Cultural evolution of killer whale calls:
background, mechanisms and consequences

Olga A. Filatova a,b,*, Filipa I.P. Samarra a,c, Volker B. Deecke d, John K.B. Ford e, Patrick J.O. Miller a and Harald Yurk f

a Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY168LB, Scotland
b Department of Vertebrate Zoology, Faculty of Biology, Moscow State University, Moscow 119991, Russia
c Marine Research Institute, Skulagata 4, 121 Reykjavik, Iceland
d Centre for Wildlife Conservation, Lake District Campus, University of Cumbria, Rydal Road, Ambleside, Cumbria LA229BB, UK
e Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, BC Canada V9T1K6
f JASCO Research Ltd, 2305-4464 Markham Street, Victoria, BC, Canada V8Z7X8

*Corresponding author’s e-mail address: alazorro@gmail.com

Accepted 2 September 2015; published online 26 September 2015

Abstract
Cultural evolution is a powerful process shaping behavioural phenotypes of many species including our own. Killer whales are one of the species with relatively well-studied vocal culture. Pods have distinct dialects comprising a mix of unique and shared call types; calves adopt the call repertoire of their matriline through social learning. We review different aspects of killer whale acoustic communication to provide insights into the cultural transmission and gene-culture coevolution processes that produce the extreme diversity of group and population repertoires. We argue that the cultural evolution of killer whale calls is not a random process driven by steady error accumulation alone: temporal change occurs at different speeds in different components of killer whale repertoires, and constraints in call structure and horizontal transmission often degrade the phylogenetic signal. We discuss the implications from bird song and human linguistic studies, and propose several hypotheses of killer whale dialect evolution.

Keywords
killer whale, dialect, culture, cultural evolution, gene-culture coevolution.
1. Introduction

Cultural evolution of behavioral traits has attracted the attention of researchers over the past 50 years, yielding several models of cultural transmission inspired by ideas from evolutionary biology (Dawkins, 1976; Mundinger, 1980; Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985; Lumsden & Wilson, 1985; Laland et al., 2000; Lachlan & Feldman, 2003). Later research identified that cultural transmission in animals is more widespread and significant than recognized earlier (Whiten et al., 2011) and there may be a variety of processes and patterns of cultural propagation that may not have any close biological analogue, such as the potential for rapid acquisition and loss of learnt phenotypes (Claidière et al., 2014; Strimling et al., 2009).

A model of cultural evolution typically includes various transmission modes, such as vertical (from parent to offspring), oblique (from a member of the parental generation other than the parent) and horizontal (among members of the same generation), and several evolutionary processes, typically mutations, drift and selection (Cavalli-Sforza & Feldman, 1981; Lynch, 1996). In a theoretical attempt to find an analogue to evolution of genetically inherited traits, Dawkins (1976) suggested to use the term ‘meme’ to refer to a unit of cultural evolution, analogous to ‘gene’ in genetic evolution. In Dawkins proposal, cultural mutations are transformations in meme structure due to random transmission errors or innovations. The frequency of different memes in populations may vary due to stochastic cultural drift if a meme is selectively neutral, or due to selection if memes have differing adaptive value. A definition of culture as an evolutionary process that specifically addresses the transmitted information was provided by Boyd & Richerson (2005) who defined culture as “information capable of affecting individuals’ behaviour that they acquire from other members of their species through teaching, imitation, and other forms of social transmission”.

Non-human animals do not have material culture approaching that of humans in complexity, transmission accuracy, and repeated refinement of transmitted traits (cumulative culture, Marshall-Pescini & Whiten, 2008), but some non-material cultural variation such as socially learned systems of acoustic signals is rather complex. The comparative study of these systems can provide insights into human language evolution. The best-studied example of such a system is bird song (Catchpole & Slater, 1995; Lynch, 1996; Marler & Slabbekoorn, 2004). Many songbirds acquire species-specific song
through vocal learning, and if birds are not exposed to the species template or are unable to hear their own song they lack the ability to produce a proper species specific song (e.g., Brainard & Doupe, 2002).

In mammals vocal learning is rare, and cetaceans are one of the few mammalian groups that are capable of it (Janik & Slater, 1997; Janik, 2014). Social learning of acoustic signals may lead to complex well-developed vocal traditions changing with time through cultural evolution (Rendell & Whitehead, 2001).

One of the species with relatively well-studied vocal culture is the killer whale (*Orcinus orca*). Killer whales occur in all oceans, but are differentiated into many populations and ecotypes. Some of these local populations display substantial variation in diet, behaviour, morphology and genetics (Ford et al., 1998; Saulitis et al., 2000; Pitman & Ensor, 2003; Foote et al., 2011) having led some researchers to propose calling them separate killer whale species (Morin et al., 2010).

The best-studied killer whale populations inhabit the coastal waters of the N Pacific. Here, three distinct killer whale assemblies or ecotypes have been described: Residents specialize on fish, live mainly in coastal waters and usually travel in large stable social units of closely related animals (Ford et al., 2000; Ivkovich et al., 2010); Transients or Bigg’s killer whales hunt primarily marine mammals and travel in smaller, more fluid social groups (Baird & Dill, 1996; Ford et al., 1998), and Offshores appear to specialize on sharks in their diet (Ford et al., 2011a) and typically occur in large groups (50+) with an unknown social structure. As the name indicates these whales are usually encountered further away from shore.

In the NE Atlantic there is evidence for two partially sympatric populations that may specialize on different prey (herring and mackerel), and one parapatric population in lower latitudes (Foote et al., 2011). Long-term studies in the Southern Hemisphere have led to the discovery of further populations and ecotypes (e.g, Antarctic: Pitman & Ensor, 2003; Pitman et al., 2007, 2011; Foote et al., 2013; Argentina: Iñiguez, 2001; Crozet Islands: Guinet & Bouvier, 1995; New Zealand: Constantine et al., 1998; Visser, 1999; Visser et al., 2000).

Genetic studies have shown that the N Pacific resident, offshore and transient ecotypes, and the NE Atlantic herring eating killer whales fall into monophyletic DNA clades (Morin et al., 2010; Moura et al., 2014). Each ecotype appears to have expanded from a small matrilineal founder group.
tens of thousands of years ago (A. Foote, pers. comm.). Expansions may lead to fixation of rare alleles found by chance within the founder group, a process called allelic surfing (Excoffier et al., 2009) that may have driven genome-wide divergence between killer whale ecotypes. This mechanism combined with local adaptations could have been the cause of rapid development of so many ecotypes in this species. In areas where ecotypes occur in sympatry reduced competition among them likely stabilized the persistence of the ecotype division and could be a possible driver of stable acoustic divergence.

Most current knowledge on killer whale acoustic communication comes from the populations in the N Pacific and the NE Atlantic, while comparatively less is known about the structure and usage of vocalizations by the other populations (Table 1).

In this article, we will review different aspects of killer whale acoustic communication to provide insights into the processes that lead to variation in stereotyped calling, which produces the extreme diversity of group and population repertoires.

2. Vocal learning

Killer whale calls as a behavioral phenotype are thought to be transmitted socially via learning rather than genetically. Resident killer whales live in stable matrilineal units that may group together, forming subpods and pods (Ford et al., 2000; Ivkovich et al., 2010). Different pods produce distinct repertoires comprising a mix of unique and shared stereotyped call types, often referred to as a dialect. Calves adopt the call repertoire of their matrilineal unit, though their fathers may belong to another matrilineal unit, often from a different pod (Barrett-Lennard, 2000; Ford et al., 2011b), and therefore have a different call repertoire than their offspring. If dialects or call repertoires were transmitted genetically through genes from both parents, offspring repertoires should be somehow intermediate between the mother’s and father’s dialect.

Furthermore, a large amount of inferential evidence in favour of a cultural instead of genetic transmission of killer whale dialects has been gathered over the last two decades. Deecke et al. (2000) showed that structural modifications to stereotyped call types occurred simultaneously in two matriline that did not interbreed. The capability of vocal learning was also raised by Foote et al. (2006) suggesting that based on recordings made in the presence of an abandoned juvenile killer whale in waters off Vancouver Island
Table 1.
Summary of studies of killer whale acoustic communication.

<table>
<thead>
<tr>
<th>Location</th>
<th>Ecotype</th>
<th>Reference</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Function</td>
</tr>
<tr>
<td>NE Pacific</td>
<td>Residents</td>
<td>Hoelzel &amp; Osborne (1986)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ford (1991)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Barrett-Lennard et al. (1996)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yurk et al. (2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thomsen et al. (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thomsen et al. (2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deecke et al. (2000, 2010)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Miller &amp; Bain (2000)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Miller et al. (2004)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foote et al. (2004)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foote et al. (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nousek et al. (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Riesch et al. (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Miller et al. (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Riesch et al. (2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rehn et al. (2007, 2011)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foote &amp; Nystuen (2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foote et al. (2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Holt et al. (2008, 2013)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wieland et al. (2010)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yurk et al. (2010)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grebner et al. (2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Filatova et al. (2012a)</td>
<td></td>
</tr>
<tr>
<td>Transients</td>
<td>Barrett-Lennard et al. (1996)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deecke et al. (2005)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Saulitis et al. (2005)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Offshores</td>
<td>Foote &amp; Nystuen (2008)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Rehn et al. (2011)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Riesch &amp; Deecke (2011)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Foote &amp; Nystuen (2008)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Rehn et al. (2011)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Filatova et al. (2012b)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Simonis et al. (2012)</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>
this whale produced calls similar to barks of sea lion, a species that frequents the area but was apparently absent during the recordings. Deecke et al. (2010) provide additional evidence for cultural rather than genetic transmission: the correlations between call similarity and association patterns were much stronger than the correlations between call similarity and genetic relatedness.
Captive studies suggest that killer whales are able to imitate calls of their unrelated tank mates. For example, Bain (1986) described a 10-year-old female captured off Iceland mimicking the calls from a 13-year-old female captured in Canada after sharing a tank for several years. Captive killer whales that were housed with bottlenose dolphins were reported to produce dolphin type chirps (Musser et al., 2014). Bowles (1988) found that a captive killer whale calf learned calls selectively from its mother, while Crance et al. (2014) demonstrated that two captive juvenile male killer whales learned new calls and altered their repertoires to match that of an unrelated adult male with whom they were most strongly associated. An adult female housed in the same facility and the only NE Pacific whale among Icelandic ones, however, did not learn any Icelandic calls (Crance et al., 2014). It is possible that killer whales learn their repertoire as juveniles, and preferentially acquire calls from animals with which they socially interact frequently. Wild calves during their first year spend most of the time in close proximity to their mothers, and later gradually begin to interact more with whales from their own and related matrilines and are therefore more likely to learn from them.

It is not clear why vocal learning evolved in cetaceans but it appears possible that this adaptation happened before the divergence of baleen and toothed whales, as both suborders possess the ability for vocal learning (e.g., humpback whales Payne & Payne, 1985; Noad et al., 2000; Garland et al., 2011; belugas Delphinapterus leucas Ridgway et al., 2012; bottlenose dolphins Tursiops truncatus Reiss & McCowan, 1993; Smolker & Pepper, 1999; Janik, 2000; Miksis et al., 2002; Sayigh et al., 2007; King et al., 2013). An aquatic lifestyle may have favoured the development of vocal learning because sound producing organs compress while diving, which inevitably leads to changes in sound structure. To produce the same sounds reliably at different depths aquatic mammals therefore need to have voluntary control over sound production (Tyack & Sayigh, 1997).

In learned communication systems (such as bird song and possibly human language) the range of signals that can be learned is limited by genetically inherited predispositions that can be more or less strict depending on the selective advantages of some cultural traits over others (Lachlan & Feldman, 2003). This likely ensures that individuals learn primarily species-specific sounds and ignore most other sounds they are exposed to in their environment. Though the fine structure of stereotyped calls in killer whales is
learned, the extent to which the general structure of their total acoustic repertoire is learned or inherent remains unknown. In order to better understand this phenomenon one has to draw on examples of other cultural signal systems in which learning occurs to a larger or lesser degree in association with the socioecology of the culture of that species.

For example, many songbirds learn their songs from fathers or neighbouring males, but despite some variation their song remains species-specific and retains a certain structure (Slater, 1989). Bird song learning is a complex behaviour dependent on an inherited system that steers and filters the learning process. Though the capacity for vocal learning is definitely inherited (and present in all songbird species), to learn most songbird species need a template and period of training, when the ability to hear its own voice is crucial (Brainard & Doupe, 2002). The learned repertoire is limited by the inherent propensity towards the species’ own song, though some birds include many ‘alien’ vocalisations in their sounds (for example, marsh warbler *Acrocephalus palustris*, Dowsett-Lemaire, 1979; and lyrebird *Menura* spp., Robinson & Curtis, 1996). The same is probably true for human languages: although languages are learned and therefore extremely diverse, there is some inherited structure common to all human languages (Pinker, 1994). Others argue, however, that these patterns are better explained as stable engineering solutions satisfying multiple design constraints (such as cultural-historical factors and the constraints of human cognition) rather than genetically inherited rules in organization (Evans & Levinson, 2009).

Important evidence of an inherent nature of general song structure comes from an experiment with zebra finches (*Taeniopygia guttata*). Young males of this species learn song from adults, so the birds raised in isolation usually sing crude, rather unstructured songs. Fehér et al. (2009) used these males as templates for young naïve males. Even in the first generation, the song of young males was more structured and more similar to normal song than that of their tutors. These first-generation tutees were then used as a template for the new groups of young males, and so on. The song constantly improved in three-four generations, gradually converging on the normal song structure.

Similar processes have been described in humans. Children raised in isolation cannot invent a language themselves (Blumenthal, 2003). Deaf children in normally-hearing families often develop ‘home’ sign systems for basic communication, but they are far from the rich vocabulary and grammar of normal sign languages (Goldin-Meadow & Mylander, 1998). But when many
deaf children gather in the same place, they create their own new sign language with rich vocabulary and complex grammar (Sandler et al., 2005; Senghas et al., 2005). A similar example is the creolisation of pidgin. Pidgin is a simplified language that develops for communication between people that do not have a language in common, and usually represents a mixture of their languages without clear rules and proper grammar. In some cases, pidgin is adopted by children as their primary language and develops into creole — a natural language with complex linguistic structure (Mühlhäusler, 1997). This supports the inherited computational ability to form proper grammar but not the use or selection of signals (lexical capacity), which is likely always learned.

Hence, it appears that these signal systems, though socially learned in terms of detail of acoustic signals, possess some inherent species template. No controlled experiment of raising killer whale calves in isolation has been undertaken. We, therefore, cannot directly determine what parts of killer whale repertoire are inherent and what parts are learned. Nonetheless, from what is known about vocal learning in many song-bird species and humans it appears that repertoire is always a learned feature, because it does not evolve in absence of an acoustic template (Brainard & Doupe, 2002; Blumenthal, 2003). Comparing known repertoires worldwide to see which traits are common across populations could help to explore this question.

3. Functions of killer whale sounds

Killer whales produce a variety of acoustic signals that can be divided into three categories of sounds: clicks, calls and whistles (Ford, 1989). Like other odontocetes, killer whales produce clicks for echolocation, which allow them to navigate and to localize objects underwater based on the acoustic properties of echoes, such as timing and frequency (Barrett-Lennard et al., 1996; Au et al., 2004). Echolocation clicks are therefore not primarily communicative signals, although individuals might still extract valuable information (such as information on foraging) from the patterns of echolocation clicks of others.

The functions of calls and whistles are less clear, though it is believed that their primary function is to maintain group cohesion and to negotiate social relationships (Ford, 1989). Most calls are stereotyped, but their structure is very diverse ranging from simple low-frequency squeaks to complex multisyllabic high-pitched calls. The sets of stereotyped calls form distinct repertoires that allow the members of different pods to recognize and track each
other on distance (Ford, 1991; Miller & Bain, 2000; Filatova et al., 2011). Killer whales often gather in large aggregations where they mix and spread over several kilometers, so they need to track the movements of their matriline members to avoid being lost. Matriline calls must be recognizable in the cacophony of sounds produced simultaneously by many groups, including those with recent common ancestry. The extreme call diversity necessary for the accurate recognition is achieved through vocal learning, similarly to other learned recognition calls found in dolphins (signature whistles, Janik & Sayigh, 2013) and birds (Berg et al., 2011a, b) and language tags (accents) in humans that function to increase cooperation (Cohen, 2012).

While the group repertoires of stereotyped calls seem well suited for their proposed function of transmitting group identity, any specific function of individual stereotyped call types remains unclear. Ford (1989) did not find any strong connection between specific call types and one of four behavioural contexts (travelling, foraging, socializing, and resting), though some calls were more or less frequently used in particular contexts. Later studies found that social context was more important than behavioural context. In NE Pacific northern resident killer whales, Weiß et al. (2007) found significant differences in vocal behavior depending on the presence and identity of accompanying whales. Matriline-specific call subtypes, aberrant and variable calls were produced at higher rates, while ‘low arousal’ call types were used less in the presence of matrilines from different pods or clans. Foote et al. (2008) reported that in multipod aggregations of southern resident killer whales the proportional production of the dominant call types of each pod decreased, while the production of call types rarely recorded from single-pod groupings increased.

Despite killer whale calls are very diverse, they can be grouped into broad categories that appear to serve specific functions related to their structure. For example, in N Pacific resident killer whales, calls with two overlapping independently modulated components (two-voiced calls, Figure 1) are generally produced at higher sound pressure intensities than single-voiced calls (Miller, 2006). The upper-frequency component is more directional than the low-frequency component (Miller, 2002; Miller et al., 2007), which potentially allows receivers to identify the orientation and direction of movement of a caller by the proportional intensity of the upper to lower-frequency components (Miller, 2002; Lammers & Au, 2003). These findings suggest that one proximate function of two-voiced calls is long-range coordination of
movements. In association with their contextual occurrence, it appears that N Pacific residents use two-voiced calls to maintain contact and monitor the position of others when several pods are present and members of different pods are mixed (Ford, 1989; Foote et al., 2008; Filatova et al., 2009, 2013b). In contrast, single-voiced calls more often are heard in close-range contact between groups and in intra-pod social contexts. N Pacific transients use single-voiced calls exclusively when communicating in their social group after a successful marine mammal kill, but produce two-voiced calls when two distant groups communicate while in the process of meeting (Deecke et al., 2005).

Close-range contact signals exist in many species that need to maintain contact in limited visibility. For example, close-range cohesion calls were described in prosimians (Macedonia, 1986), Old World (Gautier & Gautier, 1977; Cheney et al., 1995) and New World (Snowdon, 1989) monkeys, and apes (Harcourt et al., 1993). Generally, most contact calls of primates have tonal structure and rather low frequency (Oda, 1996), similarly to the low-frequency monophonic calls of killer whales.

Contact calls are also common in passerine birds, which often have to maintain contact in the dense foliage. Bird calls function to signal about food, maintain contact between individuals, synchronize movements and resolve the aggressive and sexual conflicts (Marler, 2004). So, these calls can be used in a variety of behavioural contexts, which is also typical for killer whale calls.
No evidence has been found for the referential functions of killer whale calls similar to those found in some primates (Zuberbühler, 2003), but some calls appear to be used more often in specific behavioural contexts. For example, Ford (1989) reported that in resident killer whales variable and aberrant calls were used more often than stereotyped calls during socializing. Rehn et al. (2011) found that a particular category of variable call, characterized by a rapid, up-down frequency modulations, and termed ‘excitement’ call, was found in all three ecotypes of N Pacific killer whales and in all studied populations and clans within ecotypes. These calls appear to be used mostly in high arousal states (Ford, 1989; Rehn et al., 2011). Calls with this pattern of modulation were produced by killer whales during an exposure to naval sonar (our unpublished data); they might play a role in aggressive behavior, as this was one of the most pronounced changes in call structure during aggressive behavior in captive killer whales of Icelandic origin (Graham & Noonan, 2010). The only context-specific call reported to date is the so-called ‘herding call’ produced by killer whales in Iceland (Simon et al., 2006) and Shetland (Deecke et al., 2011), that appears to be used exclusively to manipulate herring into dense schools during feeding.

Sound usage varies across activity types among individuals and groups of the different populations and ecotypes, and seems to relate primarily to the social context and the kinds of prey being pursued. NE Pacific resident (fish-eating) killer whales vocalize frequently in any activity state except resting (Ford, 1989; Holt et al., 2013). In contrast, NE Pacific transient (mammal-eating) killer whales mostly call and whistle only after a kill or when socializing and are predominantly silent during all other activity states (Deecke et al., 2005; Saulitis et al., 2005; Riesch & Deecke, 2011). Individuals belonging to the AT1 transient population (a small population of whales inhabiting Prince William Sound and adjacent areas in Alaska) are known to produce high-amplitude pulsed calls while traveling alone. These calls travel great distances and presumably are used to contact other AT1 whales at a distance (Saulitis et al., 2005). Apparent contact calling by lone whales has also been observed in West Coast Transients inhabiting waters along North America’s West Coast (our unpublished data).

In the NE Atlantic, herring-eating killer whales off the coast of Norway and Iceland are highly vocal during feeding but mostly silent during traveling (Simon et al., 2007a; Samarra & Miller, 2015a). Similä & Ugarte (1993) and Simon et al. (2007a) hypothesized that the increased production
of calls during feeding may help whales coordinate their movements and/or to herd herring. This hypothesis was supported by observed lower calling rates during feeding upon discards from fisheries, a non-cooperative feeding behaviour, in comparison to higher calling rates during cooperative ‘carousel feeding’ (van Opzeeland et al., 2005). The increased production of calls, particularly two-voiced calls, during periods of coordinated movement further supports that hypothesis (Shapiro, 2008; Samarra & Miller, 2015).

Deecke et al. (2011) reported differences in vocal behavior of seal- and herring-eating killer whales off Shetland, UK similar to those between N Pacific residents and transients. Vocal rates of herring-eating whales were higher than those of seal eating whales, and seal-hunting groups vocalized almost exclusively when surface-active or milling after a kill. An alternative function for calling during feeding could be to attract other killer whales in the area to join in a feeding event, which could benefit killer whales already feeding on very large schools of herring.

Although the function of animal sounds has to ultimately be tested through playback studies, only one study of killer whale reaction to conspecific sounds has been published to date. It revealed that Kamchatkan (NW Pacific) resident killer whales respond vocally to calls belonging to their own repertoire, but not to the calls of other pods from their community having a different dialect (Filatova et al., 2011). Behavioral responses to playbacks of other dialects varied from interest to avoidance, perhaps reflecting variation in relationships between pods which would affect their desire to communicate. In contrast, during a single playback trial with NE Pacific resident killer whales off the coast of British Columbia, a strong acoustic response to the calls of another clan was documented (our unpublished data). Thus, the function(s) of calls and the function of the variety and complexity of their structure are far from understood. The low number of reported playback studies also underline that logistical challenges of studying these animals at sea make hypothesis testing through playback experiments difficult at best.

So, the studies of different populations have suggested different functions for call categories, such as two-voiced calls that appear to play a role in long-range group coordination in mixed pod contexts in N Pacific residents, but help coordinate group movements at short ranges during feeding in NE Atlantic herring-eating killer whales. The differential use and structural variation of calls described above suggests that different call categories (e.g. single-voiced vs. two-voiced), call types, or even sub-components of calls have different functions that may be subject to diverse selective pressures.
4. Geographic and dialect variation

All known killer whale populations produce stereotyped and variable calls, though sound structure of both categories of calls may vary across populations. For example, Foote & Nystuen (2008) found differences in frequency parameters between southern resident, transient and offshore N Pacific killer whale populations: the offshore ecotype produced calls with a significantly higher minimum frequency than the other ecotypes, and the resident ecotype produced calls with a significantly higher minimum and peak frequencies than the transient ecotype. Filatova et al. (2015) demonstrated that fundamental frequencies of N Pacific resident and N Atlantic killer whale calls were similar, while N Pacific transients had significantly lower frequency calls. It is not known whether these differences are culturally or genetically transmitted, because both gene flow and social interaction between ecotypes is almost absent (Ford, 2002; Parsons et al., 2013).

Repertoire variation of stereotyped calls also occurs both geographically and across different sympatric groupings. In resident killer whales, assemblages that share at least one call type are referred to as clans (Ford, 1991). The variation in call repertoire across different clans is often highly pronounced even within the same population and community (Ford, 1991; Yurk et al., 2002). Pods within clans are also easily recognizable (Ford, 1991; Yurk et al., 2002). Within pods, there are subtle variations in the structure of shared call types across matrilines (Miller & Bain, 2000; Deecke et al., 2010) and even more subtle variations of shared call types between animals from the same matriline (Nousek et al., 2006). Although dialects are well studied in N Pacific residents, there is no clear evidence of dialects in other populations, though there are reports of group-specific repertoire differences in N Pacific transients (Deecke, 2003), Norwegian (Strager, 1995) and Icelandic (Moore et al., 1988) killer whales. Thus, although the presence of group-specific call repertoires appears to be a general feature of killer whale social communication, the lack in knowledge of group-specific sound production in other populations currently hinders our understanding of whether the existence of group dialects is universal among killer whales (and therefore likely inherent in nature).

In resident killer whales, differences between population repertoires are usually of the same order or even less pronounced than between clans of the same population, suggesting differentiation between clans is more actively selected. In fact, there is no correlation between geographic distance
and similarity of time-frequency contours of calls in four N Pacific resident populations (Filatova et al., 2012a). Apart from studies in the N Pacific, there is little understanding of how calls compare between populations, ecotypes and ocean basins elsewhere, although some efforts have been made in a few populations (Awbrey, 1982; Moore et al., 1988; Strager, 1995; Stenersen & Similä, 2004; Shamir et al., 2014; Filatova et al., 2015). Some calls and their structural characteristics may be inherent species-specific traits, such as the calls with rapid, up-down frequency modulations emitted in high arousal (‘excitement’) situations, produced by N Pacific resident and transient killer whales (Ford, 1989; Rehn et al., 2011), as well as captive Icelandic whales (Graham & Noonan, 2010).

Another common feature across all known populations is the complex structure of some calls, consisting of different syllables that can occur sequentially or simultaneously (the latter yielding two-voiced calls, also reported in all known populations) (Moore et al., 1988; Ford, 1991; Strager, 1995; Yurk et al., 2002; Filatova et al., 2007; Shapiro et al., 2011). In some populations the same syllables of multisyllabic calls may occur in different combinations (NE Pacific residents: Yurk, 2005; NW Pacific residents: Filatova et al., 2013a; Norwegian: Strager, 1995; Shapiro et al., 2011; Icelandic: our unpublished data, see Figure 2) and may change in time independently of each other (Filatova et al., 2013a). In many populations some call types occur in stable sequences, following one another significantly more often than expected (NE Pacific residents: Ford, 1989; Norwegian: Strager, 1995; NE Pacific AT1 transients: Saulitis et al., 2005; NE Pacific West Coast transients: Deecke, 2003; Icelandic: our unpublished data), perhaps representing the early stage of formation of multi-syllabic calls.

Thus, while features of call production, such as production of both single-voiced and two-voiced calls, structural complexity leading to separate components (syllables), production of discrete combinations of multisyllabic calls and discrete sequences of call types seem to be widespread, specific call type repertoires differ distinctly between groups and populations. While some calls appear to be species-specific, such as the ‘excitement’ call (Ford, 1989; Rehn et al., 2011), suggesting that certain killer whale vocalizations are at least partly the result of genetic inheritance, a mixture of inherent structural variety of components and learned composition of components may form the call type repertoires of all populations.
Figure 2. Syllabic structure of the I43 call type from Icelandic killer whales. The first syllable may vary in frequency from <2 kHz (l, low) to 2–4 kHz (m, medium) and >4 kHz (h, high). The h and l variants may occur simultaneously, yielding two-voiced sounds (I43 hlb in this figure). The second syllable appears in at least four forms (a–d) that occur in various combinations with variations of the first syllable. Spectrogram parameters: FFT size 1024, overlap 87.5%.

5. Temporal stability and change of call repertoires

It is not known exactly how fast killer whales alter their repertoires. Though some minor changes in time-frequency parameters have been documented (Deecke et al., 2000; Wieland et al., 2010), the general structure of calls is relatively stable over tens of years. Ford (1991) documented stability of call repertoires by comparing recordings from the 1980s to historical recordings made in the 1950s. Foote et al. (2008) found that the repertoires and pro-
portional call usage of the Southern Resident population described by Ford (1991) were stable for a further two decades, so the two studies combined cover almost 50 years. Therefore it appears that dialects remain stable over the life span of at least one generation.

There are few studies examining temporal changes in killer whale calls, due to the limited availability of long time data series from the same matriline. Another approach is to examine patterns of similarity in contemporary repertoires. The null hypothesis of such a study predicts that the higher call similarity should occur in recently diverged groupings, assuming random steady error accumulation. Indeed, shared call types are more similar within than between matrilines (Nousek et al., 2006) and between matrilines of the same pod than between pods (Ford, 1991; Miller & Bain, 2000).

However, this pattern does not arise on a larger scale: Filatova et al. (2012a) found no direct correlation between call repertoire similarity and geographical distance between populations, rejecting the prediction that geographically close populations should be more similar. One of the possible explanations is that call evolution may be too rapid to be phylogenetically meaningful on a population scale: call repertoires may be already so diverse that they retain no signs of common ancestry, and all similarities could be the result of random convergence.

The main function of diversity in group repertoires is considered to be transmitting the group identity (Ford, 1991; Miller & Bain, 2000; Filatova et al., 2011). Therefore, the evolution of repertoires should balance between stability over time to provide recognition of groups, and need to change to reflect the variation in social bonds between these groups. Maintaining this balance may alone provide the patterns of observed repertoire sharing and divergence, but there are other forces that have been hypothesised to affect the cultural evolution of dialects.

For example, a potential force driving diversity in some killer whale populations could be sexual selection. Barrett-Lennard (2000) showed that resident killer whale females in Northern resident population tend to choose mates with the most dissimilar dialect (though mating within pods have been observed in the Southern resident population, Ford et al., 2011b). If females choose males on the basis of acoustic divergence of a male’s call with their own sounds, sexual selection towards faster call evolution would be expected. Males from the groups that tend to change calls quickly theoretically would have more mating opportunities and this would likely surpass
the drawbacks of possible inbreeding, as for males the parental investment and therefore the cost of producing unhealthy offspring is low.

In this case, it would be beneficial to adult males to have repertoires different from the rest of the group. As killer whales possess their dialects through vocal learning, they could invent the novel calls, as bottlenose dolphins do when developing their signature whistles (Janik & Sayigh, 2013). However, multiple observations suggest that all group members share the same repertoire (Ford, 1991; Miller et al., 2004). A selection force that preserves this vocal conformity may arise from the within-group benefit of using shared calls. Miller et al. (2004) have shown that matriline members often matched each other’s call types during vocal exchanges. Sharing call types may improve the ability of callers to localise each other, because possessing a template of a non-degraded call facilitates distance estimation to a caller using propagation effects on calls. Therefore, the members of a whole matriline would benefit from changing their calls more or less simultaneously. In this case, repertoire divergence driven by the matriline members prone to the faster call change (young whales or, in case of sexual selection, males) will be slowed down by more conservative animals, i.e. conformity may be fostered by maintaining group membership which is directly associated with individual survival (Foster et al., 2012).

Another stabilising force may originate from a benefit for killer whale groups to possess markers of population identity. Bird song studies showed that culturally transmitted vocal markers of population identity may restrict interbreeding between populations. For example, a correlation between song syllable variation and allele frequency variation was found in two populations of swamp sparrows (*Melospiza georgiana*; Balaban, 1988). Theoretical study suggests that little dispersal and assortative mating increase the stability of dialects (Planque et al., 2014). The benefit of such reproductive isolation may be the selection of a mate that is better adapted to local conditions. In white-crowned sparrows (*Zonotrichia leucophrys*), nonlocal males were reported to have lower paternity and higher parasite loads than males singing a local dialect (MacDougall-Shackleton et al., 2002).

In killer whales, reproductive isolation between sympatric resident and transient populations is likely to be beneficial because both ecotypes may have specific genetically inherited adaptations to their ecological niche (Ford et al., 1998; Saulitis et al., 2000; Ford, 2002; Moura et al., 2014). It is likely
that this isolation is based at least partly on acoustic cues, because whales from different ecotypes do not interact closely, and sounds — particularly long-range two-voiced calls — are the only feature of nearby populations that whales can typically perceive. Indeed, Foote & Nystuen (2008) and Filatova et al. (2015) have shown significant differences in call structure between ecotypes. Northern and Southern resident populations in the NE Pacific present an interesting case: they belong to the same ecotype and share a substantial part of their geographical range yet retain social and reproductive isolation, which could be an epiphenomenon of their vocal divergence.

Different communicative functions of single-voiced and two-voiced calls suggest that they may follow different evolutionary pathways. For example, ecological factors may produce more stabilizing selective pressure on two-voiced calls as they may function in group movement coordination and perhaps are the first calls to be heard at a distance by different groups, giving them a higher importance in group identification. Indeed, the diversity of two-voiced calls was significantly lower than the diversity of single-voiced calls in four N Pacific resident killer whale populations (Filatova et al., 2012a).

In addition, Filatova et al. (2012a) found that single-voiced calls were more diverse and two-voiced calls were less diverse in larger resident populations. Larger populations have more matrilines, which on average would produce more call type diversity overall due to random drift. In birds, the relationship between population size and diversity has also been reported: among island populations of chaffinches, the greater diversity of syllables was observed in the largest population (Lynch & Baker, 1993).

Contrastingly to single-voiced calls, two-voiced calls were less diverse in larger populations. Two-voiced calls are the most likely population markers (as described above), and the need for unique vocal population markers increases with population size, because it is presumably harder to remember all its members. It would shift the balance of diversifying and standardizing forces in two-voiced calls in favour of the standardization. The interaction of diversifying and standardizing evolutionary forces may create the ‘maximum diversity within the permitted range’, where the ‘range’ is represented by vocal population markers or within-group conformity.
6. Proximate mechanisms of dialect change

Although changes in calls and components of calls have been documented, these observations do not explain the mechanism(s) by which these changes occur. In killer whales, several proximate mechanisms of dialect change have been proposed: random learning errors and innovations (Ford, 1991) and horizontal transmission (Deecke et al., 2000; Filatova et al., 2010, 2013a).

Random evolution suggests that recently diverged and therefore more related matrilines will have more similar calls. This is difficult to test because mitochondrial DNA does not have enough sufficient resolution and microsatellites are contaminated with paternal relatedness. Deecke et al. (2010) investigated correlation between similarity of call contours, social associations and genetic relatedness in nine matrilines of NE Pacific resident killer whales. They found that call similarity was positively correlated with association rates for two of the three call types analysed, and positively correlated with relatedness for one call type. No relationship between relatedness and association frequency was detected. The authors interpreted these results to suggest that call structure plays a role in kin recognition and shapes the association behaviour of killer whale groups. However, the results may equally suggest that kinship is responsible for some similarities while social associations are responsible for the other, i.e. some changes in call structure accumulate with time, while others are copied between associating groups. The authors used the similarity of whole call contours, so no test of correlation between similarity of different call parameters or syllable subunits was conducted.

Filatova et al. (2013a) calculated the similarity of seven syllables from three call types across 14 matrilineal units of NW Pacific resident killer whales. Based on the assumption of random steady error accumulation, the authors expected that the similarity of most syllables across matrilineal groups would correlate with each other, i.e., if one syllable is more similar in groups A and B than in C, most other syllables will be also more similar in A and B than in C. Surprisingly, very few correlations were significant.

One explanation for this may be horizontal cultural transmission of call features across matrilines, which is supported by the fact that call similarity often correlate with association patterns of matrilines (Miller & Bain, 2000; Deecke et al., 2010). Horizontal transmission is widespread in bird song (Catchpole & Slater, 1995; Brainard & Doupe, 2002) and human languages (Wang & Minett, 2005), and is the main transmission mode in humpback
whale song (Payne & Payne, 1985; Noad et al., 2000; Garland et al., 2011). In human languages, a distinction is often drawn between the transmission and diffusion of linguistic change: transmission occurs when a child learns the language, and diffusion defines the minor changes spreading among adult speakers (Labov, 2011). It is possible that a similar distinction may exist in killer whales, because most examples of vocal learning come from juvenile whales (Bowles et al., 1988; Foote et al., 2006; Crance et al., 2014).

A lack of concordance between the similarity patterns of syllables may also be caused by potential constraints on call structure. Two of the three analysed calls were two-voiced, and the NW population have the lowest diversity of two-voiced calls of N Pacific resident populations (Filatova et al., 2012a). If some call components can structurally vary only within a restricted range, then variation in the calls of diverging matrilines could soon reach that limit and the differences between them can no longer increase. This situation is analogous to ‘mutational saturation’ in genetics when many changes at a given position randomize the genuine phylogenetic signal (Delsuc et al., 2005). Saturation leads to convergence of traits in non-related taxa (homoplasy), which also occurs in morphology (e.g., Alvarez et al., 1999; Mueller et al., 2004).

Moreover, different syllables may change over time at different speeds: call change in diverging matrilines may start in one syllable while others retain similarity (e.g., see terminal component in Miller & Bain, 2000 and Figure 3). Many studies of killer whale calls used few (or even one) parameters to compare call similarity, implicitly assuming that similarity patterns for all parameters are equal, therefore this phenomenon could have been missed.

Diverse rates of change have been found in various evolving systems ranging from DNA to bird songs. For example, some DNA regions are variable while others are more conservative. This happens because some functional regions rapidly differentiate under selection and some are constrained due to background selection, while non-functional regions are not subject to selection and therefore evolve purely due to drift (but the speed of drift also varies). Similar factors may drive the evolution of different call components in killer whale repertoires. More variable genomic regions are more susceptible to saturation and homoplasy, and differences across regions can be eroded by gene flow in the same way that diffusion and horizontal cultural transmission may erode cultural differences.
Figure 3. Variation in terminal components of calls of the NE Pacific Northern Resident population. Call types N4, N5, and N9 produced by matrilines of A1 pod have very short or almost absent terminal components, while the same three call types in matrilines of A5 pod possess well-developed terminal components. Spectrogram parameters: FFT size 1024, overlap 87.5%.
The occurrence of stable and variable features was also described in bird songs (Price & Lanyon, 2002). In European starlings (*Sturnus vulgaris*) whistle-like song types showed little variation, while other song types were extremely variable (Adret-Hausberger, 1989). It is likely that in this case different song types serve different functions: the shared song is used for species recognition, while variable song types are used to increase complexity supported by sexual selection. The same could be true for killer whales, with some syllables and/or call types used as population markers and therefore more conservative, and others functioning as matriline markers and — due to sexual selection — changing much faster. Different speed of change in different syllables was proposed by Yurk (2005) based on their structural similarity.

The mechanisms described above mostly deal with call change within types but do not explain divergence between types. Modelling suggests that accumulation of random errors cannot lead to divergence between types as it only produces a gradual continuum of call parameters (Filatova & Miller, 2015). If mechanisms other than accumulation of random errors are employed in the process of call change, exploring and describing these mechanisms would be an important step towards the understanding of dialect evolution in killer whales.

In human languages, Labov (2011) reports the following proximate factors of linguistic change: a tendency to maximise variability within a system; a tendency to generalise changes among the analogous elements of the system; and the principle of least effort — ‘we speak as rapidly and with as little effort as possible, approaching always the limit where our interlocutors ask us to repeat our utterance’ (Bloomfield, 1933). The former factor works to increase the complexity, while the latter two factors lead to simplification, which may contribute to the fact that cultural traits become increasingly simple and conventional when repeatedly learned (Kirby et al., 2008; Perfors & Navarro, 2014; Tamariz & Kirby, 2015).

A tendency to maximise variability within a system may appear in the form of deliberate innovations observed both in humans (Labov, 2011) and birds (Baker et al., 2003; Slater & Lachlan, 2003). Ford (1991) suggested that innovation must play a major role in the formation of new call types in killer whale dialects, while random errors can only alter the structure of existing types. Divergence of call types could also result from predisposition to diverge from calls of other matrilines (Filatova & Miller, 2015). Grebner et
al. (2011) supposed that the evolution of killer whale calls may occur through increasing random variation in calls and then crystallizing call subtypes from the graded continuum. This would imply a deliberate tendency to increase the usage of calls with specific structure and discard calls with parameters intermediate between subtypes.

A tendency to generalise changes among the analogous elements of the system is common in human languages (Labov, 2011) both in phonology (for example, vowel quality changes) and grammar (for example, transition of irregular verbs into regular in English). This tendency was not reported in birds, but there is some indication that it may occur in killer whales. For example, in the NE Pacific Northern Resident population the N4, N5 and N9 call types produced by matrilines of A5 pod possess well-developed terminal components, while the same three call types in matrilines of A1 pod have much shorter or almost absent terminal components (Figure 3). The most likely mechanism here is that after the divergence of A5 and A1 pods the terminal components in different call types within each pod changed in a similar way, analogous to vowel shifts in different words of human languages.

The principle of least effort appears in human languages as phonetic reduction of speech forms. It is also applicable to bird songs, as the development of large repertoires is time-consuming, since singing time cannot be invested in other activities such as foraging (Oberweger & Goller, 2001). In killer whale dialects, the principle of least effort may occur in the form of shortening or reduction of some syllables. For example, in Kamchatkan killer whales, the K5 call in many pods has a shortened version, which occurs as a separate K23 or K24 type (Figure 4).

In human speech this principle leads to a permanent shortening of words; it is compensated by fusion of independent morphemes into complex words (agglutination), thus forming a ‘linguistic cycle’ of evolution of isolating languages into agglutinative languages, which are then transformed into fusional languages, and then develop back into isolating ones (Van Gelderen, 2011). In killer whale dialects, fusion of different calls is also sometimes observed. For example, in Alaskan resident killer whales, the AKS05 call can occur separately or as a part of the high-frequency component of AKS09 call; similarly, in one pod of Kamchatkan killer whales, the K10 call can appear separately or as the first syllable in the K7 call and last syllable in the K12 call (Figure 5).
Figure 4. The shortened versions of Kamchatkan killer whales K5 call: the spectrograms on the left are three subtypes of K5 call type from three different matrilinear units, while the spectrograms on the right are from K24 and K23 calls from the same units that are produced by the shortening of the central syllable. Spectrogram parameters: FFT size 1024, overlap 87.5%.

In the dialects of other populations, this phenomenon can be developed even further: for example, Norwegian killer whales produce compound calls consisting from different combinations of syllables (Strager, 1995; Shapiro et al., 2011). A similar phenomenon was observed in some Icelandic killer whale calls (our unpubl. data, Figure 2). Combinations of N7 and N8 calls in the NE Pacific Northern Resident population (Ford, 1991) and T01 and T02 calls in the West Coast Transient population (Deecke, 2003) can be
also considered compound calls because the second call in the sequence is never produced alone, but always preceded by the first. The compound calls may further develop into stable stereotyped calls with multiple syllables, and then some syllables may reduce, forming calls with one or few syllables, that may again combine into compound calls in a cycle analogous to the ‘linguistic cycle’ of human languages. Similarly to languages, killer whale dialects may be at different stages of the process of reduction/fusion of syllables yielding high diversity of repertoire structures observed in killer whale populations.

7. Conclusions

In summary, we conclude that the cultural evolution of killer whale calls is not a random process driven by steady error accumulation alone (though random errors can play a role in vocal change). Temporal change occurs at different speeds in different components of killer whale repertoires: in different syllables and in different structural categories of calls. Constraints in call structure and horizontal transmission often degrade the phylogenetic signal. The important practical consequence of these findings is the fact that similarity of repertoires is not necessarily proportional to the time that passed since divergence of their bearers. Therefore, acoustic similarity does not always indicate relatedness (either at the population or the matriline level): correlation between relatedness and acoustic similarity is not linear or even proportional and may vary across call types and syllables.

Future studies should focus on revealing whether there are any general principles that define the speed and direction of change of different call categories and different syllables, i.e. determine whether or not calls or syllables with similar structure evolve with similar speed. Defining the rate of change of various call parameters will allow us to determine which parameters are phylogenetically meaningful at different scales. It is also important to ascertain if any general rules apply to the whole species (and therefore are probably innate), or if they are unique to specific populations/metapopulations. Both large-scale comparative analysis of repertoires of different metapopulations (e.g., N Pacific vs N Atlantic) and fine-scale analysis of similarity of different components of repertoires are promising directions for future research.
Figure 5. Examples of call fusion. (a) calls from Kamchatkan killer whales: K12 including K10 as the last syllable; K10 separately; K7 including K10 as the first syllable; (b) calls from Alaskan killer whales: AKS05 call separately; AKS05 as a high-frequency syllable of AKS09. Spectrogram parameters: FFT size 1024, overlap 87.5%.
Acknowledgements

Part of the data used in this study was collected with funding by the Russian Foundation for the Fundamental Research (Grant 15-34-20392), the Rufford Small Grants Fund, Whale and Dolphin Conservation, the Fundação para a Ciência e a Tecnologia, a Russell Trust Award of the University of St Andrews, the Office of Naval Research, the Icelandic Research Fund (Rannsóknasjóður), and the National Geographic Global Exploration Fund. OF was funded by a Marie Curie International Incoming Fellowship.

References


Cultural evolution of killer whale calls


of the killer whale indicates ecotype evolution in sympatry involving both selection and drift. — Mol. Ecol. 23: 5179-5192.


