly interpolated to 5 ms temporal resolution. In each comparison (sonar signal vs. call) the contours were shifted past each other both in frequency and time in 5 ms and 1Hz steps. A matrix of root-mean-squared-differences \( D \) was calculated for each combination of offset values as

\[
D = d_{\Delta f, \Delta j} = \sqrt{\frac{\sum_{i=j+\Delta j} (s_i - c_{j+\Delta j} + \Delta f)^2}{n_{i=j+\Delta j}}}
\]  

(1)

where \( s \) is the vector of frequency values \( (s_j) \) making up the sonar contour, \( c \) is the vector of frequency values \( (c_j) \) of the call, \( \Delta f \) is the frequency offset value and \( \Delta j \) is the time offset in number of 5 ms steps and \( n_{i=j+\Delta j} \) is the number of cases when \( i=j+\Delta j \) (i.e., the number of 5 ms bins that overlapped when contours were lagged by \( \Delta j \)). The
range of values for $\Delta f$ was selected so that either sound's duration overlapped by an arbitrary value of at least 75%. The range of values of $\Delta f$ was selected so that

$$\text{MAX}(c)-\text{MIN}(s) < \Delta f < \text{MIN}(c)-\text{MAX}(s)$$

The dissimilarity metric for each pair of sonar signal and call being compared was taken as $m = \text{MIN}(D)$. A rotation test (DeRuiter and Solow 2008) randomizing the timing of the exposure period was performed to evaluate the likelihood of calls matching the sonar during each period of sonar exposure by chance alone. The rotation test maintains the call sequence as it was produced and is nonparametric, so autocorrelation in the calling data series does not bias the test statistic. To ensure that when evaluating the effects of one exposure period, the potential effects of other exposures in the same experiment did not confound the analysis, calls recorded during other exposures were removed from the record prior to running the rotation test. In each iteration of the randomization test, a mock exposure period of the same duration as the exposure was randomly shifted within the non-exposure period (excluding the actual exposure period and any period of tagging attempt on other individuals of the same sub-group). Any part of the mock exposure beyond the end of the recording was wrapped back to the start. The median value of $m$ of all calls, and the median of the 5% of sounds most similar to the sonar was recorded within each mock exposure period. The median of all calls was used to evaluate an overall shift in vocal output (i.e. all calls produced became similar to sonar signals). The 5% value is arbitrary, but is effective at limiting the contrast to calls most similar to the sonar, which may include high-quality matches (i.e. evaluates if matching calls were produced in addition to calls within the previous range of similarity). The proportion ($P$) of the median values obtained by 10,000 iterations of the rotation test lower than the median value observed during each actual exposure was calculated. Since we ran multiple
comparison tests we adjusted the critical level of significance $\alpha$ with a Bonferroni correction. We therefore conservatively evaluated the statistical significance for each individual test at a level of 0.004 ($\alpha/n$; $\alpha=0.05$, $n=12$ exposure sessions; see below).

Diving depth was measured using the tag’s pressure sensor to investigate any relationship between the occurrence of vocal matching and specific behavioral patterns (e.g., foraging during deep dives). A total of 12 exposure sessions were conducted on five tagged long-finned pilot whale (Table 1), yielding 41 h of acoustic recordings. Overlapping calls precluded contour extraction in many cases; therefore of 3,923 calls in the highest category of perceived signal-to-noise ratio only 2,556 contours were successfully extracted. Evaluating the median dissimilarity metric of all calls within an exposure period as the test statistic (Table 1), one exposure was significantly different from baseline randomizations: MFAS$_{UP}$ gm08_150c ($P<0.0001$, Figs. 1 and 2). When using the 5% most similar calls (Table 1) two exposures (MFAS$_{UP}$ gm08_150c and LFAS$_{DN}$ gm09_138b (Fig. 3), $P<0.0001$) had calls that were significantly more similar to the sonar than was observed during baseline randomizations. With this analysis we demonstrated that in two of 12 exposures pilot whales were significantly more likely to produce calls with frequency modulation similar to the sonar during sonar transmission periods.

A pattern of whistles similar to sonar transmissions was found to decrease with increasing time after individual sonar pings in false killer whales, *Pseudorca crassidens*, but not in tests of two short-finned pilot whales and a melon headed whale, *Peponocephala electra* (DeRuiter *et al.* 2012). In one of the two exposures where pilot whales matched the sonar in our study, the median difference of frequency modulation between all calls and the sonar, was less than expected based upon call production during periods without sonar transmissions. These results
indicate that the whales produced calls that matched the sonar pulses (Table 1 and Fig. 3), and in some cases may have shifted their overall call production to be more similar to the sonar (Fig. 1). This reaction appears either to be a form of vocal matching, which has been described as an individual responding to a stimulus by preferentially producing a similar signal from its repertoire (Stoddard et al. 1992) or a form of vocal mimicry, which has been described as an animal copying another species or an environmental noise (Kelley and Healy 2011). Long-finned pilot whales are reported to produce calls with simple up or down modulation patterns similar to the sonar signals used here (Taruski 1979, Weilgart and Whitehead 1990). However, the vocal repertoire of pilot whales remains so poorly defined that it is difficult to demonstrate whether or not a call similar to an artificial model is part of the pre-exposure repertoire.

Vocal matching has been reported in birds, frogs, primates, and cetaceans (Dobkin 1979, Sugiura 1998, Gerhardt et al. 2000, Janik 2000, Vehrencamp 2001, Kelley et al. 2008), and most commonly involves matching a conspecific call. Most studies in vocal matching have focused on songbirds, where a predominant function is to signal aggression (e.g., Vehrencamp 2001). Early work on vocal mimicry also focused on songbirds, where an early review defined vocal mimicry in terms of function as “the use of calls or songs of other species that are predominantly aggressive, predatory, or otherwise noxious.” and argued that vocal mimicry functions in interspecific competition and avoidance of predation (Dobkin 1979). If matching or mimicry among pilot whales functions as described above for songbirds, this might suggest that in two of our exposures, some of the pilot whales might have matched the sonar exposure as an aggressive signal to a potential threat or noxious stimulus. Matching a threat’s acoustic features might also function to communicate about the
hazard to other group members. However, other functions have been proposed for vocal matching and vocal mimicry. A recent review of vocal mimicry in songbirds concluded that there is no compelling evidence to support any of the functional hypotheses for vocal mimicry in songbirds, but rather that mimicry may be a by-product of the mechanisms by which birds learn sounds of conspecifics (Kelley et al. 2008). We have observed apparent cases of vocal matching among animals in our recordings, indication that pilot whales may also have a natural tendency to match conspecific sounds as do some other odontocetes species (Janik 2000, Miller et al. 2004, Schulz et al. 2008). Though we do not currently understand the possible function of vocal matching in pilot whales, one plausible explanation is that matching the sonar represents a by-product of pilot whales’ natural tendency to match conspecific sounds. If vocal mimicry is a by-product of vocal matching, then our conclusions of the possible implications of pilot whales matching sonar signals do not greatly depend upon whether the signals were learned or not.

Recent work on vocal matching and imitation, especially among mammals, emphasizes a diversity of potential functions including affiliative as well as aggressive displays. Many mammals produce contact calls that converge as individuals form social bonds, using matching to maintain social cohesion (Tyack 2008). In cetaceans, functions for matching include aggressive or affiliative displays (Janik 2000), reinforcing social bonds (Schulz et al. 2008), signaling individual position and coordinating movement trajectories (Miller et al. 2004), or signaling alliance membership to third parties (Janik 2000). The diversity of suggested functions limits our ability to infer the function of pilot whales matching sonar signals.

We analyzed dive behavior (Table 1), vocal behavior, and group structure
For Peer Review
during exposure periods in which pilot whales did or did not match the sonar, but no
evidence from the sonar exposures supported a particular function for vocal matching
of the sonar (Miller et al. 2011). Sonar received pressure levels (SPL) were calculated
on the tagged whale to verify the audibility of the sonar signals. The only published
audiogram for long-finned pilot whales (Pacini et al. 2010) reports hearing thresholds
between 53.1 to 124.4 dB re 1µPa. These thresholds are all well below the maximum
(150-180 dB re 1µPa) received SPL in any of the sonar exposures (Table 1). Although
the audiogram does not report hearing thresholds below 4 kHz, assuming that the
shape of the audiogram is similar to other odontocetes we expect the sonar received
levels to be above the hearing threshold at LFAS frequency. Therefore, it is fair to
assume that in all exposures the whales were able to hear the sonar stimulus.

Interpretation of the vocal matching reported here is hampered by our limited
knowledge of the vocal repertoire of pilot whales, how it develops, the function of
these calls or call matching, and whether pilot whales can learn to imitate novel
sounds that were not part of their vocal repertoire prior to exposure. If the sonar
sounds are similar to a pre-existing sound in the repertoire, which has a particular
function, then the response may be modulated by the specific function of the pre-
existing sound. On the other hand, we cannot rule out that the pilot whale call
contours that were similar to sonar were not part of the pre-exposure repertoire, but
rather reflected vocal production learning (Janik and Slater 2000) in which the subject
learned to copy the sonar signal.

Delphinids such as pilot whales are thought to rely upon social defenses
against threats such as predators (Norris and Dohl 1980), which may yield different
kinds of responses to anthropogenic stimuli compared to animals more likely to
respond with crypsis and/or flight reactions (Morisaka and Connor 2007). If vocal
matching (or mimicry) in pilot whales functions against threats, then the matching of
the sonar could indicate a more severe reaction to a stressful stimulus than if it were a
neutral by-product of vocal matching or vocal learning. We therefore face a situation
where our ignorance of basic behavioral ecology, such as the functions of vocal
matching, impedes our ability to estimate the impact of this relevant response to
anthropogenic sound by wildlife.

Acknowledgments

Thanks to the science and ships’ crews of the RV HU Sverdrup II and MS Strønstad
for the effort during data collection at sea and Paul White for contour extraction
algorithms. This study was funded by the US Office of Naval Research, the Royal
Norwegian Navy and the Norwegian Ministry of Defense, the Netherlands Ministry
of Defense, and by WWF Norway. Animal experiments were carried out under
permits issued by the Norwegian Animal Research Authority (Permit S-2007/61201),
in compliance with ethical use of animals in experimentation. The research protocol
was approved by the University of St Andrews Animal Welfare and Ethics
Committee and the Woods Hole Oceanographic Institution Institutional Animal Care
and Use Committee.

Literature Cited

Aguilar de Soto, N., M. Johnson, P. Madsen, F. Diaz, I. Dominguez, A. Brito and P.

Caldwell, M. C. and D. K. Caldwell. 1969. Simultaneous but different narrow-band
sound emissions by a captive eastern Pacific pilot whale Globicephala scammoni.


Wiley Interscience, New York, NY.


Table 1. Dissimilarity metric observed during exposure and median of 10,000 baseline randomizations. The proportion of randomization iterations that produced values below the observed statistic ($p$) and maximum dive depth and received sound pressure level range for the tagged whale during each exposure is also shown. Cases in which the median score during the exposure was significantly (Bonferroni corrected $\alpha = 0.004$) lower than during baseline randomizations are shown in **bold**. The number of contours analysed are shown inside brackets.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Exposure</th>
<th>Median dissimilarity of all calls in Hz (number of contours analysed)</th>
<th>Median dissimilarity of 5% most similar calls in Hz (number of contours analysed)</th>
<th>Exposure</th>
<th>Dive depth max (m)</th>
<th>SPL range (dB µPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>During exposure</td>
<td>Baseline randomizations</td>
<td>$p$</td>
<td>During exposure</td>
<td>Baseline randomizations</td>
</tr>
<tr>
<td>gm08_150c</td>
<td>LFAS$_{UP}$</td>
<td>403.49 (14)</td>
<td>277.77 (80)</td>
<td>0.56</td>
<td>25.2(1)</td>
<td>29.6(4)</td>
</tr>
<tr>
<td>gm08_150c</td>
<td>MFAS$_{UP}$</td>
<td><strong>46.63 (9)</strong></td>
<td><strong>298.27 (80)</strong></td>
<td>0</td>
<td><strong>10.51(1)</strong></td>
<td><strong>34.69 (4)</strong></td>
</tr>
<tr>
<td>gm08_154d</td>
<td>LFAS$_{UP}$</td>
<td>139.36 (133)</td>
<td>132.78 (778)</td>
<td>0.63</td>
<td>44.59 (7)</td>
<td>53.24 (39)</td>
</tr>
<tr>
<td>gm08_154d</td>
<td>MFAS$_{UP}$</td>
<td>161.77 (18)</td>
<td>165.37 (778)</td>
<td>0.42</td>
<td>42.76 (1)</td>
<td>63.6 (39)</td>
</tr>
<tr>
<td>gm08_159a</td>
<td>LFAS$_{UP}$</td>
<td>342.31 (8)</td>
<td>208.13 (603)</td>
<td>0.93</td>
<td>199.22 (1)</td>
<td>64.83 (30)</td>
</tr>
<tr>
<td>gm08_159a</td>
<td>MFAS$_{UP}$</td>
<td>339.29 (45)</td>
<td>197.34 (603)</td>
<td>0.93</td>
<td>95.38 (2)</td>
<td>83.29 (30)</td>
</tr>
<tr>
<td>gm09_138b</td>
<td>LFAS$_{DN}$</td>
<td><strong>86.92 (63)</strong></td>
<td><strong>267.56 (437)</strong></td>
<td>0.05</td>
<td><strong>17.28 (3)</strong></td>
<td><strong>53.48 (22)</strong></td>
</tr>
<tr>
<td>gm09_138b</td>
<td>LFAS$_{UP}$</td>
<td>487.69 (25)</td>
<td>300.54 (500)</td>
<td>0.77</td>
<td>121.72 (1)</td>
<td>91.1 (25)</td>
</tr>
<tr>
<td>gm09_138b</td>
<td>MFAS$_{UP}$</td>
<td>317.36 (60)</td>
<td>330.05 (500)</td>
<td>0.51</td>
<td>167.52 (3)</td>
<td>110.6 (25)</td>
</tr>
<tr>
<td>gm09_156b</td>
<td>LFAS$_{DN}$</td>
<td>254.05 (7)</td>
<td>292.42 (236)</td>
<td>0.34</td>
<td>93.93 (1)</td>
<td><strong>63.47 (12)</strong></td>
</tr>
<tr>
<td>gm09_156b</td>
<td>LFAS$_{UP}$</td>
<td>207.53 (28)</td>
<td>260.54 (236)</td>
<td>0.35</td>
<td>22.26 (1)</td>
<td><strong>42.31 (12)</strong></td>
</tr>
<tr>
<td>gm09_156b</td>
<td>MFAS$_{UP}$</td>
<td>556.01 (12)</td>
<td>281.48 (236)</td>
<td>0.77</td>
<td>26.85 (1)</td>
<td><strong>43.25 (12)</strong></td>
</tr>
</tbody>
</table>
Figure 1. Time series of the dissimilarity metric (m) between contours of whale calls and the MFAS signal (top) and depth (bottom) for the whole tag record for whale gm08_150c (Table 1). Each '+' symbol represents a single call. Full line box indicates the period of MFAS exposure. Calls during LFAS exposure (dotted line box) were not used in the rotation test. No good quality calls were found between the MFAS and LFAS exposures. See figure 2 for description of numbers 1 and 2.
Figure 2. Spectrogram of whale and naval sonar sounds. Top two panels show calls marked on figure 1. Example 1 (top left) is a call with a low dissimilarity value (m = 29 Hz) and example 2 (top right) is a call with a high dissimilarity value (m = 1240 Hz). Both calls were produced before any sonar exposure. Bottom panels show examples of MFASUP (bottom left) and LFASUP (bottom right) sonar signals as recorded on the Dtag. Color lines represent the extracted contours for each of the examples.

203x152mm (150 x 150 DPI)
Figure 3. Example spectrogram from deployment gm09_138b during transmission of LFAS downsweep sonar signals. Red colour represents the extracted sonar contour while green and blue colours represent extracted contours of calls produced by pilot whales. White insets show the offset position of each of the calls contour in relation to the sonar contour corresponding to the minimum dissimilarity metric value.

442x245mm (96 x 96 DPI)