Similarities in Composition and Transformations of Songs by Humpback Whales (*Megaptera novaeangliae*) Over Time and Space

Eduardo Mercado III and Christina E. Perazio

1 Department of Psychology, University at Buffalo, The State University of New York
2 Department of Psychology, University of New England

The complex songs produced by humpback whales have been cited as evidence of prodigious memory, innovativeness, sophisticated auditory scene analysis, vocal imitation, and even culture. Researchers believe humpbacks learn their songs culturally because songs appear to change rapidly, consistently, and irreversibly across whales within a population. Here, we present evidence of similarities in song structure both across populations and decades that strongly challenge claims that social learning is the main driver of variations in humpback whale songs over time. Groups of humpback whales that were not in acoustic contact (recorded in Puerto Rico in 1970, Hawaii in 2012, and Colombia in 2013–2019) produced songs in acoustically comparable cycles, suggesting that progression through sound patterns within and across songs is not simply determined by vocal imitation of innovative patterns, but may instead be controlled by production templates that prescribe how singers construct and transform songs over time. Identifying universal constraints on song production is critical to evaluating the role of vocal imitation and cultural transmission in the progressive changes that humpback whales make to their songs and for evaluating the functional relevance of such changes. The current findings illustrate how information theoretic analyses of vocal sequences can potentially obscure key acoustic qualities of signals that may be critical to understanding how vocalizers produce, perceive, and use those sequences.

**Keywords:** acoustic communication, cetacean, mysticete, cultural transmission, bioacoustics

**Supplemental materials:** https://doi.org/10.1037/com0000268.supp
cetaceans have been pointed to as evidence of prodigious memory (Guinee & Payne, 1988), innovativeness (Payne, 2000), complex cognition (Marino et al., 2007), sophisticated auditory scene analysis (Mercado, 2018b; Mercado, 2020), vocal imitation (Payne et al., 1983), and even culture (Rendell & Whitehead, 2001b).

Information theoretic analyses have been applied extensively to the vocal sequences of one cetacean species in particular: the humpback whale (Garland et al., 2013; Garland, Rendell, Lilley, et al., 2017; Kershenbaum & Garland, 2015; Miksis-Olds et al., 2008; Suzuki et al., 2006). Most analyses have focused on humpback whale “songs,” which correspond to predictable cycles (lasting ~7–30 min) within continuous rhythmic bouts of sound production (called song sessions) that may last many hours—how long humpback whales may persist in singing remains unknown, but other baleen whales have been observed to sing nonstop for 70+ hours (Clark et al., 2019), and at least one humpback whale was observed singing continuously for 20+ hours (Winn & Winn, 1978). Past analyses of humpback whale songs traditionally involve converting recordings into sequences of discrete symbols corresponding to either individual sounds (“units”), sequences of units (“phrases”), or segments of repeated phrases (“themes”). These symbolic sequences then can be used to quantify the entropy of songs (Suzuki et al., 2006), the amount of information transmitted by songs (Miksis-Olds et al., 2008), or to calculate similarities between sound patterns within songs (Eriksen et al., 2005; Murray et al., 2018). Results from such analyses have been used to argue that humpback whale songs are hierarchically structured (Cholewiak et al., 2013) and that the songs humpback whales produce in any given year are determined by processes of cultural transmission, such as vocal imitation of recently heard songs (Noad et al., 2000).

Vocal imitation abilities are relatively rare among mammals (for review, see Mercado et al., 2014; Wirthlin et al., 2019). Song learning in mammals is even more rare. In fact, humans, humpbacks, and bowhead whales (Stafford et al., 2018) are the only mammalian species known to progressively change the songs they sing throughout their life span. If humpback whales are continuously changing their songs to imitate other whales they have heard singing, then this raises a host of questions about how they keep track of the details of songs they have heard, how they decide which features to copy, why they abandon earlier versions of songs that they have mastered, and why most other cetaceans did not evolve comparable vocal systems (Parsons et al., 2008). One reason that researchers believe that singing humpback whales are learning their songs through cultural transmission is because the complex structural patterns within songs change so rapidly and consistently across whales within a population that it is difficult to imagine what other mechanisms might explain such variations (Payne et al., 1983; Rendell & Whitehead, 2001b). It has also been claimed that the yearly changes in sound sequences that whales make are irreversible (Garland et al., 2011; Payne & Payne, 1985), unpredictable, and divergent across populations (Winn et al., 1981). However, because information theoretic analyses of vocal sequences rely on assigning units or phrases to different subjectively defined categories of “unit types” or “phrase types,” the validity of these analyses depends heavily on how well these auditory categories capture the vocal variations actually produced by singers. Here, we present data suggesting that replacing units or phrases produced by singing whales with symbol sequences ob-scures variations in song cycles that are relevant to evaluating the likelihood that humpback whales socially learn songs from continuously monitoring and imitating the songs of other whales.

Interest in possible vocal dialects across populations of humpback whales has been strong from the earliest analyses of recordings of their songs (Payne & Guinee, 1983; Winn & Winn, 1978). The discovery that singers were progressively changing their songs each year did little to quell this enthusiasm (Payne & Payne, 1985; Payne et al., 1983). Singers in different populations are clearly producing qualitatively different vocal sequences in any given year (Winn et al., 1981), consistent with the possibility of local dialects resulting from cultural transmission. On the other hand, groups of singers separated by distances of 5,000+ km have been observed producing similarly structured songs (Cerchio et al., 2001; Darling et al., 2014, 2019; Darling & Sousa-Lima, 2005; Helweg et al., 1990), stretching the meaning of the term dialect potentially past its limit. Additionally, populations that show little genetic overlap commonly show evidence of shared song forms (Garland et al., 2011; Rekdahl et al., 2018), further calling into question the suggestion that song differences serve as acoustic badges of group membership. Finally, if song features are truly transmitted through acoustic contact followed by imitation, then one would expect that (a) populations not in acoustic contact would produce increasingly divergent song forms over time, and (b) the likelihood that song forms separated by multiple decades would be highly similar should be low. These predictions would hold even if the only source of song variations was drift from copying errors. The fact that singers sometimes introduce changes to songs that are clearly not copying errors, a phenomenon which has been interpreted as evidence that some singers are introducing innovations (Garland, Rendell, Lamoni, et al., 2017; Payne, 2000), should amplify the divergence of song forms over time. Consequently, evidence of acoustic similarities in song structure either across populations or decades would strongly challenge claims that cultural transmission is the primary driver of variations in humpback whale song structure over time. The first aim of this study was to assess whether there are such similarities in humpback whale song structure that information theoretic analyses have failed to reveal.

A second aim of this study was to evaluate the relationship between subjectively salient features of the vocal sequences produced by singing humpback whales and more objective measures of their acoustic properties, particularly with respect to variations in the frequency content of songs. Recent analyses suggest that singers in different populations focus most of their vocal effort within predictable frequency bands, regardless of the environmental conditions within which they are singing (Mercado, 2018a; Perazio & Mercado, 2018; Ryan et al., 2019; Seger et al., 2016). Qualitative comparisons of published spectrograms illustrating segments of song sessions further suggest that not only are singers emphasizing core frequency bands, they also may be consistently cycling through these bands in a predictable sequence (Español-Jiménez & van der Schaar, 2018; Helble et al., 2015; Mercado, 2016; Mercado & Handel, 2012). If singing humpback whales are predominately producing songs that exhibit predictable spectral cycles, then this would suggest that how they progress through sound patterns while singing is not simply a matter of which themes they have heard and learned in the past, but may instead be determined by production templates that prescribe the ways in which humpback whales transform their vocal sequences over
time. Identifying any universal properties of song composition is critical to evaluating the role that vocal imitation may play in the progressive changes that singing humpback whales continuously make.

A final aim of the current study was to examine more generally the kinds of interpretational issues that can arise when the first step in an information-theoretic analysis of vocal sequences involves subjectively discretizing, sorting, and symbolically representing sounds within sequences. Traditionally, the objectivity of such methods has been defended based on interobserver agreement with respect to how experts subjectively classify individual sounds and the fact that alternative biologically relevant partitions of vocal signals are not available. There are clear historical cases, however, where this approach has failed (e.g., early studies of bat vocal behavior before the development of ultrasonic recorders; see Griffin, 1958). Here, we show how preconceptions about both the sounds in an animal’s vocal repertoire and the ways in which sound sequences are structured can obscure potentially relevant qualities of vocal signals. Consequently, what information theoretic analyses of vocal sequences gain in rigor they may lose in scientific relevance, depending on the nature of the sequences being analyzed.

Method

Song Sample

Two different approaches were used to identify archival songs recorded from humpback whales that could provide an objective assessment of variations in singing behavior within and across populations over time. The first focused on published spectrograms of entire songs that illustrated their acoustic structure. Spectrographic representations of humpback whale songs can vary considerably depending on the analysis settings used, the quality of recordings, the size of the images that are published, and so on. Nevertheless, such variations are unlikely to increase the measurable similarity of depicted songs unless the recordings are noisy or if the images/analyses are low resolution. We elected to focus on the first published spectrogram of a full song as well as one of the most recently published spectrograms of a full song for qualitative comparisons of song similarities. The former (designated W70) depicts a song recorded on April 1–2, 1970, 22 miles west of Mayaguez Harbor, Puerto Rico (for details of recording equipment see Winn et al., 1970). Here, we show how preconceptions about both the sounds in an animal’s vocal repertoire and the ways in which sound sequences are structured can obscure potentially relevant qualities of vocal signals. Consequently, what information theoretic analyses of vocal sequences gain in rigor they may lose in scientific relevance, depending on the nature of the sequences being analyzed.

Table 1

Humpback Whale Song Recordings From Colombia

<table>
<thead>
<tr>
<th>Recording date</th>
<th>Recording length (min:s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td></td>
</tr>
<tr>
<td>8/2</td>
<td>34:41</td>
</tr>
<tr>
<td>8/12</td>
<td>34:04</td>
</tr>
<tr>
<td>9/9</td>
<td>12:10</td>
</tr>
<tr>
<td>9/13</td>
<td>41:53</td>
</tr>
<tr>
<td>2014</td>
<td></td>
</tr>
<tr>
<td>7/28</td>
<td>08:17</td>
</tr>
<tr>
<td>8/8</td>
<td>19:05</td>
</tr>
<tr>
<td>8/21</td>
<td>35:00</td>
</tr>
<tr>
<td>9/9</td>
<td>19:00</td>
</tr>
<tr>
<td>2015</td>
<td></td>
</tr>
<tr>
<td>8/27</td>
<td>24:16</td>
</tr>
<tr>
<td>8/6</td>
<td>14:01</td>
</tr>
<tr>
<td>2016</td>
<td></td>
</tr>
<tr>
<td>7/4</td>
<td>19:15</td>
</tr>
<tr>
<td>7/18</td>
<td>23:47</td>
</tr>
<tr>
<td>7/23</td>
<td>26:26</td>
</tr>
<tr>
<td>8/5</td>
<td>21:00</td>
</tr>
<tr>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>7/13</td>
<td>09:30</td>
</tr>
<tr>
<td>7/23</td>
<td>28:33</td>
</tr>
<tr>
<td>8/15</td>
<td>21:10</td>
</tr>
<tr>
<td>8/28</td>
<td>19:00</td>
</tr>
<tr>
<td>9/3</td>
<td>36:49</td>
</tr>
<tr>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>8/22</td>
<td>38:51</td>
</tr>
<tr>
<td>8/27</td>
<td>6:49</td>
</tr>
<tr>
<td>8/31</td>
<td>10:36</td>
</tr>
<tr>
<td>9/2</td>
<td>60:54</td>
</tr>
<tr>
<td>9/3</td>
<td>43:46</td>
</tr>
<tr>
<td>9/6</td>
<td>14:41</td>
</tr>
<tr>
<td>9/10</td>
<td>20:08</td>
</tr>
<tr>
<td>9/16</td>
<td>20:21</td>
</tr>
<tr>
<td>9/19</td>
<td>22:36</td>
</tr>
<tr>
<td>Total recordings analyzed</td>
<td>686:39</td>
</tr>
</tbody>
</table>
Phrase Analysis

Our qualitative analyses focused on three elements of humpback whale songs: the continuity of phrase sequencing within and across the two distant song samples (W70 and D19), variations in the composition of phrases within these two songs, and progressive changes within and across themes from each song, hereafter referred to as morphing trajectories.

Phrase Continuity

Payne and McVay (1971) provided the first detailed characterization of phrases repeated within humpback whale songs. They presented series of traced spectrograms for each identified theme to illustrate how singers progressively modified phrases as they repeated them within themes, as well as to show how consistently singers reproduced phrases across consecutive song cycles. Their approach enables one to subjectively evaluate visual similarities across spectrograms of phrases (corresponding to acoustic similarities) in a way that makes it immediately obvious how similar consecutive phrases are (Figure 1A). A limitation of this approach is that tracings of spectrograms necessarily discard some acoustic details that were present in the original vocal sequence and increase opportunities for observers’ subjective impressions of raw spectrograms as well as past experiences to affect which features are traced (including those that differentiate phrases across themes).

We adopted a variant of the approach pioneered by Payne and McVay (1971) in which spectrographic images are displayed in a time-series arrangement, so that no sounds or phrases are omitted, allowing for consecutive phrases to be visually compared. However, to avoid the introduction of subjective biases, raw spectrograms were used rather than tracings of spectrograms, and phrases were not divided into themes. An entire cycle of phrases (the song) was depicted in a circular arrangement to highlight the continuous nature of song production within singing bouts and to make all temporally adjacent phrases comparable (see Figure 1B). We also limited our segregation of sound patterns into phrases, to avoid inadvertently designating particular units as the first or last unit within a phrase, because this distinction is often subjective and inconsistent across studies (Cholewiak et al., 2013). Instead, we chose one salient pattern of units within each sequence (containing a “surface ratchet” followed by “cries”; see Figure 2A) with a subjectively clear difference from preceding patterns, designated that as a phrase, and then segmented following patterns such that the overall duration, timing, and sequential patterning of units aligned with those in the preceding segment as closely as possible. In the following, we refer to these circular arrangements of spectrograms of successive sound patterns within humpback whale songs as “singerings.” Figure S1 in the online supplemental materials provides an example singering of a full song recorded off the coast of Colombia, along with a traditional spectrographic representation of the recording (Movie S1 in the online supplemental materials) and annotations describing how the singer maintains some song features while gradually changing others.

Heterarchical Decomposition

Traditional approaches to analyzing vocal sequences produced by singing humpback whales depend heavily on the hierarchical scheme of classifying patterns proposed by Payne and McVay (1971). This approach designates sounds separated by intervals of silence as units, series of units as phrases, series of similar phrases as themes, and stereotyped sequences of themes as songs (for review, see Cholewiak et al., 2013). Payne and McVay’s terminology characterizes humpback whale songs as a nested hierarchy of clustered sound patterns, similar to how a phonologist might partition a spoken sentence with words being analogous to units. A more recent variant of this approach proposes that, for cross-species comparisons, phrases should be viewed as analogous to bird songs, with themes being analogous to repeated bird songs (Cholewiak et al., 2013). The rationale for this proposed switch in terminology came from the discovery that singers sometimes produce themes in variable orders, making it difficult to identify songs within such recordings. In both descriptive approaches, phrases play a central role because they determine what qualifies as a song. Accordingly, most published descriptions of songs include spectrograms depicting prototypical phrases (phrase types) that researchers used to identify theme types within a particular year and geographic location, as well as to compare songs within and across populations (Winn et al., 1981).

A major limitation of both approaches is that, unlike with bird songs or sentences, the criteria for partitioning patterns within humpback whale vocal sequences are arbitrary. There are not predictably longer duration pauses between phrases, themes, or songs that might be used to objectively designate a particular unit as the first or last sound within a phrase (see Figure S1 and Movie in the online supplemental materials). Pragmatically, what qualifies as a phrase becomes a matter of convention, because repeated patterns of units can sometimes include shorter patterns (referred to as “subphrases”) that are embedded within the repeated patterns. These complications have led researchers to conclude that phrases within humpback whale songs cannot be reliably identified without subjectively analyzing recordings from multiple singers in a given year and locale (Cholewiak et al., 2013). A second limitation of these approaches is that, because they depend on subjective thresholds for differentiating sound patterns, what qualifies as a phrase type is also a matter of convention. Consequently, two observers independently identifying phrase types within the same set of recordings are unlikely to identify the same number of types (Mercado et al., 2003).

One way to avoid these kinds of ambiguities in analyses of vocal sequences is to segment sequences using acoustic properties of vocalizations that relate to how sounds were produced rather than using subjective features that human observers find aurally or visually salient. Production-based representations of sounds and sound sequences have the advantage of directly corresponding to variations in the physical movements that a vocalizer makes when generating a signal, increasing the likelihood that they retain biologically relevant elements of communication signals (Adam et al., 2013; Cazau et al., 2013; Mercado, 2013; Mercado & Kuh, 1998; Mercado et al., 2010). The specific mechanisms that singing humpback whales use to produce songs remain unclear, although anatomical and acoustical analyses suggest that they do so by passing air over or through large, vibrating membranes (Cazau et al., 2013, 2016; Damien et al., 2019; Reidenberg & Laitman, 2007). Singing humpback whales do not release air bubbles underwater while singing, which implies that they are internally recirculating air while producing sounds. Anatomical observations suggest that during sound production air expired from the lungs is
Panel A: Idealized spectrograms of successive song phrases sampled from a full song cycle can be arranged vertically to reveal similarities in phrase composition (see Payne & McVay, 1971). Such representations can also reveal gradual changes in structure across consecutive phrases, such as division of a single unit into multiple units, compression or expansion of unit duration, decreases or increases in the number of units, and shifts in the frequency content or frequency modulation of units (see also Figure S1 in the online supplemental materials). Because such transformations minimally affect unit properties across consecutive phrases, it is likely that a singer producing acoustically similar units at similar time points across consecutive phrases is doing so using the same mechanisms. For example, a singer might produce a pulsive unit (top row, left; depicted as a series of vertical bands) via ingressive airflow (see also Figure 2A). If that unit is split into four separate units in the following phrase (second row, left), then the resulting units are likely also to be ingressive vocalizations. Tracking acoustic changes across consecutive phrases can thus potentially provide insights into how particular units were produced. Panel B: Circular arrangements of consecutive phrases or song segments (singerings) highlight the acoustic continuity of phrases within songs (see also Figure S1 in the online supplemental materials).
shunted into a laryngeal sac, after which it is shunted back into the lungs (Reidenberg, 2018); this cycle is thought to repeat multiple times while a singer is submerged. In principle, singing humpback whales may produce units as air travels in either or both directions. Alternating patterns of units within songs are consistent with bidirectional sound production (Mercado & Handel, 2012). If singers are producing sounds bidirectionally as air recirculates internally, then this method of sound production could contribute to the acoustic structure of phrases. Mercado and Handel (2012) proposed a heterarchical framework for describing song structure that incorporates this aspect of song production as well as changes in structure associated with the dive cycles of singers. In this framework, how phrases change over time within songs is a function of how long a singer has been submerged (the dive cycle, which is on the order of 10–20 min) as well as the direction of airflow during sound production (the recirculatory cycle, which hypothetically is on the order of 10–30 s). These mechanisms are heterarchical in that they interact to shape the acoustic qualities of successive phrases over time rather than being organized into discrete stages of nested levels that progress in a fixed order (Bruni & Giorgi, 2015; Crumley, 2015; Cumming, 2016; McCulloch, 1945). According to this heterarchical framework, humpback whale song phrases can potentially be decomposed into sounds produced as air is leaving the lungs versus sounds that are produced as air leaves the laryngeal sac.

In most animals known to produce sounds bidirectionally, acoustic features of sounds produced by airflow in one direction differ systematically from those of sounds produced by airflow in the opposite direction (because of the biophysics of sound production; Eklund, 2008). Simulations of sound production by humpback whales suggest that this is also likely to be the case within their songs (Cazau et al., 2013). To evaluate whether differences in sound patterns within songs across years and populations might be more apparent for sounds produced in one direction than the other, we compared patterns in the production of “ingressive” sounds to patterns in the production of “egressive” sounds. Classification of sounds as either ingressive or egressive was based on observations made by Winn and colleagues (1970) that singers consistently produced certain sounds just before surfacing to blow (the “ratchet” sound, presumed to be ingressive, as blows involve expiring air from the lungs) and other sounds immediately after a blow (presumed to be egressive; see Figure 2A). For each successive phrase in a song, sounds with similar acoustic features produced at corresponding temporal positions to preceding phrases

Figure 2
Predictably Timed Respiration Provides Indications of Bidirectional Sound Production by Singing Humpback Whales and by Humans

Note. Panel A: Humpback whales typically do not stop singing while surfacing, but instead breathe during silent intervals between units. Winn and colleagues (1970) observed 18 instances (five whales) in which blows produced by a singing humpback whale occurred following a “surface ratchet,” a multisecond, low-rate pulse train depicted here as a series of vertical bands. All blows were followed by production of one or two high-pitched tonal “cries” (shown here as traced horizontal bands). Assuming that blows during surfacing transfer air primarily out of and then into a whale’s lungs (as the whale inhales fresh air), cries are most likely egressive sounds (“E”) produced as air leaves the singer’s lungs. Assuming that singing humpback whales produce units bidirectionally, and that whales recirculate air back into their lungs before blowing, surface ratchets are more likely to be ingressive sounds (“I”) produced as air leaves the laryngeal sac. These designations of the direction of sound production during singing are tentative—in principle all units might be produced egressively (or ingressively). Panel B: Bidirectional sound production by a snoring human can lead to alternating sequences of pulsed and tonal sounds that are acoustically similar to the unit sequences produced by surfacing humpback whales. These cross-species similarities further suggest that singing humpback whales may produce units bidirectionally. See the online article for the color version of this figure.
were assumed to be produced using the same mechanisms (Figure 1A). These directional designations are provisional, given that internal movements of air within singing whales have yet to be directly observed. However, the acoustic qualities of sounds within each of the two categories match those seen in other mammals that produce sounds bidirectionally (e.g., donkeys and snoring humans; see Figure 2B), providing some face validity for the designations. If future studies reveal that these distinctions are unrelated to production mechanisms, then this approach reduces to segmenting repeated patterns into units produced earlier or later within each phrase based on acoustic/respiratory criteria for how to partition unit sequences.

Morphing Trajectories

Researchers noted early on that humpback whales were progressively modifying phrases within themes and across years by adding, deleting, and modifying units (Payne & McVay, 1971; Winn & Winn, 1978). Some efforts have been made to document how particular phrases change over time (Payne et al., 1983), but generally such efforts have been restricted to songs recorded in consecutive years from a single locale. Less attention has been given to how singers are morphing phrases within individual songs, other than to note that they often do so. Payne and Payne (1985) described “shifting themes,” themes in which units gradually shift in frequency, form, duration, or number as a phrase is repeated within a theme, as occurring in all of the songs they analyzed spanning 17 years. The extent and rate of unit morphing within phrases varies within and across years (Payne et al., 1983). Whether the trajectories of unit morphing are comparable across populations has never previously been assessed. We evaluated unit morphing in the two distant song samples to assess whether singers were changing songs in similar ways. If copying errors and/or innovations drive progressive changes in songs, then how singers morph units within themes should be dissimilar across the two samples. Alternatively, if whales in different populations morph units along similar trajectories, then this would suggest that the mechanisms driving progressive modifications in phrases are unlikely to be cultural.

Spectral Continuity

Recent acoustic analyses of humpback whale songs produced in multiple regions around the world have revealed that although singers are capable of producing units with frequency content ranging from 20 to 8,000 Hz, they concentrate most of their energy into three relatively narrow frequency bands (Perazio & Mercado, 2018; Ryan et al., 2019). The specific focal frequencies vary slightly over time within populations (Perazio & Mercado, 2018), and even across consecutive songs produced by individual whales (Mercado, 2018a), but all populations measured to date appear to conform to similar distributions of frequencies. Published spectrograms depicting changes in frequency usage within song cycles further suggest that singers may use these three bands in stereotypical orders while singing (Español-Jiménez & van der Schaar, 2018; Helble et al., 2015; Kowarski et al., 2018; Mercado, 2016; Mercado & Handel, 2012; Mercado et al., 2010; Mercado & Sturdy, 2017). If singers are consistently progressing through a fixed sequence of core frequency bands as they sing, then this would suggest that they are more constrained in how they can modify song structure than is currently assumed. To assess this possibility, we compared singers’ sequential usage of different frequency bands in a single locale over a period of 7 years. Time-series representations were created for individual recordings by extracting spectral peaks from consecutive units. Variations in spectral “song contours” were assessed quantitatively by fitting mathematical functions (polynomial and power functions) to the contours (using Matlab’s cftool function) and qualitatively by visually examining similarities in time-series representations across years. Shifts in frequency usage were also qualitatively compared with changes in phrase structure over time and in relation to possible heterarchical structure within songs.

Results

Phrase Continuity

The song recorded near Puerto Rico (W70) lasted 954 s, contained 325 units, and 37 phrases. The song recorded near Maui (D19) lasted 976 s, contained 334 units, and 171 phrases. Darling and colleagues (2019) identified 10 phrase types being used by singers across the North Pacific in 2012 within eight themes, and Winn classified Puerto Rican songs into six themes and phrase types. D19 contained five of the 2012 phrase types and three of the themes. Figure 3 compares representative phrases from 2012 (Darling et al., 2019, online supplemental materials) to the visually most similar spectrograms of phrases present within W70. These comparisons reveal numerous acoustic similarities between phrases identified by each group, particularly in the number and temporal distribution of units, but also in their spectral characteristics. Five out of the six themes identified by Winn and colleagues (1970) visually match phrase types identified by Darling and colleagues (2019), the only exception being a one-phrase theme (designated by Winn et al. as “Theme III”) that was composed of units from surrounding themes, which Darling and colleagues classify as a “transition” rather than as a phrase type. Furthermore, the singer recorded in Puerto Rico produced the matching phrase sequences (themes) in the same order as singers from the North Pacific. According to traditional standards for classifying two humpback whale songs as the same (i.e., matching phrase types produced in a stereotypical order; Cholewiak et al., 2013), humpback whales recorded off the coast of Puerto Rico in 1970 were singing the 2012 North Pacific song.

Singerings of W70 (Figure 4; see also Figure 1B) and D19 (see Figure 5) revealed additional acoustic similarities between these two songs, including (a) the relative timing and pacing of unit production was similar throughout the entire song cycle for both songs, visible as concentric circles within singerings; (b) changes in the number or duration of units within repeating patterns often preserved the overall timing of repeated patterns (i.e., unit durations, number, and internunit intervals appeared to be interdependent, particularly in W70); (c) units waxed and waned in duration across the song cycle (e.g., see shaded regions in Figure 5); and (d) progressive acoustic changes to sound patterns within songs were continuous and progressive across the entire song cycle (i.e., the acoustic features of each theme were related to the preceding theme, such that the entire song cycle could be classified as a single shifting theme). Figure 4, which includes all six themes identified by Winn and colleagues (1970), shows that each theme...
type is acoustically related to the one that precedes it, and that most transitions between themes involve gradual transformations (i.e., morphing) rather than discrete shifts between distinctive patterns. Note that the segmentation of songs shown in Figures 4 and 5, which was based on the acoustic similarity of consecutive unit sequences rather than on subjective classification of phrases, did not lead to the same groupings of units as shown in Figure 3. Thus, the patterns within songs revealed by singerings overlap with, but are not identical to, those evident from traditional thematic analyses.

**Heterarchical Decomposition**

Singerings were successfully partitioned into either “ingressive” or “egressive” components based on the temporal alignment of consecutive patterns and the acoustic similarity of units, starting with the “surface ratchet” phrase (designated by Winn & colleagues, 1970, as Theme VI; see Figure 3). In W70 (Figures 1 and 4), ingressive units appear as clusters along the inner circle of the singering. D19 showed a different pattern in which individual ingressive units alternated with individual egressive units throughout most of the song cycle (see Figure 5). In both songs, the time singers spent producing ingressive units relative to egressive units was similar, with egressive units being more prevalent in parts of the cycle and ingressive units accounting for a greater proportion of time in other parts. Overall, the balance of time spent producing egressive versus ingressive units as well as the alternating patterns of units were consistent with bidirectional sound production during rhythmic shunting of air between two cavities. However, the estimated period of recirculation in W70 (~25 s) was much longer than that estimated for most phrases within D19, which ranged from 5 to 21 s.

**Figure 3**

*Comparisons of Humpback Whale Phrase Types Across Oceans and Decades*

![Spectrograms showing similarities in sound patterns across themes and decades.](http://creativecommons.org/licenses/by/4.0/)

*Note.* Each of the six panels shows spectrograms of a sound pattern identified as part of a theme recorded by Winn and colleagues (1970) in 1970 off the coast of Puerto Rico (top half of each panel, with Roman numerals indicating theme designations), or by Darling and colleagues (2019) in 2012 in the North Pacific (bottom of each panel). Similarities in the number, timing, and spectral features of constituent units and sequences of units are visually evident. Spectrograms showing North Pacific themes are from Darling et al. (2019), Figure S2 in the online supplemental materials (http://creativecommons.org/licenses/by/4.0/).
Dividing units into two separate streams based on whether they were (hypothetically) produced during either inspiration or expiration (see Figure 6) revealed that acoustic morphing of units and unit sequences across the song cycles differed for these two classes of sounds. Ingressive sounds consistently contained spectral energy spread across a broad range of frequencies and changed mainly in terms of their number and/or duration throughout the song cycles; their spectrographic features generally were indicative of low pulse-rate sounds. The pulse-rate of ingressive units appeared to systematically decrease as the cycles progressed, although this could not be precisely evaluated through spectrographic images alone. This property was explored more extensively in analyses of Colombian songs described in the following text.

Acoustic morphing of egressive unit patterns was more extensive, especially for W70 (see Figure 6). Despite differing in the number and variety of egressive units being produced, both W70 and D19 showed a gradual progression from longer duration, higher pitched units, to shorter duration, medium pitched units, to more broadband, spectrally complex units (including lower pulse-rate units), to longer duration, higher pitched units,
and so on. Unit qualities associated with egressive production (shown in Figure 6) overlapped with those of ingressive units, suggesting that it may not be possible to reliably identify the direction of sound production based on the acoustic features of isolated units. The complexity of phrase structure in portions of W70 made it difficult to assess the possible alternation of individual ingressive and egressive units. Consequently, it is likely that a subset of the units within these more complex patterns that were classified as egressive were actually ingressive. This complication is addressed more directly in the quantitative analyses of spectral cycles described in the following text.

**Morphing Trajectories**

Sequences of units within W70 and D19 were progressively transformed along multiple acoustic dimensions (see Figures 1 and 4–6), as previously described by Payne and colleagues (Payne et al., 1983). The most visually salient acoustic morphing in singerings involved fluctuations in the durations of individual units over time (appearing as changes in the darkness and width of concentric rings and arcs), changes that appeared to be counterbalanced by correlated morphing of subsequent silent intervals and/or the number of units being produced. Qualitatively, longer duration units in W70 appeared to be subdivided over time to create clusters of shorter duration variants of the original unit as the song cycle progressed (Figure 6), only to merge again later in the cycle (i.e., the number of units in a subphrase was inversely correlated with unit duration). In both D19 and W70, tonal units gradually decreased in pitch and spanned a broader range of frequencies as the song progressed before resetting, a trajectory that was particularly clear for egressive units. The rate of acoustic morphing was not constant throughout a song cycle, such that certain units and sequences appeared to be repeated more times than others in both recordings. Most morphing trajectories were gradual, making it difficult to discern the direction of changes within time frames of less than 3 min.

**Figure 5**

_Singing of a Humpback Whale Song Recorded Near Hawaii in 2012_

*Note.* A singering depicting 24 segments sampled from D19 (3 kHz y-axis; width of each rectangular segment is −20 s) reveals temporal regularity and progressive morphing similar to that evident in W70 (compare with Figure 4). As in W70, acoustic dimensions that change systematically throughout the song cycle include unit duration, number, pitch, and spectral variability. Some segments of D19 (n = 20) that showed less noticeable changes in the progressive sequence were excluded to enhance the detectability of acoustic morphing across phrases. Shaded regions indicated units that were tentatively identified as ingressive vocalizations. Spectrograms are sampled from Darling et al. (2019), Figure S1 in the online supplemental materials (http://creativecommons.org/licenses/by/4.0/). See the online article for the color version of this figure.
Figure 6
Comparison of Morphing Trajectories of “Ingressive” and “Egressive” Units From Puerto Rico and Hawaii

Note. Panel A: “Ingressive” sounds from W70 consisted exclusively of low-rate pulsed sounds varying in duration, number, pulse rate, and pitch (units shown were sampled from a subset of the shaded regions in Figure 4). The “surface ratchet” (the unit above the time scale) initially divided into multiple units, which then gradually decreased in duration (narrower dark bands indicate shorter duration units), while decreasing in number. Then, both the number and duration of consecutive ingressive units gradually increased before ultimately merging back together to form the long-duration surface ratchet. Clusters of ingressive units were the norm in W70. Panel B: Ingressive sounds from D19 (sampled from Figure 5) similarly were exclusively low-rate pulsed sounds. Like ingressive vocalizations from W70, these units waxed and waned in duration and shifted in pitch throughout the song cycle. However, unlike W70, ingressive units from D19 were typically produced individually rather than in clusters, and did not vary in number. Note that visual differences between spectrograms of units from W70 and D19 reflect different analysis settings used by the investigators, as well as differences in the acoustic qualities of the units. Panel C: “Egressive” sounds from W70 (sampled from Figure 4) changed gradually in duration, number, pitch, frequency modulation, and/or complexity. Changes in units and unit sequences were more extensive and directional for egressive units than for ingressive units from W70, with repetitions of higher pitched, tonal sounds morphing into lower pitched, broadband sounds with more complex structure. For instance, in W70 pairs of “cries” (the units above the time scale) that were produced in alternation with the surface ratchet, initially shifted downward in the time-frequency plane before spawning clusters of short duration “chirps.” These multiunit clusters were then gradually replaced with longer duration, alternating frequency sweeps, which later transformed into a triad of even longer duration tonal units. This triplet then morphed into more complex frequency-modulated sounds before reverting back to cries. Panel D: Egressive sounds alternating with ratchets in D19 initially consisted of repeating cries, as in W70. This cluster of cries rapidly decreased in number, however, until only a single cry remained. The remaining cry rapidly decreased in pitch while morphing into a pulse train, after which its duration began to expand. This broadband unit then became increasingly complex before transforming back into cries.
Figure S1 in the online supplemental materials provides a high-resolution singering of a complete song cycle recorded off the coast of Colombia, along with the original recording (Movie S1 in the online supplemental materials), which shows comparable morphing trajectories to those present in W70 and D19, including progressive shifts in unit duration, pitch, number, and frequency contours that occur throughout the song cycle.

Spectral Continuity

Figure 7A shows spectrographically how a humpback whale singing off the coast of Colombia gradually changed the spectral content of its units over time. Specifically, higher pitched units gradually decreased in frequency content in multiple narrow bands in parallel, with an initial rapid decrease in the lowest frequencies followed by a more gradual decrease for most of the remainder of the song, until late in the cycle when the highest frequencies began to slowly rise and spread across a broader range. Observed trends in spectral change were simplified by measuring the peak frequency of each individual unit and then fitting multiple mathematical functions to the time-series of measurements of peak frequency (Figure 7B).

Detailed spectral analyses of song contours recorded in 2013 revealed that singers in Colombia consistently shifted their use of three core frequency bands within song cycles, both within consecutive songs and across time periods (see Figure 8). Although singers were not identified, it is highly likely that different recordings also correspond to different individuals, given that individuals generally spend limited time within the region (Perazio et al., 2013).
Singers shifted units rapidly from higher frequency ranges before spending most of their time producing units with peak frequencies in the medium and lower frequency bands (a trajectory also evident in the singing of D19). In 2013, Colombian singers also appeared to distribute their effort across a broader range of frequencies as they progressed through the song cycle, replicating the patterns of frequency spreading observed in W70 and D19 (Figures 4–7). They also tended to produce longer duration units earlier in the cycle and shorter duration units later in the cycle, such that unit pulse rate/fundamental frequency was correlated with unit duration (see also Mercado & Kuh, 1998). Qualitatively and quantitatively, consecutive songs within recordings from 2013 showed more consistency than did songs from across recordings (Figure 8 and Table 2). Nevertheless, subtle variations in the use of different frequency bands were evident even between the most similar song cycles (see Figure 8).

Converting functional coefficients (see Table 2) to $z$ scores and then calculating the Euclidean distance between $z$ scores con-
firmed that between-year differences in song contours (mean distance $= 5.0 \pm 8.0$) were larger than within-recording differences (mean distance $= 3.5 \pm 1.4$), with contours produced in 2013 being the most consistent (mean distance $= 3.0 \pm 4.0$). The specific frequencies within each band that were most often used by singing humpback whales varied from one year to the next (Figure 9; Table 2), suggesting either that singers have some flexibility in terms of which frequencies they can emphasize within songs, or that different whales use slightly different frequency bands (or both). Spectral cycles within songs produced by humpback whales off the coast of Colombia over a span of seven years generally matched those evident within W70 and D19. Specifically, portions of songs with energy concentrated in higher frequency bands ($>400$ Hz) were often followed by segments with energy focused at intermediate frequencies (between 400 and 200 Hz), which, in turn, were followed by segments with peak frequencies mainly focused at lower ranges ($<200$ Hz) and/or across a broad band of frequencies. Qualitatively, sequences of units produced within each segment were comparable to those present in W70 and D19, as were morphing trajectories, in that consecutive phrases generally showed acoustic continuity.

Notably, the dominant theme within the Colombian song from 2013 (accounting for $>50\%$ of the song cycle duration) was acoustically similar to Winn and colleagues’ “Theme IV” (see Figure 3), and to the dominant theme of the 2012 song produced by humpback whales in the North Pacific. Detailed analyses of phrases from this theme recorded in Colombia suggest that alternating series of ingressive and egressive sounds may be intermittently followed by a series of repeated ingressive sounds or by a series of repeated egressive sounds. Additionally, ingressive units with pulse rates lower than $180$ Hz, which often occurred later in a song cycle, contained many higher frequency components, some of which were more energetic than the fundamental frequency. Consequently the peak frequencies of these low-pulse rate units sometimes matched those of higher pitched units (e.g., see Figure 7B between 11 and 13 min), despite having a much lower fundamental frequency, implying that frequency content alone was not a reliable indicator of whether a unit was either ingressive or egressive.

### Discussion

Singing humpback whales continuously and collectively vary the structural and acoustic features of their songs throughout their lives. In contrast to past reports, findings from the current study suggest that neither cultural transmission nor social learning contributes significantly to how humpback whales change their songs over time. Detailed acoustic similarities exist between humpback whale songs produced across acoustically isolated populations and long-time spans that should not exist if singers are innovating sound patterns within songs or copying the innovations of others. Objective measures of song variations revealed that singing humpback whales consistently transformed the temporal and spectral properties of phrases within songs across years and populations, a potentially universal feature of songs not noted in past subjective analyses. Information theoretic analyses of humpback whale songs have obscured how singers are changing song features over time by discounting acoustic details and may similarly lead to errant

| Table 2  |
| Mathematical Function Fits for Song Contours |
| Recording date | $p_1$ | $p_2$ | a | b | $R^2$ | a | b | $R^2$ |
| 2013 | 
| 8/2 Song 1 | 0.2496 | 456 | 852.1 | $-2.552$ | .82 | 1.079 | $-6.004$ | .82 |
| 8/2 Song 2 | 0.2373 | 436 | 740.4 | $-2.235$ | .63 | 0.974 | $-4.189$ | .82 |
| 8/9 Song 1 | 0.5492 | 433.6 | 834.7 | $-2.605$ | .85 | 1.040 | $-4.689$ | .77 |
| 8/9 Song 2 | $-0.1532$ | 461.5 | 856 | $-2.346$ | .8 | 1.067 | $-4.365$ | .74 |
| 9/13 Song 1 | $-0.4309$ | 454 | 962.6 | $-2.568$ | .6 | 1.199 | $-4.507$ | .64 |
| 9/13 Song 2 | $-0.2582$ | 454 | 967.9 | $-2.588$ | .67 | 1.213 | $-4.52$ | .69 |
| 9/13 Song 3 | $-0.3409$ | 462 | 962.1 | $-2.373$ | .71 | 1.237 | $-4.781$ | .66 |
| 2014 | 
| 9/9 Song 1 | $-0.4831$ | 495.1 | 889.2 | $-2.112$ | .79 | 1.135 | $-4.309$ | .79 |
| 9/9 Song 2 | $-0.8816$ | 505.5 | 926.1 | $-2.116$ | .73 | 1.204 | $-4.425$ | .85 |
| 2015 | 
| 8/6 Song 1 | 0.5815 | 634.7 | 496 | $-1.121$ | .56 | 0.528 | $-2.72$ | .68 |
| 2016 | 
| 7/23 Song 1 | $-0.5099$ | 781.9 | 936.4 | $-2.868$ | .62 | 1.076 | $-4.746$ | .67 |
| 7/23 Song 2 | $-7.094$ | 829.4 | 639.6 | $-1.448$ | .25 | 0.793 | $-3.642$ | .73 |
| 2018 | 
| 7/23 Song 1 | $-4.472$ | 568.5 | 860.8 | $-1.859$ | .46 | 1.186 | $-4.367$ | .61 |
| 7/23 Song 2 | $-3.483$ | 583.7 | 824.5 | $-1.504$ | .36 | 1.145 | $-4.049$ | .58 |
| 2019 | 
| 9/2 Song 1 | 1.15 | 576.6 | 909.1 | $-1.167$ | .62 | 0.845 | $-3.85$ | .66 |
| 9/2 Song 2 | 0.7284 | 610.4 | 522.8 | $-1.802$ | .4 | 0.636 | $-4.121$ | .8 |

Note. Each column provides either coefficients from the best fit function to measured peak frequency values (see Figure 7B), for a linear polynomial function, $f(x) = x(p_1) + p_2$; a power function, $f(x) = a x^p$; or a goodness of fit measure for each of the power functions ($R^2$). The $p_1$ value indicates whether this frequency increased (positive values) or decreased (negative values) as the song progressed.
conclusions about how other species produce and use vocal sequences. In the following, we summarize the evidence revealed by our analyses and what that evidence implies about the ever-changing qualities of humpback whale songs.

Recurrence in Song Structure Challenges Claims of Cultural Transmission

Songs produced by different populations of humpback whales are often described as quite different (Helweg et al., 1990, 1998; Payne & Guinee, 1983; Winn et al., 1981), as are songs from a single population produced more than 6 years apart (Mercado et al., 2005; Payne & Payne, 1985). It was thus quite surprising to find such striking acoustic similarities between songs separated by such large temporal and physical distances. Given the wide variety of sound patterns produced by humpback whales within a single population over time (Payne & Payne, 1985), it is highly unlikely that two groups that are not in acoustic contact would end up producing the same themes in the same order by chance. There are
previous reports of acoustically isolated groups of humpback whales changing their songs in similar ways despite large physical separations (Cerchio et al., 2001; Darling & Sousa-Lima, 2005), occasionally with a lag of a few years between changes emerging in the groups (Garland et al., 2011). However, it is implausible that songs produced off the coast of Puerto Rico migrated across multiple populations to eventually be inherited by whales singing in Maui exactly 40 years later with minimal changes, especially given that there are no known contacts between these two populations. Similarly, the chances that singing humpback whales in Maui, Puerto Rico, and Colombia independently innovated the complex phrase designated by Winn and colleagues (1970) as Theme IV are miniscule. On the other hand, if humpback whales from different populations are sometimes singing the same or highly similar songs, then why have researchers previously failed to detect matching songs across populations in the past 50 years? We suspect that the dynamic nature of phrase morphing by singing humpback whales, combined with the widespread use of subjective, hierarchical approaches to characterizing vocal sequences, have made it difficult for researchers to detect the kinds of acoustic similarities in song structure identified in the current study.

The precise matching of unit timing across phrases recorded from Puerto Rico and the North Pacific revealed in the current analyses (see Figure 3) suggests that singing humpback whales may be inserting units into predetermined “time slots,” and that they likely reuse existing time slots for acoustically variable content (as is evidenced by the tight temporal coupling of consecutive phrases within singerrings). When zebra finches are learning to sing new songs, they are more successful at adding a new sound if the duration of the sound can fit into a preexisting rhythm (Hyland Bruno, 2017), suggesting that rhythms (which are largely ignored by information theoretic analyses) can provide a powerful organizing template for song learning and production. Not only did singing humpback whales from different populations produce similarly timed unit sequences, they also gradually changed the durations of units and interunit intervals in similar ways within song cycles. For instance, all three groups gradually increased and then decreased the durations of units within phrases, while maintaining the tempo of unit production across themes. Understanding how temporal templates shape the production of vocal sequences may be key to identifying the kinds of information that different species extract from vocal signals.

Debates about which animals have culture and what counts as good evidence of cultural transmission are ongoing (Acerbi et al., 2019; Allen, 2019; Aplin, 2019; Freeberg, 2000, 2001; Laland & Evans, 2017; Schuppli & van Schaik, 2019; Whiten, 2019). Some have pointed to the singing behavior of humpback whales as perhaps the clearest and most compelling case of communicative culture in mammals (Rendell & Whitehead, 2001a, 2001b). The progressive changes that humpback whales within a population synchronously make to their songs are so rapid and extensive that it is hard to imagine what mechanisms other than social learning could possibly explain them (Payne et al., 1983). The current findings suggest, however, that the changes that singing humpback whales are making to their songs are less arbitrary and creative than has generally been assumed. Other researchers have noted that the phrases produced by singing humpback whales over decades are “reminiscent” of phrases that were previously produced in the same population (Payne & Payne, 1985). And, some sound patterns, such as the “surface ratchet” theme (Figure 3, Theme VI), which Winn and colleagues (1970) found was highly correlated with singers blowing (see Figure 2), are widely acknowledged to recur across decades and populations (see also Mercado et al., 2003). There are also reports of entire songs reappearing in different populations (Garland et al., 2011; Noad et al., 2000; Owen et al., 2019), which have been interpreted as evidence of cultural transmission across groups with little genetic overlap. An alternative explanation for why certain themes or songs might recur within and across populations is that singers are not learning songs from each other, but are instead selecting from a fixed repertoire of pattern production templates that all humpback whales share.

**Progressive Phrase Morphing Suggests a Structural Template for Song Production**

Singerrings and song contour plots revealed that singers are not morphing phrases in arbitrary ways. Although the modifications singing humpback whales make to songs might subjectively seem like vocal flourishes or embellishments (Darling et al., 2019; Tyack, 1981), the fact that morphing trajectories were comparable across populations and years indicates that song transformations may be obligatory and not the result of creative improvisations by singers. Continuity in spectral and temporal qualities of unit sequences across consecutive phrases, in particular, indicates that singers often produce phrases and themes that are variants of recently produced sequences. Notably, the kinds of gradual acoustic morphing evident within song cycles are comparable with many of the progressive changes that singers make to songs within and across years (Cerchio et al., 2001; Payne & Payne, 1985; Payne et al., 1983; Winn & Winn, 1978). Single units split into pairs of units, units slowly expand and contract in duration, fundamental frequencies gradually drift into lower or higher bands, and frequency-modulation gradually shifts in direction or range. These transformations appear to be interdependent such that the relative duration and pacing of units is maintained across themes (see also Schneider & Mercado, 2019). Singing humpback whales are like the DJs of the sea, continuously morphing across a variety of unit combinations while maintaining a metronomic backbeat that connects all the variants temporally.

The fact that singers across populations made similar acoustic transformations within song cycles (e.g., gradually decreasing pitch, shortening unit durations, increasing bandwidth) argues against the possibility that production errors are driving morphing trajectories. Cyclical acoustic morphing of phrases within songs, which was evident from both subjective and quantitative analyses, is unlikely to be the result of mistakes. Similarly, the fact that the rate of within-song morphing varied across populations (and within individuals) suggests that these progressive transformations are controlled by singers rather than inadvertent. Past reports that humpback whales progressively modify song structure within and across seasons (Payne & Payne, 1985; Payne et al., 1983) led to the proposal that copying errors might account for, or facilitate, the gradual evolution of songs (McLaughlin et al., 2018). However, copying errors should amplify divergence across populations over time, unless preexisting biases distort errors in predictable ways (Griffiths et al., 2008), in which case all songs should converge to a common form. Given that neither of these outcomes is happening, it seems unlikely that accumulating production errors contrib-
ute to the progressive variations that humpback whales make to songs at any time scale.

Singing humpback whales are constantly changing their songs, while maintaining predictable acoustic morphing trajectories within song cycles. This behavior is more consistent with the possibility that learning (either social or individual) contributes little to collective, progressive variations in humpback whale songs over time than it is with the proposal that singers innovate novel sequences and/or inaccurately copy features they have heard other whales producing. The presence of predictable sound production across years and populations, like the consistent temporal structure of humpback whale songs, points toward the presence of production templates that guide how and when singers produce specific frequencies while singing. Acoustic similarities are evident not only in the overall structure of humpback whale songs, but also in how whales transform sounds over time. In particular, singers appear to morph phrases along similar trajectories, even when they are not in acoustic contact (Cerchio et al., 2001; Darling & Sousa-Lima, 2005). If singers modify phrases over time in ways that are guided by internal rules, and if they sometimes recycle sound patterns, then this further calls into question the role that social learning via cultural transmission or innovation may play in song production.

**Production Patterns Are Consistent With Heterarchical Structure**

Payne and colleagues (1983, 1985) noted early on that humpback whales appeared to be structuring sounds in ways that followed certain rules, leading to a characteristic song form. The trends they identified related mainly to the predictable ordering of themes, the similarity of phrases within themes, and the progressive changes in themes across years. The current findings revealed several additional properties of humpback whale songs that were shared across years and populations, some of which may relate specifically to the mechanisms singers used to produce units. Dividing units into two groups postulated to be produced either egressively or ingressively revealed clear alternations in the use of units from these two classes throughout all song cycles analyzed. Such alternation is commonly observed in animals that produce sounds bidirectionally (e.g., see Figure 1B). The time that singers spent producing “egressive” versus “ingressive” units was balanced overall, but fluctuated across a song cycle (i.e., in some phrases, egressive units were more prevalent and longer in duration, and in later segments of a song, this trend was reversed). Singers were not limited to simply alternating between individual “ingressive” and “egressive” units, but also alternated between clusters of units from these two classes (e.g., see Figure 2). These extended alternation patterns suggest that singing humpback whales are likely capable of producing multiple units during a single inspiration or expiration, but may not always do so. Bidirectional sound production may contribute to the continuity of temporal structure within songs as well as the uniformity evident across years and populations.

Heterarchical analysis of song structure, which replaces subjective partitioning of phrases with segmentation derived from models of sound production, revealed that morphing of “ingressive” units within song cycles may follow a different trajectory from morphing of “egressive” units, with transformations of “ingressive” units mainly involving gradual shifts in duration and in the pitch of low-pulse-rate units. Earlier automated analyses of song structure based on the dynamics of spectral shape within songs similarly showed that a subset of low-pitched units within songs remained relatively stable across a song cycle, whereas other higher pitched units changed more extensively (Mercado & Sturdy, 2017). Transformations of “egressive” units within song cycles were more extensive, precluding the possibility of reliably distinguishing them from “ingressive” units based on their acoustic properties.

Neither acoustic nor anatomical evidence can definitively show how singing humpback whales produce units; techniques for tracking air movements within singers will be necessary to confirm the contributions of bidirectional sound production to the varying structural features of songs. Regardless of how singers produced the alternating sequences of units described here, parallel streams of units that gradually morph along different acoustic trajectories are more consistent with heterarchical song organization than with hierarchical structure (Mercado & Handel, 2012). Hierarchical approaches to analyzing humpback whale songs treat themes as discrete components that singers replace over time, with no constraints on the qualities of consecutive themes. If songs are heterarchically organized, however, then the properties of unit sequences within a song cycle will reflect interacting physiological mechanisms that singers may modulate either in parallel or independently (e.g., by modifying “egressive” production without changing “ingressive” vocalizations). By disregarding many of the acoustic qualities that singing humpback whales gradually modify within song cycles, traditional approaches to describing song structure have failed to adequately characterize how humpback whale songs actually vary.

**Information Processing Approaches Obscure Some Patterns of Change**

Past comparisons of songs across populations of humpback whales that reported that different populations produce highly distinctive songs relied heavily on subjective comparisons of representative phrases from themes recorded over relatively short time periods (Payne & Guinee, 1983; Winn et al., 1981). The current analysis shows that songs produced by different populations of humpback whales in different years can be much more similar, suggesting that singers recycle themes and possibly entire songs over longer time scales. Singерings and song contour plots reveal the arbitrary nature of past approaches to subjectively partitioning humpback whale songs and translating them into sequences of symbols. For songs from Hawaii, Puerto Rico, and Colombia, spectral and temporal similarities between consecutive patterns are more evident than are disparities.

Detailed acoustic analyses of Colombian songs showed that singers: (a) consistently focused energy within multiple, shifting, focal frequency bands (Figure 8; see also Figure S1 in the online supplemental materials); (b) varied how energy was distributed within these bands across years (see Figure 9); (c) modulated unit duration in parallel with changes in song contours; and (d) concentrated much of their effort on units with energy focused within two to four narrow frequency ranges. Some of the transformational trends evident in the current analyses likely are present in many other recordings of humpback whale songs. When the first step in
an analyzing whale songs involves discarding all temporal and pitch-related metrics in favor of symbol sequences, however, comparisons between morphing trajectories become impossible. Comparisons within and across songs that do not consider spectrotemporal variations in the relationships between sounds, including the silent intervals between them, may fail to account for the qualities of received sounds that listeners find most salient. Similarly, relying on the subjective impressions of expert human listeners may fail to capture the properties of sounds that song learners encode. For instance, developmental trajectories of zebra finches imitating model songs can actually cause song elements to become subjectively more dissimilar from the model song before the song transforms into a more exact replica (Tchernichovski et al., 2001). In this case, subjective impressions of a young bird’s progress may underestimate how close the singer is to achieving a match.

Such complications in the analysis of vocal sequences are not unique to the songs of humpback whales. Information theoretic analyses of other animals’ songs may similarly hinder understanding of what vocalizers are doing when the acoustic properties of individual sounds and the temporal patterns of sound production are prematurely discarded. For instance, the songs of chickadees are syntactically simple, typically consisting of one tonal note (“fee”) followed immediately by a second note (“bee”). Sometimes singing chickadees leave off the “bee,” producing a song that is arguably simpler than any other in the animal kingdom, the equivalent of hitting a piano key once. From an information theoretic perspective, there is little point in analyzing chickadee songs, because they contain minimal variability or complexity. When the acoustic properties of fee-bee songs are measured precisely, how awareness of the precision with which they may change or how. Perhaps the changes singers make are a direct reaction to fluctuations in the local vocal actions of nearby singers, much like the movements of flocking birds and schooling fish, but occurring over much longer time-scales (Moussaid et al., 2009).

As noted earlier, phrase morphing within and across song cycles seems to involve gradual variations in how acoustic energy is distributed over time and within focal frequency ranges. Physiologically, this corresponds to gradual shifts in which sections of a listener’s cochlea will be maximally affected by incoming signals and when those effects will recur. Physically, acoustic morphing will influence which sounds are likely to be received after long-range propagation underwater, as well as where in a sound channel they are most likely to remain detectable (Mercado & Frazer, 1999; Mercado et al., 2000, 2007). Consequently, cycling through gradually morphed phrase variants effectively involves cycling through cochlear space (the position, width, and repetition of activation) as well as sweeping through physical space (e.g., the depths and ranges at which transmitted signals optimally propagate). It remains unclear whether progressive changes within songs are better suited to producing specific patterns of cochlear activation or to achieving efficient coverage of the underwater environments within which humpback whales sing. These two outcomes clearly interact for distant listeners (i.e., signals must be sufficiently intense at a receiver’s location to generate any detectable patterns within a listener’s cochlea).

The spectrotemporal dynamics of vocal sequences will play a critical role in listeners’ abilities to perceptually extract individual auditory streams from complex auditory scenes (Bee & Micheyl, 2008; Moss et al., 2014). Such streaming comes into play in contexts where multiple vocalizers are overlapping their signals while competing (Naguib, 2005), as well as in scenarios where multiple animals are echolocating searching for prey in the same habitat (Moss et al., 2014; Simmons et al., 2014). Because humpback whales’ songs often travel multiple kilometers in relatively
shallow water environments, they are highly prone to overlapping (Au et al., 2000), and likely represent an extreme case of the “cocktail party problem” in the wild. How the structure of humpback whale songs might contribute to (or obviate) this problem has rarely been considered (but see Zhang & White, 2017). Perceptual constraints on the segregation of parallel streams of units may have shaped the structure of humpback whale songs (i.e., the production templates that determine song form) as well as the acoustic trajectories along which songs change over time.

The complexity of humpback whale songs lies primarily in how listening whales are processing and interpreting what they hear, not in how aesthetically impressive song qualities are to human observers attempting to make sense of them. Here, we describe dynamic acoustic properties of songs that simultaneously increase the potential complexity of these sound sequences (by showing that every pattern within a song cycle may be a progressively modified variant of a preceding pattern), while at the same time simplifying them (by showing that multiple acoustic features of songs may be conserved across years and populations). These properties do not diminish the remarkable nature of what singing humpback whales are doing vocally. Any animal that can produce sounds spanning eight octaves, intensely enough that the sounds can travel 10+ kilometers, and that can do so continuously for periods of 20+ hours, is clearly deserving of respect, regardless of the complexity of the sequences they are producing. Although our results call into question the role that vocal imitation or cultural transmission may play in the dynamic qualities of humpback whale songs, they do not resolve the question of why singers are constantly on the move acoustically. Social interactions actually certainly are contributing to how variations in songs progress (Cholewiak et al., 2018; Chu & Harcourt, 1986; Winn & Winn, 1978), and learning through practice and observation seem likely to contribute to what individual singers can achieve vocally (Rothenberg, 2008). Vocal imitation abilities, which humpback whale songs seem likely to possess, may contribute to the flexibility and speed with which singers can modify their songs in ways beyond simple replication of models. For instance, self-imitation of vocal signals may increase the precision with which a singer can refine the acoustic details of the patterns a singer is “repeating,” and the ability to vocally imitate may enhance a listener’s ability to judge their distance from other whales (Mercado et al., 2014; Morton, 1998). Given that so little is currently known about how or why singing humpback whales are controlling the sounds and sound patterns they produce, information theoretic analyses of their vocal sequences should probably be avoided (Nizami, 2019), unless they provide a way to directly test theoretical predictions (see Freeberg & Lucas, 2012), and until more is known about which acoustic features are salient for intended listeners.

References


Received June 24, 2020
Revision received September 21, 2020
Accepted October 23, 2020

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at https://my.apa.org/portal/alerts/ and you will be notified by e-mail when issues of interest to you become available!