
Downloaded from: http://insight.cumbria.ac.uk/1978/

Usage of any items from the University of Cumbria’s institutional repository ‘Insight’ must conform to the following fair usage guidelines.

Any item and its associated metadata held in the University of Cumbria’s institutional repository Insight (unless stated otherwise on the metadata record) may be copied, displayed or performed, and stored in line with the JISC fair dealing guidelines (available here) for educational and not-for-profit activities provided that

- the authors, title and full bibliographic details of the item are cited clearly when any part of the work is referred to verbally or in the written form
  - a hyperlink/URL to the original Insight record of that item is included in any citations of the work
- the content is not changed in any way
- all files required for usage of the item are kept together with the main item file.

You may not

- sell any part of an item
- refer to any part of an item without citation
- amend any item or contextualise it in a way that will impugn the creator’s reputation
- remove or alter the copyright statement on an item.

The full policy can be found here.
Alternatively contact the University of Cumbria Repository Editor by emailing insight@cumbria.ac.uk.
Call diversity in the North Pacific killer whale populations: implications for dialect evolution and population history

Olga A. Filatova\textsuperscript{a,*}, Volker B. Deecke\textsuperscript{b,c}, John K.B. Ford\textsuperscript{c}, Craig O. Matkin\textsuperscript{d}, Lance G. Barrett-Lennard\textsuperscript{e}, Mikhail A. Guzeev\textsuperscript{f}, Alexandr M. Burdin\textsuperscript{g}, Erich Hoyt\textsuperscript{h}

\textsuperscript{a}Department of Vertebrate Zoology, Faculty of Biology, Moscow State University, Russia
\textsuperscript{b}Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, UK
\textsuperscript{c}Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC
\textsuperscript{d}North Gulf Oceanic Society, Homer, Alaska, USA
\textsuperscript{e}Vancouver Aquarium Marine Science Centre, BC, Canada
\textsuperscript{f}Department of Vertebrate Zoology, Faculty of Biology and Soil Sciences, St. Petersburg State University, Russia
\textsuperscript{g}Kamchatka Branch of Pacific Institute of Geography FEB RAS, Petropavlovsk-Kamchatsky, Russia
\textsuperscript{h}Whale and Dolphin Conservation Society, UK

*Corresponding author: Olga A. Filatova, Department of Vertebrate Zoology, Faculty of Biology, Moscow State University, Vorobyovy gory 1/12, Moscow 119992, e-mail: alazor@rambler.ru, tel. +7 903 2366474.

Correspondence:
Volker B. Deecke: Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, Fife, KY16 8LB, UK
John K.B. Ford: Pacific Biological Station, 3190 Hammond Bay Road, Fisheries and Oceans Canada, Nanaimo, BC V9T 6N7
Craig O. Matkin: North Gulf Oceanic Society, 3430 Main St Suite B1, Homer, Alaska, AK99603, USA
Lance G. Barrett-Lennard: Vancouver Aquarium Marine Science Centre, P.O. Box 3232, Vancouver, BC, Canada V6B 3X8
Mikhail A. Guzeev: Department of Vertebrate Zoology, Faculty of Biology and Soil Sciences, Universitetskaya emb., 7/9, St. Petersburg State University, St. Petersburg 199034, Russia
Alexandr M. Burdin: Kamchatka Branch of Pacific Institute of Geography FEB RAS, Pr. Rybakov 19-a, Petropavlovsk-Kamchatsky, 683024, Russia
Erich Hoyt: 29A Dirleton Avenue, North Berwick, Scotland EH39 4BE, UK
Although killer whale (*Orcinus orca*) dialects have been studied in detail in several populations, little attempt has been made to compare dialect characteristics between populations. In this study we investigated geographical variation in monophonic and biphonic calls among four resident populations from the North Pacific Ocean: Southern and Northern Vancouver Island residents, southern Alaska residents, and eastern Kamchatka residents. We tested predictions about call variation across populations which are due to an accumulation of random errors and innovations by vertical cultural transmission. Call frequency contours were extracted and compared using a dynamic time-warping algorithm. We found that the diversity of monophonic calls was substantially higher than the diversity of biphonic calls for all populations. Repertoire diversity appeared to be related to the population size: in larger populations, monophonic calls were more diverse and biphonic calls were less diverse. We suggest that the evolution of both monophonic and biphonic calls is caused by an interaction between stochastic processes and directional selection, but the relative effect of directional selection is greater for biphonic calls. Our analysis revealed no direct correlation between call repertoire similarity and geographical distance. Call diversity within pre-defined call categories – types and subtypes – showed a high degree of correspondence between populations. Our results suggest that dialect evolution is a complex process influenced by an interaction between directional selection, horizontal transmission and founder effects. We suggest several scenarios for how this might have arisen and the implications of these scenarios for call evolution and population history. **Keywords:** dialect, killer whale, acoustic repertoire, evolution, call type.
Vocal variation among groups of animals may occur at different levels. Variations in vocalisations between neighbouring groups of potentially interbreeding individuals are called dialects, whereas differences in acoustic repertoires over long distances and between populations that normally do not interbreed are referred to as geographic variation (Conner 1982). Dialects are common in birds (Baker & Cunningham 1985), but rare in mammals, being mostly limited to cetaceans (e.g., Ford 1991; Rendell & Whitehead 2003), bats (e.g., Boughman 1997; Esser & Schubert 1998; Yoshino et al. 2008) and humans (Labov 2001). In contrast, geographic variation in vocal repertoires is common among both bird and mammalian populations (e.g., Krebs & Kroodsma 1980; Slobodchikoff et al. 1998; Mitani et al. 1999).

Killer whales are widely distributed throughout the world’s oceans (Forney & Wade 2007). Different populations display substantial variation in diet, behaviour, morphology and genetics (Ford et al. 1998; Pitman & Ensor 2003; Foote et al. 2009; Morin et al. 2010). Rather than being genetically coded, the vocal repertoire of killer whales is thought to be learned (Bowles et al. 1988; Ford 1991; Deecke 2000; Foote et al. 2006), which leads to formation of dialects between neighbouring groups in some populations and geographic variation between distant populations.

As a species, killer whales feed on a wide variety of prey, but different populations often show a high degree of dietary specialization (Ford et al. 1998; Saulitis et al. 2000; Ford & Ellis 2006). In several regions, sympatric populations show little or no dietary overlap and represent different ecotypes (Ford et al. 1998; Saulitis et al. 2000; Pitman & Ensor 2003). Three killer whale ecotypes have been described from the North Pacific and these differ in social structure, morphology, genetics and behaviour: residents specialize on fish and live in large stable social units (Ford & Ellis 2006; Ivkovitch et al. 2010), transients hunt primarily marine mammals and live in smaller more fluid social groups (Baird & Dill 1996; Ford et al. 1998), and offshores are probably fish specialists (Ford et al. 2011) and live in large groups with an unknown social structure.
Dialects have been described for several resident killer whale populations from the North Pacific (Ford 1991; Yurk et al. 2002; Filatova et al. 2007) and for killer whales from the northeastern Atlantic (Strager 1995). Resident killer whales from the North Pacific have a complex nested social structure comprised of 1) matrilines containing a matriarch and all her descendants, which always travel together; 2) pods containing a set of matrilines that associate frequently and use a common repertoire of stereotyped calls, which represents the vocal dialect of the pod; 3) acoustically distinct clans comprised of pods which share some repertoire calls; and 4) populations or communities containing one or more associating clans (Ford 1991; Ford 2002; Ivkovich et al. 2010).

Several distinct populations of resident killer whales have been identified in the North Pacific: Southern and Northern Vancouver Island residents, southern Alaskan residents in the northeastern Pacific (Ford 2002; Matkin et al. 1999), eastern Kamchatka residents in the northwestern Pacific (Ivkovich et al. 2010) and a number of less studied putative populations around the Aleutian and Kuril Islands and in the Bering and Okhotsk seas. Genetic and other research on these populations (Barrett-Lennard 2000; Hoelzel et al. 2002) has not, to date, provided detailed information about historical relationships among them.

It has been suggested that repertoires of stereotyped calls may serve as a marker of maternal relatedness. Divergence between vocal repertoires of killer whale pods is thought to happen gradually as pods grow bigger and matrilines spend less and less time together (Ford 1991). Vocal learning involves a series of call mistakes and innovations, which gradually make vocal repertoires diverge. For these reasons, Ford (1991) suggested a direct relationship between the maternal ancestry of different pods within clans and the degree of similarity of their vocal repertoires: the more distant the common maternal ancestry between pods, the fewer calls shared within their repertoires. This prediction was confirmed by Deecke et al. (2010), who showed that the similarity of one call type across matrilines was correlated with matriarch relatedness in spite of substantial male-mediated gene flow.
These findings suggest that populations with more recent common maternal ancestry must have more similar repertories as well. Although killer whale dialects have been described in detail in several populations, few studies have examined vocal variation between killer whale populations. Yurk (2005) compared distribution of call syllables among subpopulations and showed that distinct lower frequency syllables were shared between clans but differed between populations, whereas upper frequency syllables varied between clans of the same population. Foote & Nystuen (2008) showed that the frequency parameters of calls varied across ecotypes (resident, transient and offshore). In this paper, we investigate the variation of whole frequency contours across the three North Pacific resident killer whale populations.

Comparing killer whale acoustic repertoires is complicated by the fact that killer whale sounds are not structurally homogenous. Killer whale sounds comprise several distinct structural categories, common to all killer whale populations studied to date. Killer whale sounds include whistles, echolocation clicks and pulsed calls. Most pulsed calls are highly stereotyped and can be easily divided into call types (Ford 1991) with varying degrees of variability within types. Many call types have an overlapping, independently modulated high-frequency component (Fig. 1); this phenomenon is usually referred to as “biphonation” (Wilden et al. 1998; Fitch et al. 2002) or, when two independent sources are responsible, “two-voiced calling” (Zollinger et al. 2008). Since the mechanism responsible for this pattern in killer whales is unknown, we use the former term in this paper.

Biphonic sounds have been described in mammals as diverse as canids (Wilden et al. 1998; Riede et al. 2000, Volodin & Volodina 2002), primates (Fisher et al. 2001, Brown et al. 2003; Riede et al. 2004) and cetaceans (Tyson et al. 2007). While the functional significance of biphonation in calls is not readily understood, its presence in the vocalisations of different species suggests a potentially important communicative role. Proposed functions include the enhancement of individual recognition (Aubin et al. 2000; Fitch et al. 2002; Volodina et al. 2006), or honest signalling of physical condition (Fitch et al. 2002). For killer whales, it has been
suggested that differences in the directionality of the lower- and higher-frequency components in biphonic calls can provide information on the orientation of a caller relative to a listener (Miller 2002). Differences in usage of biphonic and monophonic calls in diverse social contexts suggest that they may have distinct functions in killer whale communication with biphonic calls functioning mostly as group identifiers and monophonic calls serving as short-range contact signals (Filatova et al. 2009). Moreover, biphonic and monophonic calls show substantial differences in source levels (Miller 2006) and structure (Filatova et al. 2007), which suggests that they should be considered two distinct structural categories.

In this study we examined geographical variation in monophonic and biphonic calls among four resident populations from the North Pacific Ocean: Southern and Northern Vancouver Island residents, southern Alaska residents, and eastern Kamchatka residents (Fig. 2). We tested predictions about call variation across populations which are due to an accumulation of random errors and innovations by vertical cultural transmission. First, we compared the diversity of monophonic and biphonic calls within these populations. Second, we compared the similarity of monophonic and biphonic calls between each pair of populations. Finally, we measured call diversity within pre-defined call categories – types and subtypes, and examined whether call diversity within these categories differed across populations.

**Methods**

**The study populations**

Southern Vancouver Island resident killer whales (referred to below as SR) inhabit the coastal waters of British Columbia and Washington State. The core area of this population is in the waters of southern Vancouver Island, but they sometimes range south to Monterey Bay, California. The population consists of the single acoustic clan – J-clan (Ford 1991), which comprised 86 individuals in 2003 (van Ginneken et al. 2005).

Northern Vancouver Island resident killer whales (referred to below as NR) inhabit the coastal waters of British Columbia and southeastern Alaska, from southern Vancouver Island...
north to southeastern Alaska (approximately 48°N to 58°N). The population is comprised of three acoustic clans: A-clan, G-clan and R-clan (Ford 1991).

Southern Alaskan resident killer whales (referred to below as AR) range from southeastern Alaska to Kodiak Island (Matkin et al. 1999) and possibly into the Aleutian Islands and the Bering Sea (Allen & Angliss 2010). This population includes two acoustic clans: AB-clan and AD-clan (Yurk 2002).

Eastern Kamchatka resident killer whales (referred to below as KR) were encountered along the eastern coast of Kamchatka peninsula from Avacha Gulf to Karaginsky Gulf and east to the Commander Islands. This population comprises three acoustic clans: Avacha clan, K19 clan and K20 clan (Filatova 2007).

Some overlap exists in the ranges of the northeastern Pacific populations: SR overlap with NR in the waters of Vancouver Island, and NR overlap with AR in southeastern Alaska (Ford et al. 2000). Despite this overlap, none of the populations have been observed to mix (Ford et al. 2000).

Data collection

Sound recordings used for this study were taken from the existing long-term databases. Recordings of Southern residents were made from 1980-2009, Northern residents from 1988-1999, southern Alaskan residents from 1984-2008, and eastern Kamchatka residents from 2000-2009. All recording systems had a flat frequency response from at least 0.1 to 7 kHz, although in most cases this extended up to 20 kHz. We only included recordings that had sufficient frequency bandwidth and signal-to-noise ratio to clearly display all call features.

The recordings were made from small (4-9 m) boats. Photographs were taken during all recording sessions and compared to identification catalogues to confirm pod and population identity, as described in Bigg et al. (1990).

In the recording sessions made directly for this study, all approaches to the whales were conducted following procedures to minimize disturbance. To take photographs, the boat
approached at slow speed at 45° to the whale’s course when they were traveling and left the
group immediately after the photographs of all group members were obtained. To make sound
recordings, we moved the boat 200-300 m ahead of the animals and waited until they passed us.
If the whales were feeding or milling, we stayed at a distance of 100-300 m from them to avoid
disturbing their natural behaviour. When the animals showed strong avoidance behaviour (e.g.
change in the direction of movement away from the boat), we stopped our activities and kept a
distance of at least 500 m from the group.

Acoustic and statistical analysis

We classified calls according to existing catalogues (Ford 1987; Yurk et al. 2002;
Filatova et al. 2004). For Alaska and Kamchatka, however, some call types were split and others
were added according to the results of more recent studies (see Filatova et al. 2007; Yurk et al.
2010). Two calls from each type/subtype were used for the analysis, with some exclusion of rare
call types for which we were not able to obtain at least two call samples of adequate quality. If a
call type had no subtypes, two samples from this call type were used. For call types that fell into
discrete subtypes, two samples from each subtype of this type were used. When possible, the
pairs of call samples from the same type/subtype were selected from different encounters and
different years to cover the presumed variation in the call structure. For calls which did not fall
into discrete subtypes but showed apparent group-specific variations (e.g., N12, see Ford 1991),
two calls from the opposite sides of the structural continuum were selected. In total, 348 samples
of 174 call types/subtypes were used for the analysis: 34 SR monophonic, 28 SR biphonic, 34
NR monophonic, 62 NR biphonic, 48 AR monophonic, 40 AR biphonic, 46 KR monophonic, 56
KR biphonic. Because our primary interest was call evolution, rare calls were of equal interest to
common calls, and we did not weight call similarities by the frequency with which each call
occurred.

Call contours were extracted using a custom-made MATLAB (The Mathworks, Inc.,
Natick, MA) routine for manually tracking frequency contours of each frequency component
(available online at www.russianorca.org/sound_pro.htm). After the operator selected enough points to track all modulations of the contour from the fundamental frequency and harmonics, the algorithm performed the generalization of frequency points by dividing them by the band number and joined them into a set of frequency measurements of the fundamental frequency, which were then smoothed and interpolated to produce a vector of frequency measurements with the sampling interval 0.01 s (Fig. 3). For biphonic calls, contours were extracted both from the low- and the high-frequency components. Calls were identified as biphonic if they contained the overlapping high-frequency component (Fig. 1).

Similarity of calls was measured using dynamic time-warping. Dynamic time-warping is an algorithm developed for the automated recognition of human speech that allows limited compression and expansion of the time axis of a signal to maximise frequency overlap with a reference signal (e.g., Itakura 1975). For this study, we used a modified version of the warping algorithm of Deecke & Janik (2006). Percent similarity of contours was calculated by dividing the smaller frequency value by the larger value at each point and multiplying by 100:

$$S(i) = \frac{\min[M(i), N(i)]}{\max[M(i), N(i)]}\times 100$$

where M is the reference contour and N the input contour. From the resulting similarity matrix, a cost matrix was constructed that kept a running tab on the similarities of the elements making up the curves while adding up these costs to give a final number called the “similarity” between the contours. In our algorithm, each element of the cost matrix was obtained by comparing the weighted sum of similarity values from two columns and two rows distant from the weighted diagonal.

Because the algorithm of Deecke & Janik (2006) only allows expansion or compression of the time axis by a factor of three, the algorithm cannot be used to compare calls that differ in length by more than a factor of three. In this case, their similarity is considered zero percent. This constraint biased the results in comparisons where many short or long contours were present in the repertoire of one population but not the other. To avoid this, we developed an
additional algorithm that stretched the shorter contour through interpolation to make it one point longer than 1/3 of the longer contour.

For each pair of contours within each sample set, we measured their relative similarity in frequency using this dynamic time-warping algorithm. By generating all possible pairwise comparisons between call samples from sample sets, we used N samples to generate 0.5*N*(N - 1) comparisons. Since the correlation structure of this data set was unknown, we assumed that all data points generated with the same sample were correlated. To achieve independence between the analysis units, we calculated the mean similarity for each call sample and used it as the unit of analysis for the further comparison.

For the measurements of **intra-population call similarity**, we calculated the similarity between each pair of calls from each major category (monophonic/biphonic) within each population.

For the comparison of **intra- and inter-population similarity of calls**, we calculated the inter-population similarity for each pair of populations as a set of similarity values between each call from the first population and each call from the second population. The median inter-population similarity obtained by this method would depend not only on the true similarity between populations, but also on the intra-population call similarity of each population. To get a less biased inter-population similarity measure, we divided the median inter-population similarity by the median of the pooled intra-population similarity values for each pair of populations.

To measure **call similarity within types and subtypes**, we divided each of the six intra-population sets of similarity values into the following three subsets: a) similarity values between pairs of calls from the different types; b) similarity values between pairs of calls from the same type; and c) similarity values between pairs of calls from the same subtype. To test for differences in the type/subtype threshold between monophonic and biphonic calls, the combined
sets from all three populations were used. To test the differences in the type/subtype threshold across populations, subsets b) and c) were compared for each pair of populations.

Statistical analysis was performed using R software (R Development Core Team 2010). Distribution of similarity values in most cases differed significantly from normal, so we used the non-parametric two-tailed Mann-Whitney U-test for all statistical comparisons. Bonferroni correction was applied in cases of multiple pairwise comparisons. A Mantel matrix permutation test (Schnell et al. 1985) was used to estimate the correlation of call similarity and geographical distance between populations. Distances were approximate based on the core summer distribution as the full extent of population ranges is unknown. In all statistical comparisons the significance level was accepted to be 0.05.

**Results**

**Intra-population call similarity of monophonic and biphonic calls**

We compared levels of similarity among monophonic calls with the levels of similarity among biphonic calls. Within each of the four populations, call similarity was significantly less among monophonic calls than it was among the biphonic calls in each population (monophonic vs biphonic, Mann-Whitney U-test, KR: $U = 109, N_1 = 46, N_2 = 56, p < 0.0001$; AR: $U = 89, N_1 = 48, N_2 = 40, p < 0.0001$; NR: $U = 98, N_1 = 34, N_2 = 62, p < 0.0001$) except SR in which the difference was close to significant ($U = 338, N_1 = 34, N_2 = 28, p = 0.051$). Differences between populations within these two categories were less pronounced (Fig. 4) though also significant in all cases except AR vs KR and NR vs SR monophonic, AR vs NR biphonic (Table 2).

SR had the highest median similarity (that is, the lowest call diversity) for monophonic calls followed by NR, AR and KR. KR had the highest median similarity for biphonic calls followed by AR, NR and SR. Including the higher-frequency component into the analysis of biphonic calls increased the call similarity within all populations (Table 1). In this case, KR again had the highest median similarity, followed by NR, AR and SR.
Intra- and inter-population similarity of calls

We compared intra- and inter-population similarity of monophonic and biphonic calls in the each pair of populations (Table 3). Differences in intra- and inter-population similarity of monophonic calls were non-significant for all comparisons. Intra-population similarity of biphonic calls was significantly higher than inter-population similarity for all comparisons (Table 3).

The fact that the inter-population similarity of monophonic calls did not differ significantly from the intra-population similarity means that monophonic calls are equally diverse within and between populations. For this reason we did not compare the inter-population similarity of monophonic calls between pairs of populations.

Inter-population similarity of biphonic calls was the highest between AR and KR, followed by KR-SR, NR-KR, NR-SR, SR-AR and NR-AR (Table 3). After dividing this value by the intra-population similarity to obtain the “true” similarity measure, SR and NR were the most similar, followed by NR-KR, SR-AR, AR-NR, AR-KR and SR-KR (Fig. 5A).

Inter-population similarity of biphonic calls compared by analyzing both the lower-frequency and the higher-frequency components was the highest between KR and NR, followed by KR-AR, SR-AR, NR-AR, SR-NR and SR-KR (Table 3). After dividing this value by the intra-population similarity, SR and AR were the most similar, followed by NR-KR, AR-NR, SR-NR, AR-KR and SR-KR (Fig. 5B). The correlation between approximate geographical distance and call similarity measured by the lower-frequency component and by both the lower-frequency and the higher-frequency components was non-significant (Fig. 5).

Call similarity within types and subtypes

We measured the similarity between monophonic and biphonic calls from different types, calls from the same type, and calls from the same subtype for each population (Table 4). Similarities between calls from different types followed the pattern revealed by the comparison of call similarities within populations: similarities between monophonic calls were significantly
lower than between biphonic calls ($U = 826, N_1 = 86, N_2 = 71, p < 0.0001$; Fig. 6). The
difference in similarity between monophonic and biphonic calls within types and subtypes was
non-significant. The differences in similarity of calls within types and subtypes across different
populations were non-significant.

Discussion

The comparison of monophonic and biphonic calls from the four North Pacific resident
killer whale populations revealed a pronounced difference in the degree of similarity between
monophonic and biphonic calls for each population. For all four populations, the diversity of
monophonic calls was higher than the diversity of biphonic calls. Differences between
populations in the diversity of call types from the corresponding category (monophonic or
biphonic) were less pronounced. This result suggests that monophonic and biphonic calls have
different principles of evolution that are shared among the different resident populations. This
supports the suggestion that monophonic and biphonic calls are discrete categories.

Our results are consistent with previous studies which found differences in source levels
(Miller 2006), directionality (Miller 2002) and usage (Filatova et al. 2009) between monophonic
and biphonic calls. Miller (2006) measured source levels of different killer whale sounds and
showed that monophonic calls exhibited mean source levels lower than biphonic calls. This
variation in intensity suggests that killer whale pulsed calls fall into two functional groups:
“long range” biphonic calls with a mean estimated active space of 10–16 km in sea state zero
and “short-range” monophonic calls with an active space of 5–9 km (Miller 2006). Miller
(2002) showed that the relative energy in the high-frequency components of biphonic calls was
significantly greater when animals were moving toward the hydrophone array than away from it.
It is likely that this difference could help listening whales to determine the direction of
movement of a caller. Filatova et al. (2009) showed that the proportion of biphonic calls in the
vocalisations of the eastern Kamchatka residents increased when more than one pod was present
in the area. The combination of these findings suggests that biphonic calls function mostly as

338 group identifiers and help whales to define group affiliation and monitor the position of group

339 members over long ranges. The function of monophonic calls is less clear, but, like killer whale

340 whistles (Thomsen et al. 2002), they may serve as short-range communication signals.

341 The similarity of diversity levels in monophonic and biphonic calls in four populations

342 raises the question of whether the repertoire structure in resident killer whales is genetically or

343 culturally inherited. It is now generally accepted that killer whales acquire the detailed structure

344 of stereotyped calls in their repertoire through vocal learning (Bowles et al. 1988; Ford 1991:

345 Deecke et al. 2000; Foote et al. 2006); however, other aspects of the vocal repertoire could be

346 innate. Many songbirds learn their songs from fathers or neighbouring males, but despite some

347 variation their song remains species-specific and retains a certain structure. Moreover, gradual

348 differences in the function, usage and raw structural difference of the songs versus calls appear

349 to be innate in songbirds (Marler 2004). The same is probably true for human languages:

350 although languages are learned and therefore extremely diverse, there is some inherited structure

351 common to all human languages (Pinker 1994; but see Evans & Levinson 2009).

352 The comparison of stereotyped call repertoires among populations produced rather

353 unexpected results. The classical theory of call change through random drift predicts that the

354 highest call similarity should occur in populations that are geographically close and therefore

355 may share the recent common ancestors. However, our analysis revealed no direct correlation

356 between call repertoire similarity and geographical distance (Fig. 5). There are several possible

357 explanations for this. First, it is possible that call evolution is too fast to be phylogenetically

358 meaningful on a population scale: call repertoires may be already so diverse that they retain no

359 signs of common ancestry, and all similarities could be the result of random convergence. It has

360 been suggested in resident killer whales that females choose mates with the most dissimilar

361 dialects (Barrett-Lennard 2000). This would drive sexual selection towards faster call evolution

362 to provide greater resolution in the recognition of kin (e.g., allowing discrimination between first
and second degree cousins). The opposite selection force may be caused by the need of killer whale groups to possess markers of population identity. Biphonic calls of all three clans of the eastern Kamchatka resident population have certain frequency features in common (Filatova et al. 2007). In southern Alaska resident killer whales, distinct lower frequency syllables are shared by clans within this population (Yurk 2005). In Northern residents, most stereotyped whistle types are structurally identical in two of the three acoustic clans (Riesch et al. 2006). Northern and Southern resident populations in British Columbia share a substantial part of their geographical range but retain strong behavioural reproductive isolation, although it is not obvious if the isolation is based on acoustic or other cues.

The interaction of these opposite evolutionary forces (diversifying and standardizing) may lead to the “maximum diversity within the permitted range”, where the “range” is represented by vocal population markers. This scenario is consistent with the fact that in our study the diversity of biphonic and monophonic calls between populations was very similar to that within populations, suggesting that every population had already reached some optimal level of diversity for each call category. Moreover, the intra- and inter-population diversity of monophonic calls did not differ significantly, suggesting that for monophonic calls the pressure to standardise is lower.

It is interesting to note that the diversity of monophonic and biphonic calls appears to be negatively correlated. This pattern is also related to the population size: monophonic calls are more diverse and biphonic calls are less diverse in larger populations. Although the SR population of about 86 animals (van Ginneken et al. 2005) is the smallest of the four populations, it has the highest diversity of biphonic calls and the lowest diversity of monophonic calls, followed by the NR with 216 animals (Ford et al. 2000), KR (650 individuals; T.V. Ivkovich unpublished data) and AR (more than 1000 individuals; Allen & Angliss 2010). In human languages, speaker population size was shown to be a significant predictor of phonemic diversity, with a smaller population size predicting smaller overall phoneme inventories.
Our results demonstrate that the diversity of monophonic calls follows the same pattern as phonemic diversity in human languages, which suggests that their evolution is driven by the same stochastic processes that also affect human phonemes (Labov 2001). By contrast, diversity in biphonic calls shows the opposite pattern. This may be caused by the fact that in larger populations the inbreeding risk is lower, and the need for unique vocal population markers is higher, which shifts the balance of diversifying and standardizing forces in favour of the latter. Therefore, we suggest that the evolution of both monophonic and biphonic calls is caused by an interaction between stochastic processes and directional selection, but the relative effect of directional selection is greater for biphonic calls.

An alternative hypothesis suggests that call similarity between populations does reflect their ancestry, but the ancestry is not directly correlated with geographical distance. Killer whales are highly mobile and phylogeographic structure could very easily be disturbed by long-distance movements. For example, genetic studies showed that the resident killer whales of the North Pacific are more related to the North Atlantic killer whales than to sympatric transient populations (Morin et al. 2010). This suggests a complex and multi-stage history of population formation and colonisation of the North Pacific Ocean. Moreover, the southern Alaskan resident population possesses two haplotypes of the control region of mitochondrial DNA; one matches with the single haplotype of the Northern resident population, the other with the single haplotype of the Southern resident population (Barrett-Lennard 2000) and the eastern Kamchatkan resident population (Hoelzel et al. 2007). It is possible that the AR population retains genetic diversity of an ancestral population, while the NR, SR and KR populations have reduced diversity due to a founder effect. Founder effects have been shown to reduce the diversity of syllables in bird songs (Baker & Jenkins 1987) and phonemic diversity in human languages (Atkinson 2011). This is in agreement with the reduced diversity in KR biphonic and NR and SR monophonic calls, but it is in contradiction to the high diversity in SR biphonic calls. However, the information currently available does not allow testing of this hypothesis, and further study of genetic and acoustic
similarity is required to reveal the population history of North Pacific killer whales. A comparison of repertoire similarity with mitochondrial haplotype similarity across populations can provide an important insight into population history and dialect evolution.

Similarities of biphonic stereotyped call repertoires measured by the low-frequency component and by both the low- and the high-frequency components were generally not consistent, although they agreed in some aspects (Fig. 5 A, B). It appears that the evolution of low-frequency and high-frequency components of biphonic calls is not always parallel, suggesting that it may be influenced by different factors. The inclusion of the high-frequency component in the analysis of the intra-population similarity always increased the similarity values (Table 1), indicating that the high-frequency component is less diverse within populations than the low-frequency component. It is possible that the high frequency component may be a more stable and reliable marker of population relatedness, than the more diverse and variable low-frequency component. Alternatively, the stability of the high-frequency component may be related to the radiation pattern of biphonic calls. The higher-frequency component is more directional than the lower-frequency component (Miller 2002), so it is clearly audible only when the signaller is orientated towards the receiver. Therefore, the more omnidirectional lower-frequency component would appear to be more useful for the long-range recognition of pod members. This may result in the higher contour variability in the lower-frequency component to make the call more discernible. Consequently, the type-specific variation in the contour shape of the higher-frequency component may be redundant because the call type is already identifiable by the lower-frequency component. In that case, the function of the higher-frequency component may be restricted to marking the orientation of a signaler while the lower-frequency component marks the pod membership.

The diversity of calls within type and subtype categories had a high degree of correspondence between populations, despite the fact that the initial categorisations were made by different researchers (Ford 1991; Yurk 2002; Filatova 2004). Unlike the graded vocalisations
of some other odontocetes (e.g., Weilgart & Whitehead 1990), killer whale pulsed calls are
highly stereotyped showing little variation within call types, but there are consistent differences
between them. Differences within type/subtype diversity between populations were non-
significant. Moreover, the difference in diversity between monophonic and biphonic calls was
non-significant within types and subtypes. This suggests that the observers had rather similar
ideas of what they meant by call type and subtype. No one has yet provided a satisfactory
definition of “call type” in killer whales, and the most common description of the categorisation
process refers to “the distinctive audible characteristics of the calls”. Call structure changes
subtly but continuously over time (Deecke et al. 2000), and call type divergence is thought to be
a gradual process (Bigg et al. 1990; Ford 1991). Consequently, calls of different matrilines can
differ to a greater or lesser extent, and it is not always obvious where to place the border. Deecke
and Janik (2006) performed an automatic neural network categorisation of calls recorded from
North Pacific transient killer whales and identified 8 monophonic and 5 biphonic call types. Our
study provides an equally objective approach to call categorisation by comparing contours and
using a threshold similarity level to delineate call types.

In conclusion, our results suggest that divergence of vocal repertoires may not result
solely from the accumulation of random errors and innovations by vertical cultural transmission.
Repertoire diversity appears to be related to the population size: monophonic calls are more
diverse and biphonic calls are less diverse in larger populations. Call similarity across
populations does not correspond with geographical distance. All this suggests that dialect
evolution is a complex process subject to an interaction between directional and non-directional
agents of structural change. These may include opposing selecting forces to diversify and
standardise vocal repertoires, horizontal transmission of calls, as well as random drift.

Acknowledgements

This research was supported by the Russian Fund for Fundamental Research (grant 11-
04-00460-a), the Whale and Dolphin Conservation Society, Vancouver Aquarium and
Department of Fisheries and Oceans Canada. The North Gulf Oceanic Society provided support for collection of recordings in Alaska. VBD was supported by a Marie-Curie Intra-European Fellowship. We are grateful to Harald Yurk for providing the recordings of Alaskan killer whales and supervision in their classification. We are grateful to Mayuko Otsuki for selecting the samples of Southern resident calls.

References


Yurk, H., Filatova, O. A., Matkin, C. O., Barrett-Lennard, L. G. & Brittain, M.

2010. Sequential habitat use by two resident killer whale (*Orcinus orca*) clans in Resurrection Bay, Alaska, as determined by remote acoustic monitoring. *Aquatic Mammals*, 36, 67-78.

Table 1. Median time-warped frequency contour similarity (in %) between monophonic and biphononic call types of different resident killer whale populations in the North Pacific, calculated by the low-frequency component (LF) and by both low- and high-frequency components (LF+HF).

<table>
<thead>
<tr>
<th>Population</th>
<th>Median similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>monophonic</strong></td>
<td></td>
</tr>
<tr>
<td>KR</td>
<td>45.75</td>
</tr>
<tr>
<td>AR</td>
<td>46.58</td>
</tr>
<tr>
<td>NR</td>
<td>54.47</td>
</tr>
<tr>
<td>SR</td>
<td>55.01</td>
</tr>
<tr>
<td><strong>biphonic LF</strong></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>65.08</td>
</tr>
<tr>
<td>SR</td>
<td>60.55</td>
</tr>
<tr>
<td>KR</td>
<td>72.16</td>
</tr>
<tr>
<td>AR</td>
<td>66.38</td>
</tr>
<tr>
<td><strong>biphonic LF+HF</strong></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>72.36</td>
</tr>
<tr>
<td>NR</td>
<td>72.87</td>
</tr>
<tr>
<td>SR</td>
<td>71.19</td>
</tr>
</tbody>
</table>
Table 2. Results of Mann-Whitney U-test comparison of intra-population call similarity of monophonic and biphonic calls across different resident populations.

<table>
<thead>
<tr>
<th>Populations</th>
<th>U</th>
<th>N₁</th>
<th>N₂</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR-KR</td>
<td>1068</td>
<td>48</td>
<td>46</td>
<td>0.789</td>
</tr>
<tr>
<td>AR-NR</td>
<td>403</td>
<td>48</td>
<td>34</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>NR-KR</td>
<td>322</td>
<td>34</td>
<td>46</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SR-NR</td>
<td>447</td>
<td>34</td>
<td>34</td>
<td>0.109</td>
</tr>
<tr>
<td>SR-KR</td>
<td>416</td>
<td>34</td>
<td>46</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SR-AR</td>
<td>429</td>
<td>34</td>
<td>48</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>AR-KR</td>
<td>606</td>
<td>40</td>
<td>56</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>AR-NR</td>
<td>1205</td>
<td>40</td>
<td>62</td>
<td>0.814</td>
</tr>
<tr>
<td>NR-KR</td>
<td>915</td>
<td>62</td>
<td>56</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SR-NR</td>
<td>345</td>
<td>28</td>
<td>62</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SR-KR</td>
<td>217</td>
<td>28</td>
<td>56</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SR-AR</td>
<td>292</td>
<td>28</td>
<td>40</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
Table 3. Results of Mann-Whitney U-test comparison of intra- and inter-population similarity, median time-warped frequency contour similarity (in %) and “true” median similarity between repertoires of different resident killer whale populations in the North Pacific.

<table>
<thead>
<tr>
<th>Populations</th>
<th>Mann-Whitney test</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U</td>
<td>N\textsubscript{intra}</td>
</tr>
<tr>
<td>AR-KR</td>
<td>2188</td>
<td>94</td>
</tr>
<tr>
<td>AR-NR</td>
<td>1634</td>
<td>82</td>
</tr>
<tr>
<td>NR-KR</td>
<td>1727</td>
<td>80</td>
</tr>
<tr>
<td>SR-NR</td>
<td>909</td>
<td>68</td>
</tr>
<tr>
<td>SR-KR</td>
<td>1663</td>
<td>80</td>
</tr>
<tr>
<td>SR-AR</td>
<td>1717</td>
<td>82</td>
</tr>
<tr>
<td>AR-KR</td>
<td>2037</td>
<td>96</td>
</tr>
<tr>
<td>AR-NR</td>
<td>1614</td>
<td>102</td>
</tr>
<tr>
<td>NR-KR</td>
<td>2825</td>
<td>118</td>
</tr>
<tr>
<td>SR-NR</td>
<td>1719</td>
<td>90</td>
</tr>
<tr>
<td>SR-KR</td>
<td>1480</td>
<td>84</td>
</tr>
<tr>
<td>SR-AR</td>
<td>837</td>
<td>68</td>
</tr>
<tr>
<td>AR-KR</td>
<td>1447</td>
<td>96</td>
</tr>
<tr>
<td>AR-NR</td>
<td>1590</td>
<td>102</td>
</tr>
<tr>
<td>NR-KR</td>
<td>2567</td>
<td>118</td>
</tr>
<tr>
<td>SR-NR</td>
<td>1455</td>
<td>90</td>
</tr>
<tr>
<td>SR-KR</td>
<td>647</td>
<td>84</td>
</tr>
<tr>
<td>SR-AR</td>
<td>874</td>
<td>68</td>
</tr>
</tbody>
</table>
Table 4. Median and mean (in parentheses) time-warped frequency contour similarity (in %) between all calls within the vocal repertoire belonging to the different call types, between calls belonging to the same call type and calls of the same subtype for different resident killer whale populations in the North Pacific. For biphonic calls, similarity of the low-frequency component and both components combined are given separately.

<table>
<thead>
<tr>
<th>Population</th>
<th>different types</th>
<th>same types</th>
<th>same subtypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>44.99 (45.87)</td>
<td>89.19 (84.03)</td>
<td>91.40 (89.95)</td>
</tr>
<tr>
<td>AR</td>
<td>45.53 (44.82)</td>
<td>90.44 (85.17)</td>
<td>91.83 (86.26)</td>
</tr>
<tr>
<td>NR</td>
<td>51.40 (50.12)</td>
<td>85.34 (82.47)</td>
<td>90.79 (90.04)</td>
</tr>
<tr>
<td>SR</td>
<td>53.51 (52.88)</td>
<td>91.56 (86.77)</td>
<td>92.84 (91.74)</td>
</tr>
<tr>
<td>KR</td>
<td>71.15 (68.34)</td>
<td>92.62 (91.26)</td>
<td>93.87 (92.04)</td>
</tr>
<tr>
<td>AR</td>
<td>65.88 (63.61)</td>
<td>88.69 (84.40)</td>
<td>93.76 (92.17)</td>
</tr>
<tr>
<td>NR</td>
<td>64.46 (64.97)</td>
<td>90.73 (89.10)</td>
<td>93.45 (91.45)</td>
</tr>
<tr>
<td>SR</td>
<td>59.30 (57.16)</td>
<td>87.66 (76.44)</td>
<td>93.97 (91.26)</td>
</tr>
<tr>
<td>KR</td>
<td>77.42 (75.05)</td>
<td>92.88 (91.43)</td>
<td>93.53 (92.51)</td>
</tr>
<tr>
<td>AR</td>
<td>72.21 (70.97)</td>
<td>89.63 (89.66)</td>
<td>94.74 (93.97)</td>
</tr>
<tr>
<td>NR</td>
<td>72.44 (71.74)</td>
<td>92.51 (90.69)</td>
<td>95.08 (92.58)</td>
</tr>
<tr>
<td>SR</td>
<td>70.35</td>
<td>82.94</td>
<td>95.40</td>
</tr>
</tbody>
</table>
1. Spectrograms of monophonic (left) and biphonic (right) calls. Note the low-frequency component (LFC) with multiple harmonics in both sounds and the high-frequency component (HFC) which is not a multiple of the LFC in the biphonic call.

2. Map of the North Pacific Ocean showing the home ranges of the resident killer whale populations investigated in this study.

3. Example of frequency contours extracted from the low-frequency component (LFC) and the high-frequency component (HFC) of a K27 call from the eastern Kamchatka resident population.

4. Intra-population time-warped frequency contour similarity between monophonic and biphonic calls of four resident killer whale populations. Horizontal lines represent medians, boxes interquartiles, and whiskers a 90% confidence interval.

5. “True” similarity (inter-population divided by intra-population time-warped frequency contour similarity) of biphonic stereotyped call repertoires plotted by the approximate geographical distance between four resident killer whale populations. (a) – similarity measured by the low-frequency component; (b) - similarity measured by both the low- and the high-frequency components.

6. Time-warped frequency contour similarity between monophonic and biphonic calls from different types, from same types and from same subtypes. Horizontal lines represent medians, boxes interquartiles, and whiskers a 90% confidence interval.