Aquatic Mammals 2007, *33*(2), 202-213, DOI 10.1578/AM.33.2.2007.202

Characterizing Patterns within Humpback Whale (*Megaptera novaeangliae*) Songs

Sean R. Green,[1] Eduardo Mercado, III,[1] Adam A. Pack,[2, 3] and Louis M. Herman[2, 3]

[1]*Department of Psychology, University at Buffalo, State University of New York, Buffalo, NY 14260, USA; E-mail: srgreen@buffalo.edu*
[2]*The Dolphin Institute, 420 Ward Avenue, Suite 212, Honolulu, HI 96814, USA*
[3]*University of Hawai'i Manoa, Honolulu, HI 96822-2294, USA*

Abstract

Male humpback whales (*Megaptera novaeangliae*) produce long songs that contain predictably repeated sound patterns. Other animals (including humans) identify patterns in acoustic sequences based on regularities in transitions between sounds. The present study examined transitional probabilities within humpback whale songs to determine whether relative acoustic changes from unit to unit are sufficient for identifying repeating patterns within humpback whale songs. To identify such patterns, four humpback whale songs were analyzed by first classifying song units using a self-organizing map, and then calculating transitional probabilities based on these classifications. Two separate analyses of transitional probabilities were conducted: one involved units classified based on their absolute acoustic features (e.g., duration, peak frequency, and amplitude) as well as changes in these features relative to adjacent units, and the other used units classified based on the relative changes alone. Both analyses revealed repeated sequences of units within humpback whale songs, but the analysis based on relative changes alone yielded a larger number of predictable transitions. This finding suggests that relative acoustic changes within humpback whale songs may provide robust indicators of repeating patterns.

Key Words: mysticete, baleen whale, cetacean, singing, display, vocalization

Introduction

Male humpback whales (*Megaptera novaeangliae*) produce long series of sounds, which are known as song. These sound sequences have a predictable structure that can be described in terms of a hierarchy of components ranging from individual sounds (units), through phrases and themes, to the song itself, which may repeat many times in the course of a song session (Payne & McVay, 1971). The properties of humpback whale song suggest that its function may depend on the sequential organization of song units.

One aspect of humpback whale song that has impeded understanding of its temporal structure is the fact that, while whales in a given location typically sing quite similar songs (Winn & Winn, 1978; Guinee et al., 1983; Payne & Guinee, 1983; Payne & Payne, 1985), the individual sounds that constitute the songs can vary considerably from one year to the next (Payne & Guinee, 1983; Helweg et al., 1998; Cerchio et al., 2001; Mercado et al., 2005). Therefore, it is unlikely that the function and informational content of humpback whale songs, if any, depend solely on the invariant acoustic features of the song components. One source of information that could be stable despite changes in the absolute features of song units is the change from one unit to the next. The aim of the present study is to investigate the possibility that such changes might be useful for identifying patterns within a song.

Little is known about what listening whales learn from songs, particularly with respect to the temporal structure. Listening whales undoubtedly detect whales that are singing nearby, and probably acquire information about the approximate location of the singer. If the patterns within a whale's song convey specific information to other whales, the listeners should have some way of detecting patterns despite the changing repertoire of song units. At a minimum, they need to be able to know what parts of the song constitute a pattern and to identify patterns that may have different meanings or functions.

One mechanism through which whales might learn to recognize these patterns is statistical learning. Research suggests that when humans interpret speech, they consider not only the features of an individual sound but also its context—that is, the sounds or silent intervals that precede or follow the utterance. Research using connectionist models of spoken word recognition demonstrated that a phoneme's temporal context can guide recognition and syntactic processing (Rumelhart & McClelland, 1986). Consistent with this finding, infants use the statistical distribution of phoneme transitions as cues to word boundaries (Saffran et al., 1996; Aslin et al., 1998). To illustrate, consider the phrase "humpback whale song." The word "whale" occurs many times in this document. As a result, "wh" is often followed by "a," which is usually followed by "le." Therefore, the "wh" - "a" and the "a" - "le" transitions occur relatively frequently. In contrast, transitions across a word boundary—in this case, "ck" – "wh" and "le" – "s"—occur less frequently because there are a variety of different words that could precede or follow "whale." In a similar fashion, one might expect to find high transitional probabilities between humpback whale song units that occur within a phrase or theme and low transitional probabilities between units that cross a theme or phrase boundary. Prior research has used analyses of transitional probabilities to demonstrate that themes within humpback whale songs are produced in a predictable order (Payne et al., 1983), but this approach has not previously been extended to the analysis of unit sequences.

In the case of whale song, unit classification is a prerequisite for analysis of transitional probabilities within phrases since it is not possible to determine the likelihood of a transition from Unit A to Unit B without first developing a method for clearly determining which song units are to be classified as Type A or Type B. For this process, the choice of dimensions along which one classifies the song units is important. If the songs are analyzed using measures of acoustic features that are incidental to whatever pattern is present, then the analysis is unlikely to identify patterns. Conversely, if certain measures are integral to patterns contained within the song, analyses based on classification schemes that focus on those measures should be more likely to identify patterns. There have been many previous attempts to categorize whale song units. Early investigations of whale sounds often relied on subjective human judgments of the acoustic signal (Winn & Winn, 1978) or on quantitative analyses of acoustic features (Clark, 1982; Chabot, 1988; Mednis, 1991; Potter et al., 1994). More recently, a type of neural network known as a self-organizing map, or SOM (Kohonen, 1990), has been used to objectively classify whale song units (Walker et al., 1996; Mercado & Kuh, 1998; Suzuki et al., 2006).

A SOM consists of multiple processing units (called nodes) that each respond selectively to particular input features. Like other neural networks, SOMs can be trained to classify inputs—in this case, measures of humpback whale song units—by repeatedly presenting those inputs to the SOM. With training, nodes in the SOM become retuned such that they selectively respond to features that are prevalent within the input set. For example, if a SOM is trained with measures of whale song units, some nodes may become active only when the inputs correspond to low frequency, high amplitude, and long duration sounds. Other nodes may become active only when the inputs are measures from high frequency, short duration sounds. Consequently, song units can be described in terms of the SOM nodes that they activate. The advantages of SOM-based classifications of song units are that they are quantitative, automatic, and objective. Furthermore, since SOMs classify a data set on a clearly defined set of dimensions, the success or failure of otherwise identical SOMs at producing useful classification schemes when different measures are used provides an indication of how useful those measures are for describing the data.

In the current study, SOMs were used to classify song units as a first step toward determining transitional probabilities within songs (see also Suzuki et al., 2006). The goal of these analyses was to evaluate whether measures of relative changes in units can be useful for identifying repeated patterns in whale songs. The hypothesis is that relative changes from one unit to the next, rather than the absolute features of each individual unit, can, in principle, serve as a cue to song structure. It is important to note that these analyses do not address the question of whether this approach will be useful for characterizing all humpback whale songs. It is clear that songs produced by different whales in various contexts, years, and geographical regions may differ in many ways. Whether characterizations of songs based on relative changes across units can encompass the full range of structural variability seen across individuals, years, and populations in an informative way is an empirical question that is only worth pursuing if there is some evidence supporting the hypothesis noted above. The aim of this study is to test this hypothesis.

Comparative evidence suggests that perception of relative change may play a greater role in mammalian auditory perception than has generally been recognized. For example, dolphins appear to be sensitive to both relative and absolute features of tone sequences (Ralston & Herman, 1995) and variations in amplitude (Richards et al., 1984). Similarly, human infants can recognize melodies based on relative frequency changes despite overall changes in absolute frequency (Plantinga & Trainor, 2003). If humpback whales are also sensitive to patterns of relative change within sound sequences, then one would expect that relative changes across song units might be as important as (or even more important than) the absolute acoustic features of individual units. Conversely, if patterns of relative change are inconsistent across phrases, it is less likely that they contribute to the function of humpback whale song.

Materials and Methods

To investigate the usefulness of basic acoustic features of units (e.g., intensity, frequency, and duration) and relative change in these features between units as cues to the temporal structure of humpback whale song, two separate analyses of whale songs were conducted. In the first analysis, a SOM classified song units using acoustic features of units as well as the change in these features from one unit to the next prior to statistical analysis of the song sequences. To gauge the contribution of relative changes in acoustic features, a second SOM used only the relative change across units as a basis for classification. If the changes in acoustic features from one unit to the next were incidental to the structure of phrases, then statistical analyses based on this second SOM should be less effective at identifying patterns within songs. If, on the other hand, relative changes were a primary source of cues to phrasal structure, then statistical analyses based on the second SOM should be equally or more effective at identifying patterns within songs.

Classification of Song Units

Four humpback whale songs recorded in 1992 by researchers at the Kewalo Basin Marine Mammal Laboratory were analyzed in the current study; the songs were recorded on different days (29 January, 13 February, 13 March, and 24 March) from whales wintering off the coast of Hawaii. Song 1 was 739 s in total duration. Song 2 was 900 s long, and Songs 3 and 4 were 604 s and 1,081 s long, respectively. These songs were chosen based on recording quality and because earlier subjective and quantitative analyses established that these songs contained recognizable patterns (Mercado et al., 2003). They are structurally comparable to humpback whale songs that have been described in many prior reports (i.e., they all contain themes produced in a predictable order, which are composed of repeated phrases, which consist of stereotypical sequences of units). This sample is not suitable for statistically assessing whether the chosen songs are representative of songs from this region or time period, or for comparing features of these songs with those of songs from other regions or time periods, and no attempts at such comparisons are made in the current analyses.

Recordings were sampled at a rate of 22,050 Hz (note, however, that spectral components above 4,000 Hz were rarely present in these particular recordings). An automated sound segmentation program, running in *MATLAB*, isolated song units from recordings based on signal amplitude compared to the amplitude of the silent intervals between units. The threshold for setting the start point of a unit was arbitrarily defined as the point where the signal amplitude over a 5-ms interval was three times the amplitude of the "background noise" separating the previous two units, and the stop point was defined as the point where the average amplitude for a 5-ms segment dropped below twice the noise value. In cases where the start and stop point for a given unit did not coincide with the subjective judgment based on listening to the sound and viewing its waveform and spectrogram, the unit boundaries were changed to match the subjective judgment. This was done because background noise occasionally led the algorithm to assign unit boundaries incorrectly. While the choice of unit boundaries will have some impact on each analysis, it should be noted that the same set of unit boundaries was used for each simulation, so comparisons between the two simulations were unlikely to have been systematically affected by the choice of boundaries.

Once unit boundaries were set, the program computed the duration, amplitude, and peak frequency (as measured from the power spectral density function) for each unit. The temporal separation between each unit and the relative change in frequency, duration, and amplitude between a unit and its neighbors also were calculated. Automated segmentation was verified by listening to recordings and visually monitoring amplitude plots.

Once individual units had been isolated, input feature vectors were generated for each unit. For the first analysis, each vector contained eleven elements: absolute peak power spectrum density, duration, amplitude, the duration of the gap between the preceding and following unit, and the relative change in peak power spectrum density, duration, and amplitude relative to the preceding and following units. In the second analysis, the measured elements of absolute peak power spectrum density, duration, and amplitude were omitted. For all units analyzed, each vector element was normalized such that the mean value across all vectors was 1. This was done to prevent elements with a large range of values (e.g., peak frequency, which ranged from 16 to over 3,000 Hz) from dominating the network relative to elements that covered a narrower range of values. The mean-squared amplitude of the sound waveform served as a measure of sound intensity. Note that this measure is not intended to reflect the source level of the sound but, rather, the received sound level at the recording location. Variations in distance from the singing whale, recording sensitivity and other factors will cause received sound level to differ from source level. Nevertheless, past analyses have shown that received levels vary systematically across units (Mercado et al., 2003; Au et al., 2006). Measurements of received level are useful for assessing what listening whales would hear.

Relative changes in frequency, duration, and amplitude were represented as log ratios to ensure that high values would not have disproportionate effects on the mean relative to their reciprocals, thereby distorting the SOM. Another advantage of the log ratio transformation is that symmetrical relationships are represented with symmetrical values. That is to say, if the frequency of one unit is 10 times higher in frequency than the previous unit, then the measure of the frequency of the second unit relative to the first would be 1 and the measure of the relative frequency of the first unit relative to the second unit would be -1.

Network Implementation

SOMs were implemented via the *newsom* function in *MATLAB*'s Neural Network Toolbox. Each SOM consisted of 36 nodes, arranged in a 6 × 6, two-dimensional hexagonal grid. SOMs were trained for 200 epochs (the term "epoch" refers to the successive presentation of each of the input vectors to the SOM in a random order). Note that all samples were used in training. Unlike supervised neural networks, which are typically presented with a subset of the data in order to train them to produce specific outputs, the SOM learns in an unsupervised fashion, without user feedback. Consequently, it is not necessary to pretrain SOMs with a subset of the data.

At the beginning of training, nodes within the SOM were initially given random values for each of the input dimensions. During training, a standard algorithm calculated which node was most similar to the input (using *MATLAB*'s *linkdist* function). The most similar node and surrounding nodes were then automatically modified so that they all became more similar to the input. Two parameters, called the learning rate and the neighborhood size, controlled how the selectivity of nodes was adjusted. The learning rate (set at 0.2) determined the degree to which nodes were modified after responding to an input, and the neighborhood size (set at 1) determined how many nodes were modified. This process was repeated for each of the inputs within each epoch.

How a trained SOM classifies song units is determined by what the neural network learns during training. Consequently, individual nodes within a trained SOM do not always correspond to subjective categories. One way to assess how a SOM is classifying song units is to determine which input vector generates the strongest response at each node. The unit that most strongly activates a particular node is generally prototypical of the kinds of units that will activate that node. For example, Figure 1 shows a 6 × 6 grid of spectrograms, each of which corresponds to the unit that best activates the node in the corresponding location of a trained SOM. For this particular SOM, short duration units were likely to activate a node on the right side of the map, and long duration units were likely to activate a node on the left side of the map. In the current analyses, every song unit was classified based on the node within a trained SOM that it activated most strongly. Specifically, each unit was assigned a number based on which of the 36 nodes within the SOM became active when the unit's feature vector was used as an input. In this way, sequences of units were automatically transformed into sequences of numbers.

Analysis of Transitional Probabilities

After sequences of song units were transformed into sequences of numbers using SOMs, transitional probabilities within these sequences were analyzed to determine whether they contained repeating patterns like those found via subjective analyses of song structure. Based on prior studies, one would expect hierarchical structure to be evident in the temporal distribution of units in the song. Viewed in the context of a 6 × 6 map of nodes, instances of a given phrase or repeated sequence should follow similar trajectories within the SOM multidimensional space. Moreover, within a given phrase, the probability that one class of song unit follows another should not be random, but, instead, should be determined by the internal structure of the phrase. The transitional probability from one unit to the next should thus be much higher than would be predicted by chance. In contrast, unstructured regions of humpback song, or boundaries between phrases, should contain more low-probability transitions. To examine whether this was the case, the frequency of occurrence for each of the possible two-unit pairings, or bigrams, was calculated, followed by the frequency of occurrence for each bigram as a percentage of all bigrams beginning with the same unit. This provided a measure of the transitional probability for each unit transition in each of the songs, where transitional probability is defined as the proportion of occurrences in which one unit type is followed by another (Aslin et al., 1998).

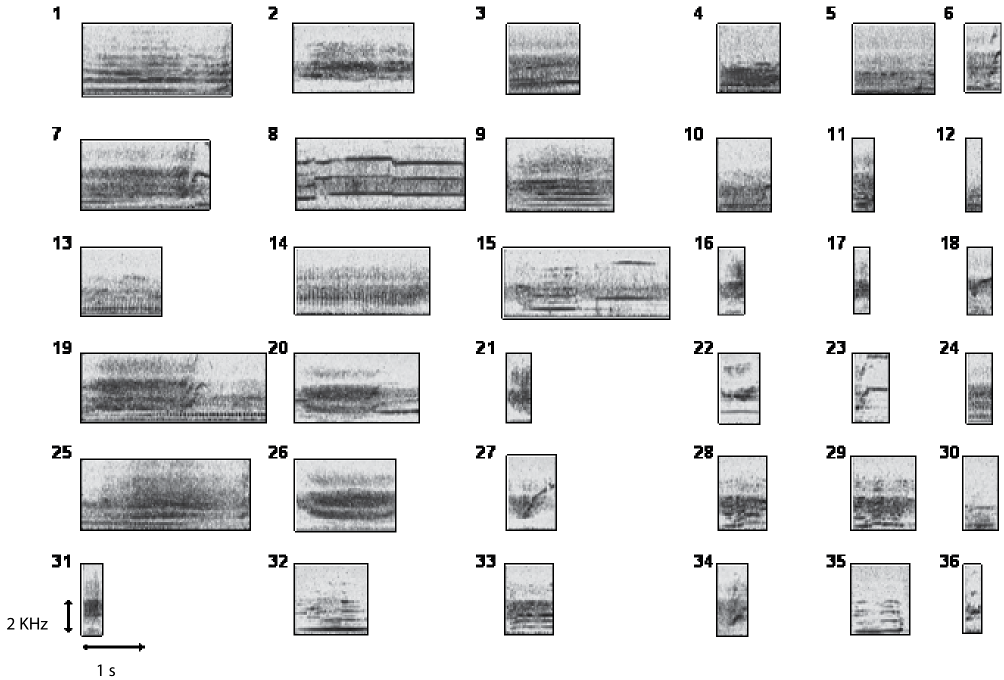

Figure 1. Spectrograms of the humpback whale song units that most strongly activate each of the 36 nodes in a SOM trained to classify units based on relative changes; note the scale located by the leftmost node on the bottom row.

Results

Feature vectors were generated for 1,183 song units collected from the four songs. The distributions of peak frequency, duration, and amplitude for these units are shown in Figure 2 for each song. The distribution of peak frequency was bimodal, with a local minimum at approximately 250 Hz and a wide plateau from approximately 1,200 to 2,000 Hz. The distribution of duration was positively skewed and unimodal, with a peak at approximately 0.5 s. These distributions are similar to those found by Mercado et al. (2005), who also showed a large number of units with peak frequencies around 250 Hz and a plateau extending to approximately 2,250 Hz, although these regions were not separated by a local minimum as they were in the current analysis. Duration in the earlier study also was positively skewed, though the modal duration was greater than 1 s. The fact that the current sample contains more units of shorter duration could indicate that the threshold used here for determining start and stop times was higher or that the sample used by Mercado et al. happened to contain more units of longer duration. Amplitude showed a fairly symmetrical but ragged distribution; variations in whale distance and in the quality of the recordings undoubtedly affected these distributions. When the relative change from one unit to adjacent units is plotted for frequency, duration, and amplitude for each song, the distributions are symmetrical and differ mainly in the number of values near zero (see Figure 3).

The spatial distribution of nodes within the SOM trained with both directly measured features and relative changes in units revealed that individual nodes were selective for variations in peak frequency, duration, and amplitude. Specifically, the upper left corner of this map contained nodes that responded best to units with peak frequencies below 500 Hz, while nodes grouped along the lower and rightmost edge of the map responded best to units with higher peak frequencies. Nodes that responded best to long duration units were found in the upper right corner of the map. Nodes that responded best to high amplitude units were found in the lower right quadrant. The spatial structure of the trained SOM suggests that the distributions of unit frequency, duration, and amplitude were not systematically related. If they had been, then nodes selective for units with correlated features should have been present (e.g., if most long duration units contained low peak frequencies, then units selective for these features should have been found in the same regions of the SOM).

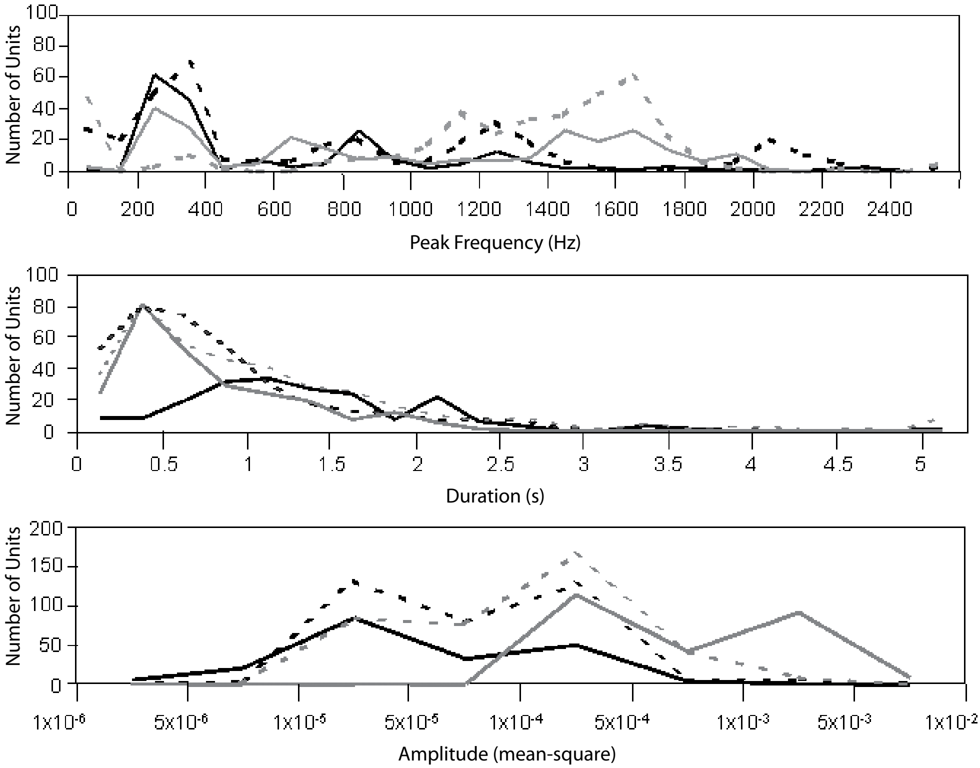

Figure 2. Distribution of measures of peak frequency (upper), duration (middle), and amplitude (lower) of humpback whale song units recorded in 1992; amplitude is the mean square of the amplitude values for the .WAV file for each unit. The solid black line indicates measures from Song 1 (29/1/92), the solid gray line corresponds to Song 2 (13/2/92), the dashed black line indicates measures from Song 3 (13/3/92), and the dashed gray line corresponds to units from Song 4 (24/3/92).

Analysis of Temporal Structure Based on Direct Measures and Relative Changes

When song units were classified using the SOM described above, the distribution of unit transitions was not randomly distributed within songs ($\chi^2_{(1,295)} = 3{,}535$, $p < 0.01$). Most unit transitions tended to involve particular subsets of SOM nodes. This indicates that the temporal arrangement of units within each song plays a role in song organization and that this temporal structure is discernable through analyses of SOM-classified units.

Song 1 (recorded on 29 January 1992) contained prolonged stretches of high-probability transitions consisting almost entirely of repetitions of three-unit sequences. Analyses of other songs did not reveal repeating patterns; this result does not imply that longer patterns were not present, only that they were not evident in this particular analysis. High-probability transitions in these songs typically occurred in isolation or in pairs, suggesting that two- or three-unit sequences may be a fundamental component of their phrase structure.

Although the analysis of transitional probabilities did not show many individual patterns longer than three or four units, the repeating trigrams found in Song 1 were startlingly clear and merit further discussion. In Song 1, there were 26 repeats (along with several near misses) of a single three-unit sequence, associated with Nodes 6, 3, and 15 respectively, comprising approximately one-third of the song. Considering all units in terms of their nearest nodes, those associated with Nodes 6, 3, and 15 were dominated by Song 1 units, in spite of the fact that Song 1 was the shortest by far of the four songs analyzed. This demonstrated that the SOM was sensitive to at least some patterns within humpback whale song, although it appears that for Songs 2 (recorded on 13 February 1992), 3 (recorded on 13 March 1992), and 4 (recorded on 24 March 1992), which were somewhat less clear than Song 1, the net was not sensitive to overall theme and phrase structure. This may be due to the comparative difficulty of determining a unit's start or stop point, or even that a unit exists, amid at times considerable background noise. Further study using a larger sample of whale song may

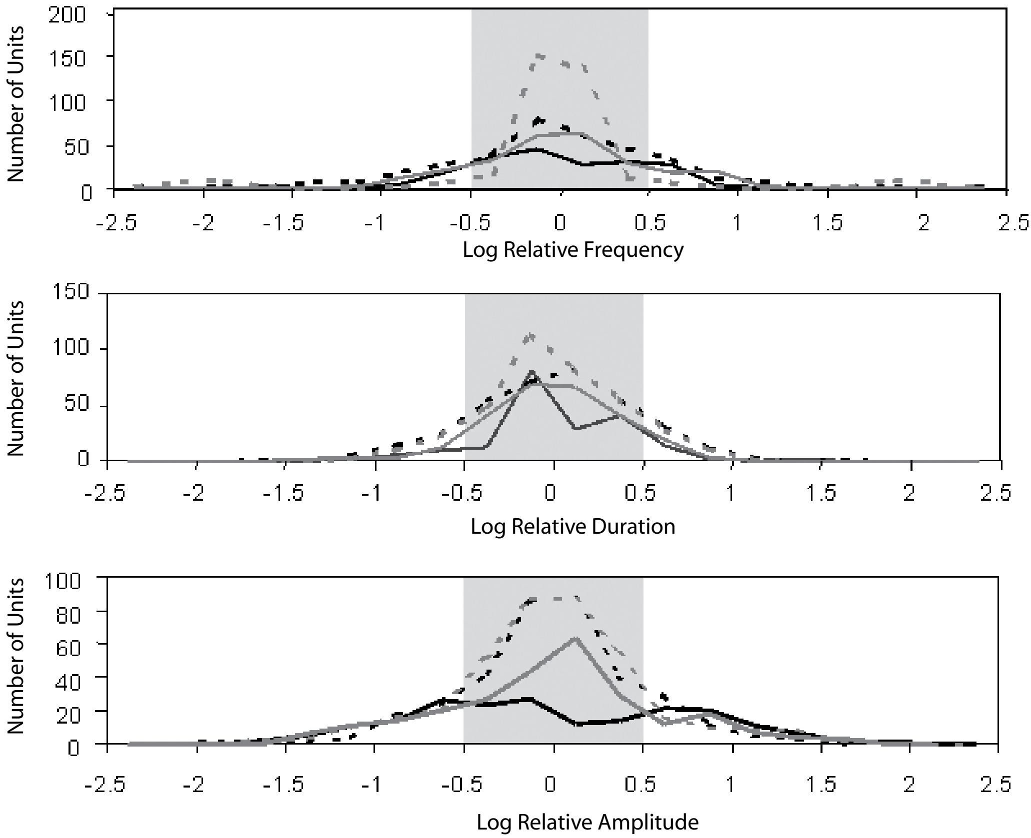

Figure 3. Distribution of measures of relative log frequency (upper), duration (middle), and amplitude (lower); positive values indicate that a unit has a higher value than the previous unit. The solid black line indicates measures from Song 1 (29/1/92), the solid gray line corresponds to Song 2 (13/2/92), the dashed black line indicates measures from Song 3 (13/3/92), and the dashed gray line corresponds to units from Song 4 (24/3/92). The region from -0.5 to +0.5 has been shaded to illustrate the symmetry of the distributions.

give a clearer understanding of which whale songs are more amenable to SOM analysis.

The repeating three-unit sequence in Song 1 described above is informative because it provides a way of ascertaining which acoustic dimensions of units are important for establishing the temporal pattern. The SOM differentiated between individual units based on differences in the 11 features represented by the input vector. Initially, it was unclear whether directly measured features or relative changes in features were more important as cues for discriminating between units. To address this question, and to investigate whether each of the three sound features—peak frequency, duration, and amplitude—are important in this repeating pattern, features of the units in the repeating sequences were analyzed. Features that are important for identifying this repeating pattern should be consistent across repetitions of a unit, but different from instances of the other units to accentuate the cyclical pattern.

As shown in Figure 4, the peak frequency, duration, and amplitude of the three units overlapped considerably, suggesting that there were few discrete, stereotypical acoustic features that defined the components of the pattern. In contrast, for measures of relative change in these features, as well as for the silent intervals between units, the range of at least one of the three units did not overlap with its neighbors. This result indicates that in this case, relative changes in the peak frequency, duration, and amplitude were more important for classifying whale song units in repeating patterns compared to directly measured features.

Analysis of Temporal Structure Based on Relative Changes Alone

To further assess the role of relative changes across song units in the temporal structure of whale songs, a second analysis was conducted in which a SOM was trained to classify units based on input

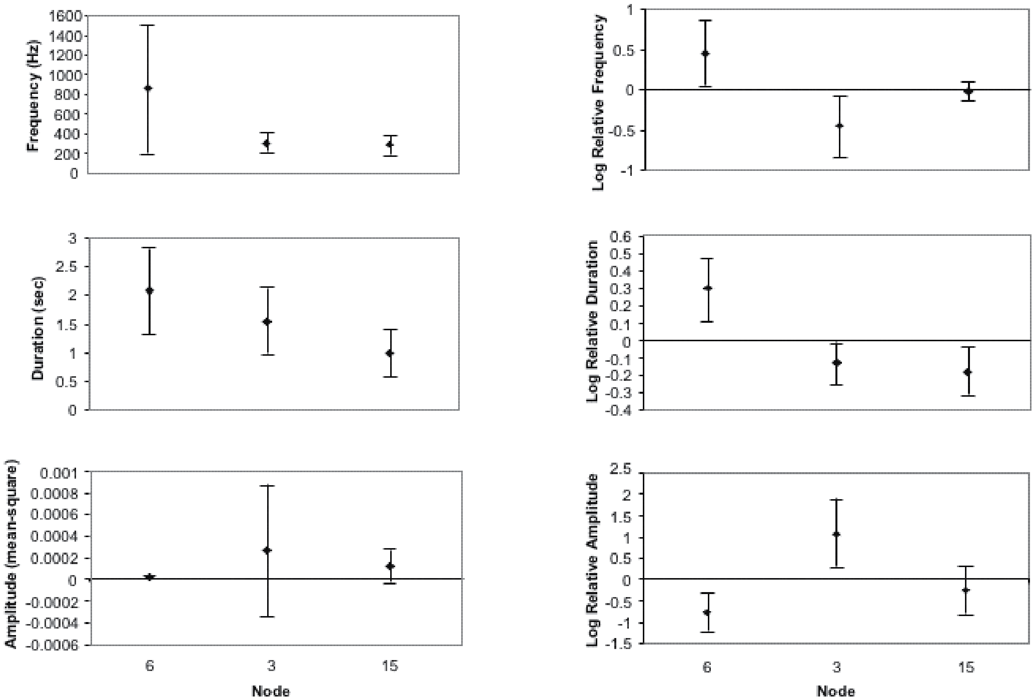

Figure 4. A comparison of the range (error bars indicate ± 2.5 SD) of directly measured and relative changes for 23 repetitions of the 6-3-15 unit sequence; there is considerable overlap in the range of the directly measured features, but the ranges of relative features do not overlap for some elements of the sequence.

vectors that did not contain direct measures of the acoustic features of those units. The same set of isolated whale song units used in the first analysis was also used in this analysis; however, each input vector used to train the SOM in this analysis contained only eight of the original 11 elements (duration of preceding gap, duration of following gap, preceding/following log relative frequency, preceding/following log relative duration, and preceding/following log relative amplitude)—that is, the three elements containing direct measures of the duration, peak frequency, and amplitude of each unit were no longer included as part of the input vector.

Nodes in this second trained SOM were less obviously organized than in the previous analysis. Figure 1 shows spectrograms of a representative unit for each map node. Units that lasted longer than one of their neighbors tended to occupy the upper left quadrant of the SOM, and relatively shorter units occupied the lower right quadrant. The relative change in amplitude values of the nodes showed the most coherent distribution, with local maxima in the lower-left quadrant and local minima in the upper-right, separated by a band of no change or rising/falling trends. It is possible that this pattern of unit distribution did not emerge in the previous analysis because differences in amplitude related to distance or recording differences overwhelmed patterns across units. The other corners of the map were also amplitude maxima and minima, revealing further evidence of an influence of relative amplitude. This analysis suggests that classification of song units based solely on relative features reveals patterns that are similar to, but distinct from, those obtained when both direct measures and relative changes are considered.

Transitional probabilities in sequences generated by this second SOM were not randomly distributed overall ($\chi^2_{(1,295)} = 6{,}025$, $p < .01$). Relative to the first analysis, more of the unit types (as classified by the SOM) were followed by a distribution of units that would not be likely to occur by chance ($p < .05$; 94% vs 75% of unit types). As with the first analysis, sequences generated by the second SOM revealed a series of repeating triplets in Song 1 (20 repetitions of the sequence 1-32-35) as well as several similar triplets. Across all songs in this analysis, there were 270 (23%) high-probability transitions (a transition A-B was considered high-probability if 25% or more of the bigrams that start with Unit A were of the Type A-B). In comparison, 136 (12%) high-probability transitions were found in the first analysis. The number of high-probability transitions for each song is shown in Figure 5. A list of the most frequent of these sequences appears in Table 1. Together, high-transitional probability sequences with transitional probabilities greater than 25% comprised 453 units, or 38% of the total content of the four recordings.

With the exception of the repeating sequence in Song 1 mentioned above, all of the most common high-probability bigrams listed in Table 1 occurred in more than one song. This argues against the idea that these patterns are artifacts or that these sequences are an idiosyncrasy of Song 1. The nature of these high-probability transitions can be understood by considering the nodes that are associated with them (in the following description, numbers refer to units that activate particular nodes, as shown in Figure 1). For example, in the 33-18 sequence, which occurs 17 times, the "33" unit is higher in amplitude and lower in frequency than its immediate neighbors, whereas the "18" unit is lower in amplitude and higher in frequency than its neighbors. Another common sequence,

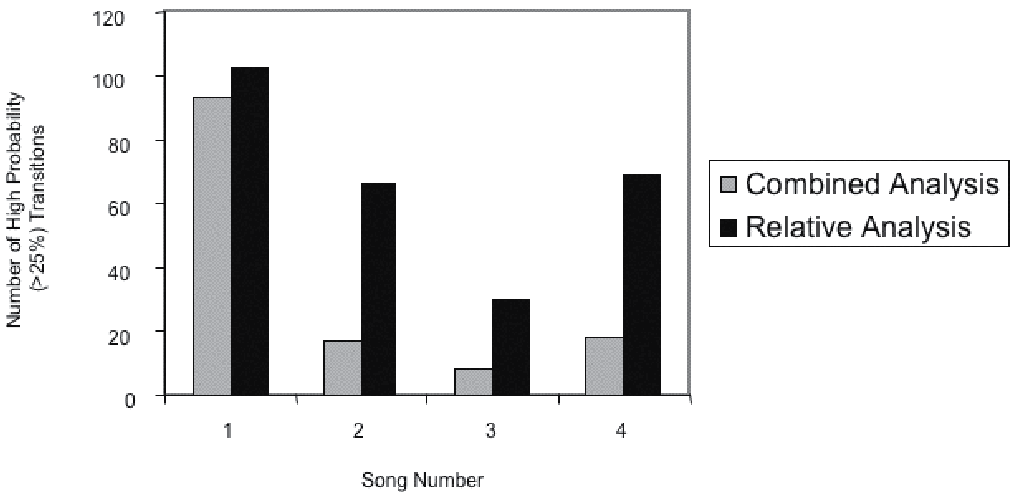

Figure 5. Number of high-probability (> 25%) transitions found using two different SOMs for classifying units within humpback whale song; fewer high-probability transitions were found when units were classified using directly measured acoustic features, gap duration, and relative change in features (gray bars) than when units were classified using gap duration and relative feature change alone (black bars).

Table 1. Tally of common repeating sequences in songs revealed when units were characterized in terms of relative changes

Sequence	Total	Song 1	Song 2	Song 3	Song 4
1-32-35	20	20	0	0	0
33-18	17	8	3	4	2
30-13	17	2	7	2	6
14-36	16	3	5	4	4
17-31	16	1	0	4	11
25-6	13	1	7	0	5
19-18	10	1	5	2	2
23-15	9	0	5	3	1
23-5	9	0	4	1	4
36-16	9	1	3	0	5
5-24	8	0	1	0	7
9-30	8	0	3	0	5
10-19	8	1	5	2	0
12-26	8	1	4	1	2
1-32	8	5	2	0	1
15-24	7	0	3	1	3

the 14-36 sequence was not marked by clear differences between the units except in duration relative to the following node. The "36" unit was shorter than the node that preceded it, while the "14" unit in fact had the longest duration relative to its following unit of all the network nodes. The 14-36 bigram thus indicates a decrease in duration in the absence of a marked change in frequency or amplitude.

As can be seen from these examples, the magnitude of feature change across the whale song units varies from one high-probability sequence to the next. Some permutations of change in duration, frequency, and amplitude are over-represented relative to others. The SOM trained with relative features of song units was more effective at isolating these patterns compared to the first SOM that took into account both direct and relative measures of whale song features.

Discussion

Despite the fact that individual humpback whale songs tend to contain repeated patterns, the basic units and patterns that make up these sequences typically vary from year to year (Payne & Guinee, 1983; Payne et al., 1983; Helweg et al., 1998; Cerchio et al., 2001). The current analyses investigated whether classifying songs based on the relative change between units could effectively detect patterns within whale song. This would in principle allow the properties of units to be flexibly changed, either to accommodate environmental constraints (Norris, 1995; Mercado & Frazer, 1999) or to allow for seasonal variation for other reasons. The current analyses demonstrate that classifying units in terms of their differences from preceding and following units reveals patterns in the temporal sequence of song units that are not evident when units are classified based on their directly measured acoustic features. Specifically, the second analysis, in which only the relative changes from one unit to the next were used as a basis for classifying whale song units, outperformed the first analysis.

Examination of transitional probabilities across song units revealed certain predictably repeating elements in both analyses, most notably a repeating three-unit phrase in Song 1. This phrase involved a complex series of changes to amplitude, peak frequency, and duration across units. Consistent with the idea that repeated patterns can be best characterized by the relative change between units within each pattern, transitions from one unit to another within a phrase were often more consistent than the absolute features of the units themselves. Additionally, the second analysis revealed a greater number of repeating units, possibly phrases or subphrases. It should be noted that these analyses do not support conclusions about whale song in general since other songs by other singers in other years may show substantially different patterns. These analyses instead demonstrate that classification by relative acoustic features can successfully identify patterns. Therefore, methods of classification based on relative measures may be more effective in detecting temporal patterns that would be relevant to listening whales as an alternative to or in conjunction with methods that depend on features of the units in isolation.

In addition to serving as potentially effective cues in quantitative analyses of whale song structure, patterns within songs that are defined by relative changes may have more utility for whales because environmental conditions and propagation-related distortion can obscure the acoustic features of song units. For example, the perceived loudness of a unit will vary depending on the listener's distance from a singer. Because many song units show a gradual increase in amplitude at the beginning of the unit and a gradual decrease at the end (Au et al., 2006), song units from distant whales may seem shorter in duration if only the middle of the unit reached the threshold for detection. Pitch should also be affected, both because different frequencies do not propagate equally well across long distances (Dusenbery, 1992; Mercado & Frazer, 1999) and because there may be individual differences in pitch among singing whales.

If the patterns defined by relative change across units prove to be stable across years, further study of these patterns could provide the basis for a new understanding of phrasal structure in whale song. Under previous systems for describing songs (e.g., Payne & McVay, 1971), changes in acoustic features from one year to the next may have obscured similarities in phrases across years associated with relative changes across units. If so, this raises questions as to the relationship between absolute acoustic features and patterns of change across units. For instance, to what extent do repeated phrases defined by relative changes vary with respect to the absolute acoustic features of their constituent units? For such phrases, do the acoustic features change in some predictable way depending on year, location, or other factors such as a whale's acoustical or social environment?

The answers to the above questions may provide some indication of the function of repeated phrase structure within songs, which has been the subject of considerable debate. Some researchers argue that the songs are mating displays (for review, see Helweg et al., 1992; Au et al., 2000), while others argued that humpback whale song is used for echolocation (Mercado & Frazer, 2001) and echo-ranging (Clark & Ellison, 2004). Guinee and Payne (1988) argued that patterns in whale song made the songs easier to remember. Given the various functions that acoustic signals serve throughout the animal kingdom, it is likely that other functions are possible. If researchers can identify fundamental organizing principles in whale song that generalize over time and across geographic location, this knowledge may clarify why whales modify the acoustic features of their repertoire over time. To illustrate, the mating display hypothesis makes no claims about song structure other than that "better" singers should have a better chance of mating or discouraging competitors. In contrast, if humpback whales use their songs for echolocation, then the acoustic characteristics of units should be tailored to environmental conditions rather than to the proclivities of listening females. Alternatively, if songs convey information, whales should employ certain themes or phrases in certain contexts, locations, or situations.

Acknowledgments

We thank the many researchers and volunteers at the Kewalo Basin Marine Mammal Laboratory who took part in obtaining the recordings analyzed in this study. All song samples were obtained under National Marine Fisheries Service Permit Nos. 223, 576, 810, and 941. We would also like to thank Ann Zoidis, Jennifer Schneider, and one anonymous reviewer for their comments, suggestions, and criticism of previous drafts of this paper. This work was supported in part by a grant from Earthwatch to L. Herman.

Literature Cited

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month old infants. *Psychological Science*, *9*, 321-324.

Au, W. W. L., Mobley, J., Burgess, W. C., Lammers, M. O., & Nachtigall, P. E. (2000). Seasonal and diurnal trends of chorusing humpback whales wintering in waters off western Maui. *Marine Mammal Science*, *16*, 530-544.

Au, W. W. L., Pack, A. A., Lammers, M. O., Herman, L. M., Deakos, M. H., & Andrews, K. (2006). Acoustic properties of humpback whale songs. *Journal of the Acoustical Society of America*, *120*(2), 1103-1110.

Cerchio, S., Jacobsen, J. K., & Norris, T. F. (2001). Temporal and geographic variation in songs of humpback whales, *Megaptera novaeangliae*: Synchronous change in Hawaiian and Mexican breeding assemblages. *Animal Behaviour*, *62*, 313-329.

Chabot, D. (1988). A quantitative technique to compare and classify humpback whale (*Megaptera novaeangliae*) sounds. *Ethology*, *77*, 89-102.

Clark, C. W. (1982). The acoustic repertoire of the southern right whale: A quantitative analysis. *Animal Behaviour*, *30*, 1060-1071.

Clark, C. W., & Ellison, W. T (2004). Potential use of low frequency sounds by baleen whales for probing the environment: Evidence from models and empirical measurements. In J. A. Thomas, C. T. Moss, & M. Vater (Eds.), *Echolocation in bats and dolphins* (pp. 564-581). Chicago: Chicago University Press. 605 pp.

Dusenbery, D. B. (1992). *Sensory ecology: How organisms acquire and respond to information.* New York: W. H. Freeman and Company. 558 pp.

Guinee, L. N., & Payne, K. B. (1988). Rhyme-like repetitions in songs of humpback whales. *Ethology*, *79*, 295-306.

Guinee, L. N., Chu, K., & Dorsey, E. M. (1983). Changes over time in the song of known individual whales. In R. Payne (Ed.), *Communication and behavior of whales* (pp. 59-80). Boulder, CO: Westview Press. 643 pp.

Helweg, D. A., Frankel, A. S., Mobley, J. R., & Herman, L. M. (1992). Humpback whale song: Our current understanding. In J. A. Thomas, R. A. Kastelein, & A. Ya. Supin (Eds.), *Marine mammal sensory systems* (pp. 459-483). New York: Plenum. 788 pp.

Helweg, D. A., Cato, D. H., Jenkins, P. F., Garrigue, C., & McCauley, R. D. (1998). Geographic variation in South Pacific humpback whale songs. *Behaviour*, *135*, 1-27.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, *78*, 1464-1480.

Mednis, A. (1991). An acoustic analysis of the 1988 song of the humpback whale, *Megaptera novaeangliae*, off Eastern Australia. *Memoirs of the Queensland Museum*, *30*, 323-332.

Mercado, E., III, & Frazer, L. N. (1999). Environmental constraints on sound transmission by humpback whales. *Journal of the Acoustical Society of America*, *106*, 3004-3016.

Mercado, E., III, & Frazer, L. N. (2001). Humpback whale song or humpback whale sonar? A reply to Au et al. *IEEE Journal of Oceanic Engineering*, *26*, 406-415.

Mercado, E., III, & Kuh, A. (1998). Classification of humpback whale vocalizations using a self-organizing neural network. *IEEE World Congress on Computational Intelligence*, *2*, 1584-1589.

Mercado, E., III, Herman, L. M., & Pack, A. A. (2003). Stereotypical sound patterns in humpback whale songs: Usage and function. *Aquatic Mammals*, *29*(1), 37-52.

Mercado, E., III, Herman, L. M., & Pack, A. A. (2005). Song copying by humpback whales: Themes and variations. *Animal Cognition*, *8*, 93-102.

Norris, T. F. (1995). *Effects of boat noise on singing behavior of humpback whales (*Megaptera novaeangliae*).* Unpublished Master's thesis, San Jose State University, San Jose, CA.

Payne, K., & Payne, R. (1985). Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, *68*, 89-114.

Payne, K., Tyack, P., & Payne, R. (1983). Progressive changes in the songs of humpback whales (*Megaptera novaeangliae*): A detailed analysis of two seasons in Hawaii. In R. Payne (Ed.), *Communication and behavior of whales* (pp. 9-57). Boulder, CO: Westview Press. 643 pp.

Payne, R., & Guinee, L. (1983). Humpback whale (*Megaptera novaeangliae*) songs as an indicator of "stocks." In R. Payne (Ed.), *Communication and behavior of whales* (pp. 333-358). Boulder, CO: Westview Press. 643 pp.

Payne, R., & McVay, S. (1971). Songs of humpback whales: Humpbacks emit sounds in long, predictable patterns ranging over frequencies audible to humans. *Science*, *173*(3997), 585-597.

Plantinga, J., & Trainor, L. J. (2003). Long term memory for pitch in six-month old infants. *Annals of the New York Academy of Science*, *999*, 520-521.

Potter, J., Mellinger, D., & Clark, C. (1994). Marine mammal call discrimination using artificial neural networks. *Journal of the Acoustical Society of America*, *94*(3), 1255-1262.

Ralston, J. V., & Herman, L. M. (1995). Perception and generalization of frequency contours by a bottlenose dolphin (*Tursiops truncatus*). *Journal of Comparative Psychology*, *109*, 268-277.

Richards, D. G., Wolz, J. P., & Herman, L. M. (1984). Vocal mimicry of computer generated sounds and vocal labeling of objects by a bottlenose dolphin, *Tursiops truncatus*. *Comparative Psychology*, *98*, 10-28.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning past tenses of English verbs. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing. Vol. 2: Psychological and biological models* (pp. 216-271). Cambridge: MIT Press. 632 pp.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606-621.

Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Information entropy of humpback whale songs. *Journal of the Acoustical Society of America*, *119*(3), 1849-1866.

Walker, A., Fisher, R. B., & Mitsakakis, N. (1996). *Singing maps: Classification of whalesong units using a self-organizing feature mapping algorithm (*DAI Research Paper No. 833).

Winn, H. E., & Winn, L. K. (1978). The song of the humpback whale *Megaptera novaeangliae* in the West Indies. *Marine Biology*, *47*, 97-114.