International Journal of Comparative Psychology

Title

Using a Self-Organizing Map (SOM) and the Hyperspace Analog to Language (HAL) Model to Identify Patterns of Syntax and Structure in the Songs of Humpback Whales

Permalink

https://escholarship.org/uc/item/0vc7j5g2

Journal

International Journal of Comparative Psychology, 25(3)

ISSN

0889-3675

Authors

Kaufman, Allison B.
Green, Sean R.
Seitz, Aaron R.
et al.

Publication Date

2012

DOI

10.46867/ijcp.2012.25.03.01

International Journal of Comparative Psychology, 2012, 25, 237-275.

Using a Self-Organizing Map (SOM) and the Hyperspace Analog to Language (HAL) Model to Identify Patterns of Syntax and Structure in the Songs of Humpback Whales

Allison B. Kaufman
California State University, San Bernardino, U.S.A.

Sean R. Green
University of Buffalo, The State University of New York, U.S.A.

Aaron R. Seitz and Curt Burgess
University of California, Riverside, U.S.A.

Two different fully automated models were used to examine syntax and structure in humpback whale song. Songs were initially classified via a Self-Organizing Map (SOM), and then examined, via the Hyperspace Analog to Language (HAL) model, for evidence of a type of higher level organization - global co-occurrence - found in human language. HAL was able to identify particular “classes” of song units which were used interchangeably to form patterns in the song, not unlike the use of *noun-verb-direct object* in human language, where the noun, verb, or direct object can be any one of many possibilities from that particular class. Further, HAL identified specific patterns unique to the songs and their respective geographical areas. These patterns provide support for the idea that humpback whale songs are unique to specific region and may be transmitted culturally.

In 1971, researchers Payne and McVay demonstrated, for the first time, that the sounds made by humpback whales while “singing” were organized and sequential. Subsequent attempts to characterize song showed syllables, which could be combined to motifs, motifs which could be combined to create phrases, and phrases which were joined to themes yielding a distinct hierarchical structure (e.g., Winn & Winn, 1978). When phrases and motifs were compared over the years, additions or subtractions of syllables, or syllable types, were evident (e.g., Payne & Payne, 1985). Later studies revealed differences between songs recorded in the North Atlantic, North Pacific, and South Pacific (Winn et al., 1981).

The debate over the function, if any, of the structure and repeated elements in whale song is reflected in broader disagreements over the function of whale song itself (Darling, Jones, & Nicklin, 2006; Helweg, Frankel, Mobley, & Herman, 1992; Herman & Tavolga, 1980). General consensus was, and remains, that songs are used primarily for courtship purposes, as they are sung almost exclusively by males in breeding season. A fairly recent study showed that 58% of migrating male humpback whale singers were found associating with conspecifics; and singers

The authors would like to thank the Animal Behavior Archive maintained by the Macaulay Library at Cornell University, Eduardo Mercado III, Khaleel Razak, and James Kaufman. Part of the work detailed here was done to fulfill the requirements of a dissertation by ABK in the department of Neuroscience at the University of California, Riverside. During much of the time while data was being analyzed, ABK was supported by a Dissertation Year grant from the University of California, Riverside. Correspondence concerning this article should be addressed to Allison Kaufman, Department of Psychology, California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA 92407, U.S.A. (akaufman@csusb.edu).

were more likely both to join with and stay longer with groups of whales containing mother-calf pairs and no other males (although note that this particular study was carried out on migratory whales, as opposed to whales on breeding grounds; Smith, Goldizen, Dunlop, & Noad, 2008). However, other hypotheses do exist: (a) that the song is part of a search mechanism, possibly even long-range sonar (Frazer & Mercado, 2000; *cf.* Au, Frankel, Helweg, & Cato, 2001); (b) that song plays a part in sexual selection, for example, the ability to sing a lengthy song without breathing could be a sign of greater fitness (Chu, 1988); (c) that song creates an organizational structure, possibly even one that occurs cooperatively among males in mating season (Darling et al., 2006); or (d) that song is indicative of a whale's position in a dominance hierarchy (Darling, 1983). Parsons, Wright, and Gore (2008) provide an excellent review of these and other hypotheses.

Currently, many studies liken whale song to the song of territorial songbirds, and in this tradition song elements are often labeled in isolation based on inspection of spectrograms or identification of acoustically invariant cues (e.g., the "fee-bee" song of the chickadee; Weisman & Ratcliffe, 2004). For many bird species, stereotypy of song features, song matching and song copying play important roles in territorial conflict. To the extent that whale song is like bird song, fidelity in reproducing learned songs and the flexibility in adopting new songs across years and breeding seasons is likely to be important, while the actual content of song may be highly redundant and/or arbitrary (Eriksen, Miller, Tougaard, & Helweg, 2005; Newman, Yeh, & Price, 2008; Rendell & Whitehead, 2001; Riesch & Deecke, 2011; Scarl & Bradbury, 2009; Sockman, Salvante, Racke, Campbell, & Whitman, 2009; Winn et al., 1981; Yurk, Barrett-Lennard, Ford, & Matkin, 2002).

Other studies of whale song have focused on the information content of whale song rather than its suitability for sexual advertisement (e.g., Suzuki, Buck, & Tyack, 2006). The ability of a song to carry information is limited if a song is either highly repetitive or nearly random (McCowan, Doyle, Jenkins, & Hanser, 2005; Suzuki, Buck, & Tyack, 2005). Thus, the song qualities that may be useful for unambiguously declaring a small amount of information (e.g., stereotypy) would be less well-suited for conveying a complex message. Studies of the information content of whale song have tended to focus on the internal structure of songs rather than the changes from year to year or region to region, although the sudden change in song for one population as a result of contact with another population has been interpreted as evidence for "revolution" in humpback whale song (e.g., Noad, Cato, Bryden, Jenner, & Jenner, 2000).

Clues to the function or role of humpback whale song may be found in the changes to the song across years and across geographical areas. Typically, these changes are described as progressive (Payne & Payne, 1985), although songs may replace one another in "revolution" rather than "evolution" in the course of a number of years (Noad et al., 2000). The changeable nature of whale song, coupled with the spread from population to population, places constraints on how whales learn new songs. As Rendell and Whitehead (2001) point out, this pattern of song change is inconsistent with parent-to-offspring transmission of culturally distinct whale song. However, it seems that male humpbacks must learn at least some aspects of their song from each other, and that the benefit of learning new songs outweighs the cost in terms of time and cognitive resources required to do so. Therefore, it seems reasonable to ask: what has changed from year to year or from place to place in order to make the new song an improvement? Does singing a new song help the whale fit in with its neighbors, stand out from its neighbors, or allow its song to be heard more effectively? In order to investigate these questions, it is necessary to examine whale song in terms of its cultural, social, environmental and syntactic context.

To investigate the possibility that humpback whale vocalizations are context dependent, we turn to tools that were developed to investigate an undeniably context dependent communication system – human language. Despite recent studies to the contrary (e.g., Crockford & Boesch, 2005; Ouattara, Lemasson, & Zuberbühler, 2009; Zuberbühler, Cheney, & Seyfarth, 1999), human language and animal communication have been traditionally seen as categorically different (Christiansen & Kirby, 2003; Hauser, Newport, & Aslin, 2001; Jackendoff & Pinker, 2005; Nowak & Komarova, 2001; Pinker & Jackendoff, 2005; Premack, 2004; Weissengruber, Forstenpointner, Peters, Kubber-Heiss, & Fitch, 2002). However, maintaining this distinction in a methodological approach may be detrimental – human language research is significantly more advanced than animal communication research, and if, from a computational perspective, animal communication streams have statistical regularities, then the human models can be applied to animal research. In human language, the presence of a particular word in an utterance or text is governed by a multitude of local and long-range contextual influences. For a concrete example, occurrences of the word "whale" in this paper are not constrained simply by the preceding few words, but by a combination of constraints ranging from the syntactic structure of the sentence to the gist and topic of the entire paragraph or text.

Recently, computer modeling has come to the forefront of research on the communication of non-human animals, with efforts to quantitatively detect, segment, and classify vocalizations in species including bottlenose dolphins (Buck & Tyack, 1993; Janik, 1999; McCowan, 1995), killer whales (*Orcinus orca*; Brown & Hodgins-Davis, 2006; Deecke, Ford, & Spong, 1999), humpback whales (*Megaptera novaeangliae*; Suzuki et al., 2006), bats (*Microchiroptera* spp.; Skowronski & Harris, 2006), prairie dogs (*Cynomys gunnisoni;* Placer, Slobodchikoff, Burns, Placer, & Middleton, 2006), beaked whales (*Mesoplodon spp.*; Mellinger, 2008), and songbirds (Gentner, Fenn, Margoliash, & Nussbaum, 2006). Continuing in a similar vein, the Hyperspace Analog to Language (HAL) is an automated technique for identifying syntactic patterns within any structured sequence. Work with HAL has demonstrated that a broad range of human language capacity can be accounted for by simple inductive learning. As a result, models such as these are ideal for use in animal communication as they "bootstrap" the "rules" of language/communication from the already existing language stream. In other words, these models can take a given language stream and extrapolate the rules used to create it. It is our aim to apply HAL to whale song in order to determine whether humpback whale song – like human language – is shaped by statistical regularities beyond the simple repetition of phrases and motifs. One type of statistical regularity that HAL could look for is the patterns that distinguish nearby groups of whales from one another. Previous study of whales in the area and time period used for this study (the Caribbean in the early to mid-1970s) suggested that while whales in the Caribbean all sang similar song, song content changed during that time period and there were comparatively subtle differences between songs recorded from different locations within the same year (Winn & Winn, 1978; Winn et al., 1981). If HAL can find differences between samples of whale song recorded at different times and locations, future studies might make use of HAL to better understand the syntactic and structural changes in whale song as it evolves.

In principle, data from large-scale studies of humpback whales using HAL could inform current theories of whale song function. Because HAL classifies song elements based on their context, it may be sensitive to changes in song characteristics that arise from many contextual factors. These factors include the arrival of novel song from nearby populations or clans, changes in the acoustic or ecological environment, or changes to the social context (e.g., the presence of nearby mates or allies). Note that this is a novel use for the HAL technique, as it has until recently only been applied to human language (Kaufman, 2010; Kaufman, Colbert-White, & Burgess, under review).

The promise of such a view is that language can be seen as starting with a set of basic cognitive processes rather than a complex linguistic system, and a cognitive or behavioral basis is a far more parsimonious explanation for the existence of regularities and rules in animal communications. As evidence for the conclusion that a human brain is not required for language-related cognition, the ability to use syntax in order to interpret commands has been documented in bonobos (Savage-Rumbaugh, 1993; Williams, Brakke, & Savage-Rumbaugh, 1997), sea lions (Schusterman & Gisiner, 1988) and bottlenose dolphins (Herman, Kuczaj, & Holder, 1993; Herman, Richards, & Wolz, 1984). The latter finding is particularly relevant, since it demonstrates that a cetacean evolutionary history is compatible with the ability to perceive and interpret syntactic structure.

HAL uses word order in language to compute co-occurrence values between words in a particular body of text (a *corpus*). This provides the basis for contextual learning in the model. The use of a high dimensional model such as HAL is relatively novel to the animal communication literature; beyond other work with the HAL model (Kaufman, 2010; Kaufman et al., under review; McCowan, Doyle, Kaufman, Hanser, & Burgess, 2008), none of the models currently in use are based on a global co-occurrence theory. HAL is a different kind of memory and language model, which, by using a global co-occurrence algorithm, is able to identify the similarities of a word's contextual use across a large body of text. For example, the words *cat* and *dog* often appear in the same sentence in human language (*I have a cat and a dog*). This is a local co-occurrence. However, *cat* and *dog* can also show a global co-occurrence pattern (*I have a cat that I'm in charge of feeding. She has a dog that she's in charge of feeding*). In global co-occurrence, *cat* and *dog* occur in the same context – in both sentences they are "the things that can be had and must be fed". This also makes them essentially (grammatically) substitutable. A traditional, local co-occurrence model would still identify *cat* and *dog* as being related, but would do so based only on the sentence *I have a cat and a dog.* In a more distinct example, the words *street* and *road* rarely occur in the same sentence, but are used in virtually the same contexts (*Make a left on that street. Make a left on that road*). The contextual similarity of the words *street* and *road* is reflective of their global co-occurrence, which would not be produced by a simple local co-occurrence procedure. While much valuable information can be obtained from both local co-occurrence and conditional probabilities, they do not capture higher-order contextual relationships because they do not encode words in this broader, more diverse, contextual fashion (Burgess, 1998; Lund & Burgess, 1996).

Method

Background

A variety of techniques rooted in information theory (Shannon, 1948) and other mathematical concepts (such as entropies, neural networks, and statistical regularities) have also been shown to provide much insight into the analysis of communication streams. Experimental procedures as early as the 1970s attempted to understand how animals communicated via information theory (e.g., Menzel, 1973). The HAL model works because it is provided with a segmented stream of words, much like what infants hear during language development. Previous experiments have demonstrated that infants heavily rely on statistical co-occurrence information at the phonetic level in learning word segmentation (Aslin, Saffran, & Newport, 1999; Estes, Evans, Alibali, & Saffran, 2007; Saffran et al., 2008), and a segmented language stream is the required input for the HAL model. The previously discussed hierarchical nature of humpback whale song is indicative of a similar type of organization, in which acoustic units are distinct from each other. Songs are therefore parsed into segments for analysis as described below. After segmentation, the units must be classified; the most accurate method for doing this has been the topic of much research and debate (Clemins, Johnson, Leong, & Savage, 2005; Janik, 1999; Melendez, Jones, & Feng, 2006; Murray, Mercado, & Roitblat, 1998; Rickwood & Taylor, 2008; van der Schaar, Delory, Catala, & Andre, 2007). Currently, the favored techniques are either segmentation by humans or segmentation by some variation on a self-learning neural network such as a Self Organizing Map (SOM), which is the technique used here and detailed below. The use of the SOM ensured appropriate input to the HAL model.

Throughout the description of the methods in the upcoming sections, it may be useful to the reader to remember the path of transformation taken by the data → Raw data → input vector → SOM → HAL → Color coded clusters → Color coded sequences

The Corpus

The corpus used in this experiment resulted from vocalization data obtained from the Macaulay Library (Cornell Laboratory of Ornithology) and classified by a Self Organizing Map (SOM). Original data were obtained via DVD data disks and included 325 distinct recordings made by 17 scientists. The recording dates ranged from 1964 to 2006 and included locations in both Atlantic and Pacific Oceans. Twelve songs from three locations were originally selected for analysis (three recordists, years ranging from 1970-1976), although only 11 songs were used in the final analysis (reasons for this are discussed below). The internal sequence order of each song was maintained (with some exception, discussed below).

Twenty one recordings of whale song were used for this study, comprising 13.4 hours of song. These songs were recorded between 1970 and 1976 at the following locations in the Caribbean: the Lesser Antilles, Puerto Rico, the Turks and Caicos Island. These recordings were obtained through the Animal Behavior Archive maintained by the Macaulay Library at Cornell University. See Table 1 for further information about these songs. In total, these songs have 9545 song units. These songs were sampled at a rate of 44,100 Hz. Two song recordings, #117762 (822 units) and #128296 (291 units) were not used in all analyses. The rationale for omitting these songs and the analyses for which they were omitted will be described in analysis and results sections.

Table 1

Description of song samples used in analysis

Song ID	Year	Location	Length (units)	Length (seconds)	Recorded by
110847	1973	Lesser Antilles	920	2875	Perkins, P.J.
110858	1973	Lesser Antilles	954	2954	Perkins, P. J.
117762	1970	Puerto Rico	822	2942	Perkins, P. J.
117770	1970	Puerto Rico	723	2681	Perkins, P. J.
117774	1970	Puerto Rico	762	2444	Perkins, P. J.
117775	1970	Puerto Rico	1271	3459	Perkins, P. J.
128296	1975	Puerto Rico	291	874	Llungblad, D. K.
118102	1974	Turks and Caicos	600	1887	Levenson, C.
118118	1974	Turks and Caicos	693	1902	Levenson, C.
118119	1974	Turks and Caicos	565	1902	Levenson, C.
118171	1976	Lesser Antilles	947	2971	Perkins, P. J.
118172	1976	Lesser Antilles	997	1972	Perkins, P. J.

SOM Methodology

In order to train an SOM to classify individual whale song units, it is first necessary to create a set of input vectors, i.e., groups of numerical values representing the song units. This process presents two challenges, isolating individual units and creating input vectors to represent these units. Segmentation was achieved here by using the following: (1) an automated estimate of each sound's start and stop points based on changes in sound amplitude, (2) visual inspection of the waveform and spectrogram, and (3) listening to the sound. For more detail on the methods, see Green, Mercado, Pack, and Herman (2007).

Once the individual song units were obtained, the following elements of the input vector were directly computed from the sound. The timing of the unit was represented as its duration as well as the time interval between the song unit and its neighbors (three elements). The amplitude of the song unit was represented as the root mean square of the entire sound unit as well as the root mean square of each fifth of the song unit (six elements). The spectrum of the song unit was represented by dividing the spectrum into eight bands and calculating the proportion of energy contained within each band (8 elements). See Table 2 for a description of these input vectors. While the duration of the units themselves seemed to be consistent, the time between units seemed to be highly dependent on the units to follow and/or other units in the pattern. This is hypothesized to be one of the ways humpbacks may be able to recall and repeat such large portions of song (Handel, Todd, & Zoidis, 2009).

In order to represent the relationship between a sound and its immediate neighbors, "relative" acoustic measures of the song unit were calculated by comparing each element with the corresponding element of the song unit's nearest neighbors. For each of the elements described above except for "following interval" and "preceding interval," the logarithm of each element divided by its nearest neighbors was calculated (16 elements were calculated relative to the preceding unit and 16 elements were calculated relative to the following unit, for a total of 32 relative acoustic measures). Thus, in total, each song unit was represented as a vector with 53 elements. In order to prevent vector elements with large values (such as duration) from dominating the network at the expense of elements with smaller values (such as amplitude), the input vectors were normalized so that the range for all elements across the data set was between zero and one.

An SOM consisting of 25 nodes arranged in a 5x5 hexagonal grid was trained on each of the input vectors representing the data set. Training consisted of 200 epochs, and each input vector was presented to the network once during each epoch. When the SOM was presented with an input vector, the SOM program, running in MATLAB, used MATLAB's *linkdist* function to calculate the distance between the input vector and each of the nodes in 53-dimension space. The neighborhood size, learning rate and other network parameters were set to the MATLAB default values for the newsom function.

After training, the location of input vectors and nodes were compared and each vector was assigned a number based on the closest node (in terms of Euclidean distance). Euclidean distance was calculated by summing the squares of the distance on each of the 53 dimensions and taking the square root of the result. These numerical assignments were then used as category labels for the song units themselves. For instance, a song unit whose input vector most closely resembled Node 24 of the

SOM would be labeled a type 24 song unit. This classification system allows us to represent each whale song as a series of numbers. These sequences were in turn used as input for the HAL model.

Table 2

Numerical values calculated from whale song recordings for each whale song unit used to create the input vectors that were presented to the SOM

Element	Description
1, 2	Duration of gap preceding and following the song unit
3	duration
4	Amplitude (root mean square)
5-9	Amplitude of song unit quintiles as a proportion of total
10	Peak frequency
11-18	Frequency in eight bands as a proportion of total 1) 0- 100 Hz 2) 100-250 Hz 3) 250-500 Hz 4) 500-750 Hz 5) 750-1000 Hz 6) 1000-2000 Hz 7) 2000-3000 Hz 8(Over 3000 Hz
19-21	Quadratic, Linear and Intercept values of a quadratic curve fit of the peak frequency over time within the unit. Frequency values were calculated for each 100 ms of the unit.
22-37	Unit values for elements 3-18 relative to the preceding unit (Logarithm of current unit value divided by the preceding unit).
28-53	Unit values for elements 3-18 relative to the following unit (Logarithm of current unit value divided by the following unit).

HAL Methodology

Global co-occurrence values in the HAL model are computed by use of a sliding window (typically ten words long for human language corpuses), which assigns a co-occurrence value to each pair of words in the window based on the number of intervening words. A matrix is created by encoding these values (see Figure 1a for an example matrix). Co-occurrences between a particular word and those which precede it are encoded in rows, while those that follow it are encoded in the columns of the matrix.

By including co-occurrence values for words that occur both before and after a target word, the model provides a contextual perspective beyond that of conditional probabilities or measures of entropy, which only encode co-occurrences in one direction. Once a matrix is formed, each word can be represented by a vector comprised of its row co-occurrence values and its column co-occurrence values (see Figure 1B). These vectors can be visualized and grouped by using either multi-dimensional scaling or hierarchical cluster analysis. The placement of words in these visualizations is a function of the similarity of their contextually driven vectors (see Figure 2), and words that are used in similar contexts will cluster in the visual output.

In the analyses presented here, Ward's (1963) method of hierarchical cluster analysis was deemed most useful for visualization of clusters. Ward's analysis is unique from other methods of cluster analysis in that the clusters it creates are based on attempts to minimize the error sum of squares between the groups (Aldenderfer & Blashfield, 1984; Ward, 1963). This approach is particularly useful in visualization of a HAL analysis because the reduction to two dimensions naturally causes information loss, and, although the loss is much less than one might expect (Burgess, 1998; Lund & Burgess, 1996), Ward's method provides an ANOVA-like reliability for the clusters in high-dimensional space. One of the major disadvantages with using Ward's method is that it tends to give solutions with many small, distinct clusters (Aldenderfer & Blashfield, 1984). The analyses performed here were not troubled by this, which may provide additional support for the robustness of the data.

A) Example sentence: Studying animal vocalizations is fun.

	Studying	animal	vocalizations	is	fun
Studying	0	0	0	0	0
animal	4	0	0	0	0
vocalizations	3	4	0	0	0
is	2	3	4	0	0
fun	1	2	3	4	0

B) Vector for “animal” – 4 0 0 0 0 0 0 4 3 2

Figure 1. **A)** Example of HAL matrix of the sentence “Studying animal vocalizations is fun”. Window size 5, co-occurences between a particular word and those that precede it are encoded in rows, while those that follow it are encoded in columns. **B)** Example co-occurrence vector for the word “animal” in the above matrix. Row co-occurrence values are followed by column values.

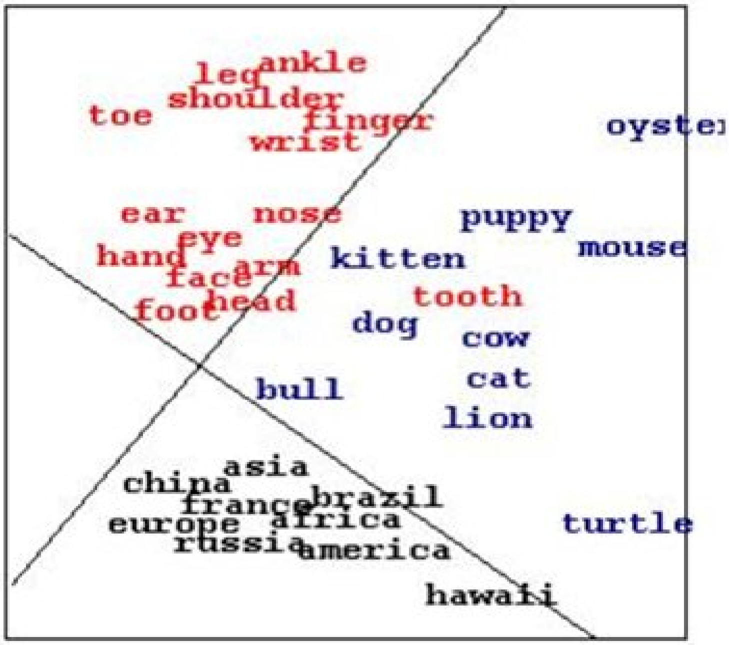


Figure 2. Placement of words in a visualization of HAL vectors. From Lund and Burgess (1986).

HAL Parameters. When creating a HAL matrix and vectors, there are several options or parameters which can be set. The three options relevant to the experiments here are “window size,” “cut,” and “limit.”

Window size controls the number of words on either side of the target word used in the calculation. It is a measure of the length of the sliding window. The default window size (and size which has been found to be most appropriate for use with human language) is ten. Smaller corpuses and/or corpuses in which less sophisticated co-occurrence is expected warrant smaller window sizes (Burgess, 1998; Lund & Burgess, 1996).

The cut function removes all words in a corpus that occur with a frequency at or below a certain level (for example, a cut at three removes all words that occur only once, twice, or three times in the corpus). This allows for the removal of words for which there are not enough occurrences to create fully correct representations. By varying where the cut is with the size of the corpus, it is possible to even out and subsequently compare corpuses by forcing HAL to focus on the top X% most frequent words, regardless if one corpus has 10,000 words and is cut at 30 and another has 1,000 words and is cut at two.

Limits serve to equalize frequency from the top down – in many corpuses (human or otherwise) a small set of words are a lot more frequent than others. The limit parameter tells HAL to “look” only at the first X number of instances of each word when calculating its co-occurrence, thereby controlling for a frequency effect (Burgess, 1998; Lund & Burgess, 1996).

As the corpus used in this study was large, a default window size of ten was deemed appropriate (Lund & Burgess, 1996). In this corpus, unit frequency ranged from as high as 500 to 600 occurrences of a particular unit to as few as 14 occurrences of a unit – although in most cases the minimum number of occurrences of any particular unit in a corpus was around 150. In cases such as this, it is often beneficial to limit the matrix at the mean or median frequency. The limit acts as a quota; as HAL moves through the corpus calculating co-occurrence values, it disregards any occurrences of a particular unit after that unit’s limit has been reached. The limit on all of the matrices used was 250, unless otherwise indicated. For more information on matrix building and the theory behind HAL parameters such as window size and limit, see Lund and Burgess (1996).

Mixing the Corpus. Because it was advantageous to limit the HAL matrix at 250, it was necessary to “mix” the corpus to avoid problems with songs that contained particularly high numbers of any specific unit. For example, if the first song in the corpus contained 250 occurrences of a particular unit, HAL would reach the limit for that unit almost immediately, and, thus, would ignore any occurrences of that unit in any other song for the remainder of the corpus. In order to avoid this problem, and to balance the first and second halves of the songs, songs were “mixed” as follows (parts A and B designate the first and second halves of each song):

Split 1
Song1, partA
Song2, partB
Song3, partA
Song4, partB
etc…

Split2
Song1, partB
Song2, partA
Song3, partB
Song4, partA
etc…

The two splits were then concatenated to create a final corpus to be used in the experimental procedure. A template was created and used each time a corpus needed to be mixed. Additionally, the splits created were used in the split-half reliability calculations discussed later.

Visualization. The mixed corpus was subjected to HAL analysis, as described previously. The result of this analysis was a 53-dimensional space containing twenty-five 53 element vector representations (25 being the number of distinct units identified by the SOM - the equivalent of words — and 53 being the number of elements in each vector; or the characteristics of the actual sound). Following the procedures discussed in the introduction, these vectors were visualized via Ward's Cluster Analysis and "mapped onto" the original corpus via color coding as described below.

Alternate Corpuses. When HAL analysis is used with human language, there is no need for confirmatory analysis. Because we know the language, it makes sense to us that "cat" and "dog" should be more likely to have similar contexts than "cat" and "file." This luxury is not afforded when working with a corpus comprised of animal vocalizations, and, therefore, two separate techniques were used to validate the results.

Three new corpuses were created by removing songs. Song numbers 117762 and 128296, henceforth referred to as song62 and song96, were removed, as they were more irregular due to inordinately high proportions of a particular unit or group of units (see discussion in Miksis-Olds, Buck, Noad, Cato, & Stokes, 2008). Song96 was also much shorter than any other song. These removals created the corpuses "without117762" (or wo62), "without128296" (or wo96), and "without117762and128296" (wo62and96), in addition to the "full" corpus. A comparison of all four corpuses was used as a test for robustness of the procedure, in addition to being used in the final corpus selection process detailed below.

Song62 and song96 were removed in the acoustic stage, prior to unit classification by the SOM. Because of this procedure, however, there was an inherent problem with comparing the four corpuses. When classification is done by different SOMs, the identification of each acoustic unit (i.e., the unit's "name") is different. For example, a particular acoustic unit could have been called unit1 by the full SOM, unit16 by the wo62 SOM, unit22 by the wo96 SOM, and unit25 by the wo62and96 SOM.

Color Coding and "Mapping" Clusters onto the Sequential Corpus. To solve this problem, a color coding scheme was used as a way of marking clusters when they were output from HAL. Once put into effect, the usefulness of this color coding actually extended far beyond the confirmatory analysis, and it became a vital part of the methodology. Once HAL analysis of a corpus was completed and the vectors were visualized using cluster analysis, each of the clusters was assigned a color. For example, in Figure 3, a cluster analysis of the full corpus, the cluster containing unit4, unit5, and unit9 is assigned the blue color. Consequently, in the sequential listing of the units which made up the full corpus, all of the units four, five, and nine were highlighted blue. In addition, because the temporal order of the corpuses was maintained, the acoustic units were still ordered the same, regardless of what they were called. In this way it was possible to place the corpuses next to each other in a document, creating rows consisting of the same acoustic unit *across* the corpuses, regardless of what it was called. Once all the clusters had been assigned colors and the colors transferred to the corpuses (the "mapping on" of the clusters), a consistency of color across an entire row would be indicative of an equivalent cluster across the different SOMs (see Figure 4). This procedure could be used for comparison across corpuses by individual song as well as along the complete corpus, and was used throughout experimentation.

Split-Half Analysis. A split-half confirmation of the clusters produced by the HAL model was also performed. Each of the songs was split into an A section and a B section, and two "splits" were created from alternating A's and B's, as discussed before. However, the two splits were now treated as two separate corpuses, instead of being concatenated to one corpus. Because there is no established statistical method to compare categorical data, Cohen's kappa (1960) was used as an experimental approximation. Cohen's kappa is an inter-rater reliability statistic used to compare agreement between two raters who code behavioral observations. In this case, the two versions of cluster membership (the full and one of the split halves) were each considered to be a "rater" and the cluster membership for each unit was considered the rating for each particular behavior. As a result, the model was created such that two "raters" (the two versions of cluster membership) were being compared on 25 "instances of behavior" (the 25 units), for which they had assigned one of eight "types of behavior" (the clusters).

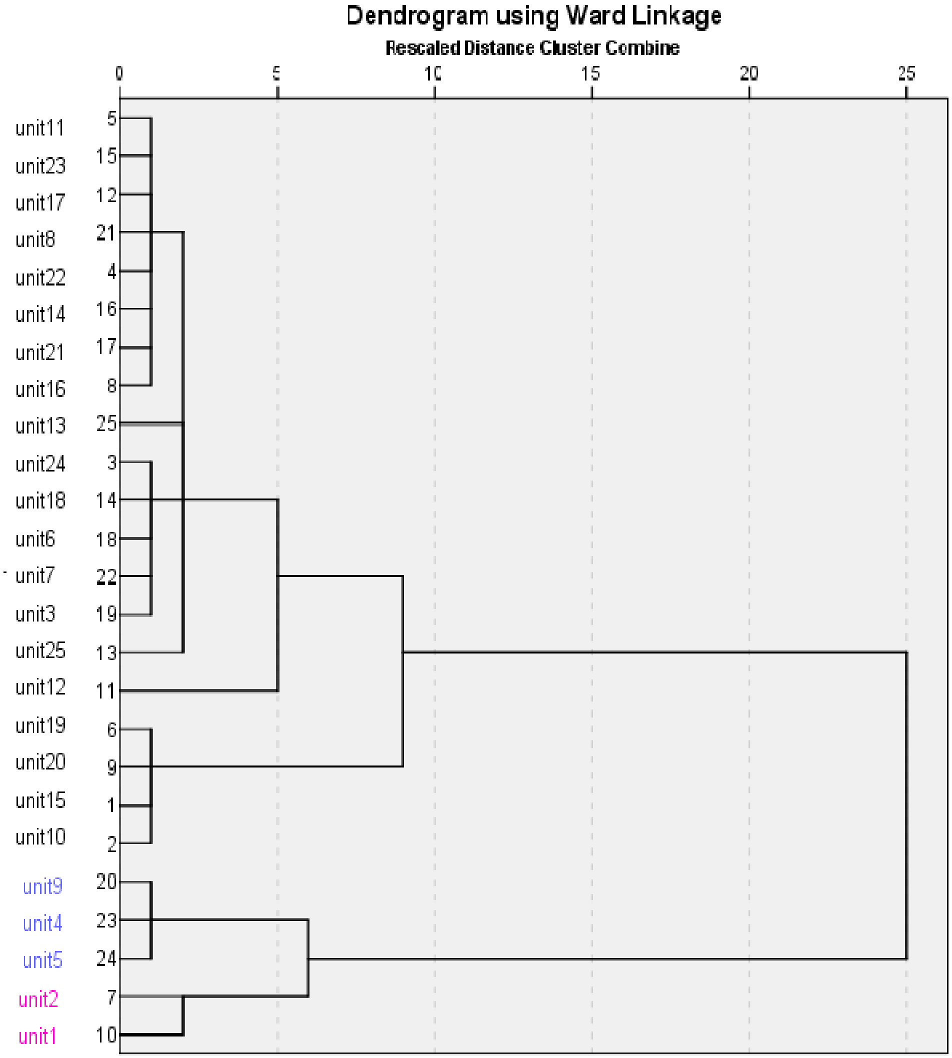
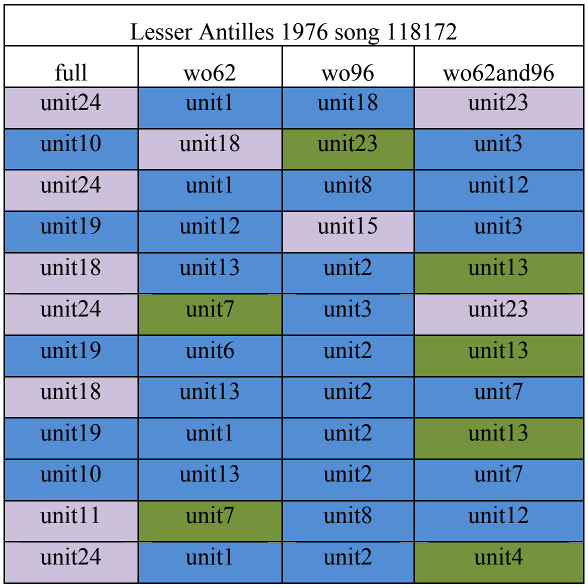


Figure 3. Wards cluster analysis of HAL vectors from the "full" corpus. Song units cluster based on global co-occurrence patterns, and clusters can then be given color codes.



Lesser Antilles 1976 song 118172			
full	wo62	wo96	wo62and96
unit24	unit1	unit18	unit23
unit10	unit18	unit23	unit3
unit24	unit1	unit8	unit12
unit19	unit12	unit15	unit3
unit18	unit13	unit2	unit13
unit24	unit7	unit3	unit23
unit19	unit6	unit2	unit13
unit18	unit13	unit2	unit7
unit19	unit1	unit2	unit13
unit10	unit13	unit2	unit7
unit11	unit7	unit8	unit12
unit24	unit1	unit2	unit4

Figure 4. Mapping of the color coding scheme onto the corpus. In this example, the same song is color coded with clusters obtained from each of the four corpuses (i.e., the "full" corpus grouped unit10 and unit19, while the "wo62" corpus grouped unit1, unit6, unit12, and unit13). As the sequence is maintained in all cases, each cell in a particular row is the same acoustic unit, regardless of the unit label assigned by the SOM. HAL, grouping the units by sequential co-occurrence, should maintain the same order with regard to cluster membership, and this can be read across rows as matching color patterns. Note this is only a small portion of a song.

Final Analysis

Two types of analysis were conducted, one that examined the frequency of occurrence of units and classes, and one that examined the overall entropy of the songs. Frequency of occurrence of each class was examined as a measure of change in song composition and perhaps dialect, whereas entropy analysis was performed to establish the existence of non-random sequences of information, and then to identify and compare these sequences across songs.

Frequency Analysis. Although it is unknown if frequency of occurrence of unit or class plays a role in humpback whale communication, there is some precedent for it being studied as a variable. Suzuki et al. (2006) used frequency of occurrence to estimate the true distribution of units in humpback song, thus generalizing over a particular area. In killer whales, frequencies of occurrence of whistles within the sequence and transition patterns between whistles in sequences are nonrandom (Riesch, Ford, & Thomsen, 2008). Frequency also plays a large part in calculations of conditional probabilities in other studies of humpback whale song (Green et al., 2007).

Entropy Analysis. To examine the extent to which the sequences of sequences were ordered, we conducted a Monte Carlo analysis where the order of units were compared to randomly generated sequences that preserved the frequency of each unit type. This analysis was performed on each of the songs individually, using the clusters produced by inputting the entire corpus into HAL. This was done initially to identify empirically if any of the songs were in fact anomalous (song62 and song96 were suspected to be, but were removed for a test of robustness; they could not be removed entirely without an empirical justification). Furthermore it provided a measure of the entropy of each sequence.

In this analysis, the frequency of each unit for each song was computed. These frequencies values were used as probabilities to generate 5000 random "songs" of the same length of the original song. Then for the true song and each of the random songs, sequences between two and ten items were segmented and the number of repeats of each sequence in each song was calculated. We could then compare the number of times that each sequence appeared by chance to the number of times that a given sequence occurred in the true song to determine the likelihood that a given sequence occurred by chance. Upon identifying the sequences which existed within each song (via the entropy analysis algorithm), further data analysis was based on sequences which occurred five or more times in a particular song (henceforth referred to as "the sequences found"). Comparison of these data was largely based on qualitative techniques which allow for the sorting and characterization of data (Bogdan & Biklen, 1998).

The Elman Corpus

Because one of the goals of this study was finding semantic patterns within humpback whale song, it seemed relevant to perform a similar procedure on human language. The corpus chosen was one created by Elman (1990) when developing his Simple Recurrent Network (Elman, 1990, 1993, 1995). This corpus was chosen because it has been subject to HAL analysis previously and, despite the substantial difference between the mechanisms of the two models, yielded similar results (Burgess & Lund, 2000). In addition, the corpus consists of approximately the same number of words in it as units in the humpback whale song. The "Elman Corpus" was subjected to the same HAL analysis as the humpback whale song, although many of the elaborate confirmatory procedures did not have to be undertaken since it was a known language. HAL vectors were subject to cluster analysis, and clusters were color coded and mapped on to the original corpus, just as had been done in the case of the whale song (see Figure 5).

The color coding of the Elman corpus provided a visual example of the patterns we sought in the humpback whale corpus. As shown in Figure 5 and introduced in the previous section, it is possible to assign colors to each HAL cluster and then, upon returning to the sequential text, demonstrate "banding" patterns which are a visual representation of sentence structure. In Figure 6A, nouns are purple, verbs are blue, and direct objects are pink. Therefore, a typical sentence, regardless of which actual words were used, should be visually represented as *purple-blue-pink,* or *noun-verb-direct object.* Figure 6 shows two banding patterns on the same sequence – 6A being the "correct" pattern, i.e., colors assigned by the researcher, who knew which words were nouns, verbs, and direct objects, and 6B being the groupings provided by HAL, which successfully created similar banding patterns. These banding patterns were what we were looking to identify in the whale song results, and are discussed in more detail below.

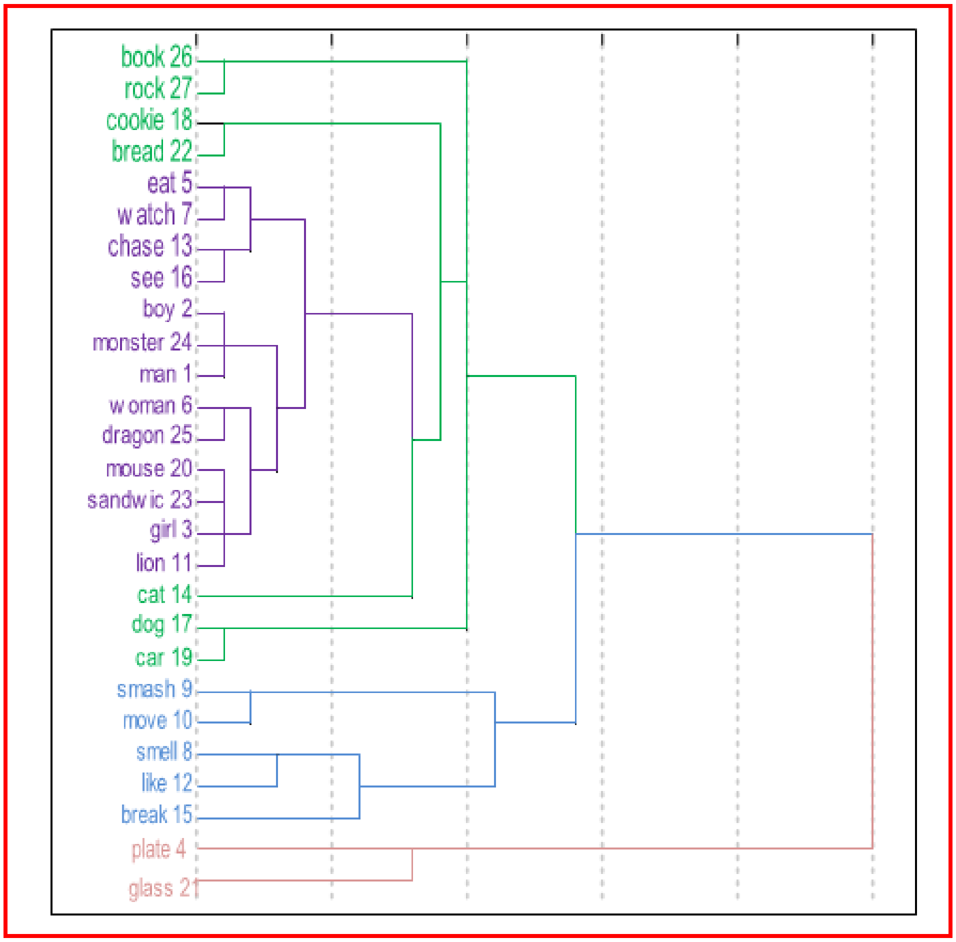
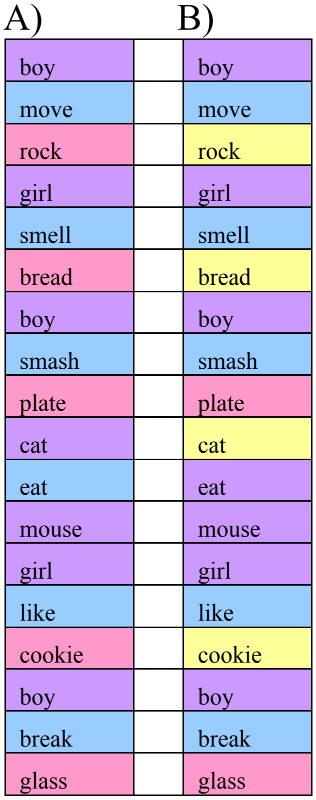


Figure 5. Color coded cluster analysis of the "Elman corpus." Clusters represent words which have high co-occurrence values (i.e., are often used in the same context)



A)		B)
boy		boy
move		move
rock		rock
girl		girl
smell		smell
bread		bread
boy		boy
smash		smash
plate		plate
cat		cat
eat		eat
mouse		mouse
girl		girl
like		like
cookie		cookie
boy		boy
break		break
glass		glass

Figure 6. Elman corpus color coding. **A)** The "correct" answers, manually coded. Animate subjects in purple, verbs in blue, and inanimate direct objects in pink. **B)** The clusters created by HAL. Note the similarity; HAL created four groups instead of three, primarily via an extra division within the inanimate direct objects.

Results

Final Corpus Selection

A comparison of all possible sequences of lengths two to ten to the number of actual sequences of lengths two to ten was used to measure entropy. Low entropy (high organization) songs were characterized by having a small percent of the possible sequences present. The possible sequences were calculated, for comparison purposes, by randomizing the sequence order, keeping frequency consistent. The entropy analysis was run on the entirety of each of the four corpuses (full, wo62, wo96, wo62and96), plus each individual song in each of the following conditions: full, wo62split, wo62lumped, wo96split, wo96lumped, wo62and96split, wo62and96lumped. Split and lumped versions referred to alternate readings of the Ward's Cluster Analysis, which does not always provide an exact solution.

Graphs of the entropies at each sequence length for songs in the full corpus condition were compared (see Figure 7). This comparison showed that song62 was clearly anomalous, having a very different structure than any of the other songs, namely entropy that increased with sequences of length 3-4 and then decreased at lengths of 6-7. This is the only song whose entropy is not highest (most random) at sequences of two units. Logically speaking, one would expect to see the highest entropy at the 2-unit sequence, as sequences of two units are very easy to "make," and therefore a higher proportion of the ones possible (represented by the random condition) would appear.

It was therefore determined that the wo62 corpus would be used for all further analysis. The split version of the corpus was used because it generally yielded lower entropies. The clustering pattern produced by this corpus yielded six classes, which were assumed to be grammatical and/or semantic classes and used to characterize patterns in the songs.

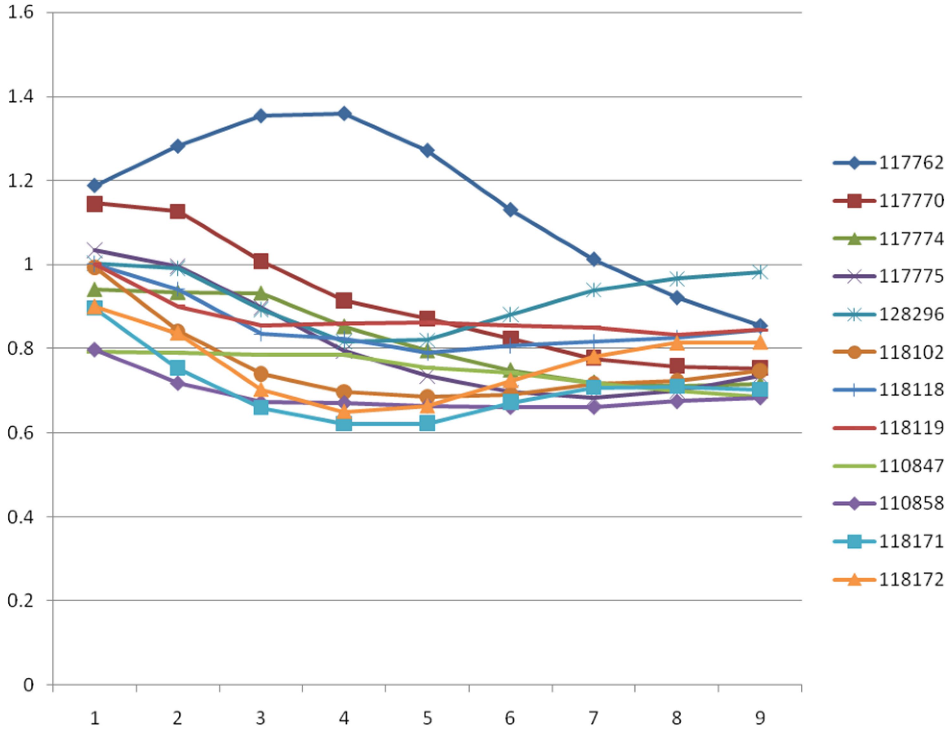


Figure 7. Graph of entropy values for sequences of lengths 2 to 8 for each of 12 humpback whale songs. Note that Song 117762 is anomalous as compared to the other songs.

Confirmatory Analysis

After the wo62 corpus was chosen, a split-half procedure using the "mixed" halves referred to above was completed in order to confirm reliability of the overall technique. Using Cohen's kappa (1960), each of the split halves was compared to the full corpus, and kappas of 0.4 and 0.5, respectively, were obtained. Comparison of the wo62 corpus to randomized versions of split1 and split2 yielded trivial kappa values of 0.01 and 0.1, respectively. This final comparison to a randomization of each of the splits was done because there is no precedent for the use of Cohen's kappa in a situation like this. In clinical psychology, kappas of this magnitude (0.4 to 0.5) would indicate "moderate" agreement (Cicchetti & Sparrow, 1981).

Grammatical Classes Cause Patterns

The human language system is composed of grammatical classes from which we are able to choose words to piece together sentences. For example, to create a basic sentence structured *subject+ verb+ direct object,* we are able to choose from a large vocabulary of subjects, verbs, and direct objects; many of which are interchangeable. Consider, for example:

Boy sees dog
Boy sees cat
Boy sees mouse
Boy walks dog
Girl sees dog

and so forth. Imagine you did not know what a dog, cat, or a mouse was. Given one sentence from the above set, you would have very little information. Given the first three, you would have more. Dog, cat, and mouse are *contextually* similar in that they are all things a boy can see, in addition to belonging to the class that is direct objects. Given all five sentences, you could then surmise that the act of seeing is not exclusive to the male gender, as the girl can also see the dog (and, therefore, she can most likely see the cat and mouse as well). In addition, you would learn that not only can the boy see the dog, he can also walk it – and therefore most likely the girl can as well. These inferences are all made possible by contextual similarities within the sentences.

To demonstrate the technique used in this study in more "comfortable" format, the Elman corpus discussed above was used. After classes were color coded and mapped onto the original corpus, it is possible to compare HAL's accuracy to the correct grammatical categories – a luxury we do not have when dealing with animal vocalizations. HAL does fairly well; there is a noun cluster (purple) and a verb cluster (blue). HAL splits direct objects into two categories – pink and tan, with the pink cluster representing breakables (Figure 6). The categorization is fairly comparable to the "correct" categorization of the three classes directly to the left (which, in and of itself, is not always exact – for example, the pattern is briefly interrupted at "cat eat mouse" because "mouse" has received the overall classification of noun, not direct object).

However, what if we did not know the words? In Figure 8 the words are removed and we just see patterns of color. These color patterns represent the patterns of grammatical classes that occur in our language.

There is similar patterning throughout the humpback whale corpus. Because these patterns are based on *classes* distinguished by the HAL model, they are not recognizable simply by examining the order of the units. Likewise, the colored patterns from the Elman corpus come from the grammatical classes (whether they were created by HAL or by the experimenter's hand), not from the words themselves.

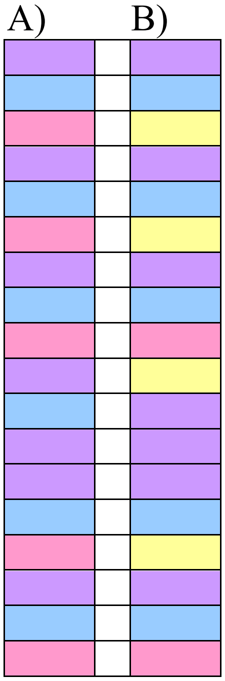


Figure 8. Removal of words from figure 6 to show patterns of grammatical classes. Note the patterned triplets (purple-blue-pink in column **A** and purple-blue-tan/purple-blue-pink in column **B**)

Frequency Measures

Initial analysis of the songs was accomplished by examining the frequency of occurrence of each of the classes. Because the number of songs available for comparison here is low, these results must be confirmed by additional analysis. In addition, the development of statistical techniques that could be used for a comparison across songs, despite the units being categorical data, is a necessity. Lastly, three classes – Class Green, Class Pink, and Class Yellow – consist of one unit, which makes generalization beyond mere presence or absence impossible to justify. However, there are some conclusions about class frequency which can be supported.

Turks and Caicos have the most distinctive repertoire. Class Orange is almost completely unique to this region, and Class Pink and Class Dark Purple never appear there at all.

The Lesser Antilles songs seem to differ with the year of recording (and are generally regarded as different groups throughout this experiment). This is particularly apparent in the use of specific units from Class Dark Blue; while the overall usage of Class Dark Blue remains roughly the same over time, in 1973 the Class Dark Blue units are nearly all units 20 and 21 (unit20 = 10.0%, unit21 = 12.3%, unit1 = 0%, unit12 = 1.1%). However, in 1976, this is completely reversed, and the composition of Class Dark Blue is almost entirely units one and

12 (unit20 = 3.5%, unit21 = 1.5%; unit1 = 10.4%, unit12 = 11.1%). If general meaning conveyed by Class Dark Blue remains the same, this finding may be support for the inherent substitutability of units within a particular class.

Lastly, the songs in Puerto Rico appear to be slightly more diverse than songs in other regions. In the other locations approximately 70 percent (TC = 72%, LA73 = 68%, LA76 = 70%) of all units can be accounted for by the two Classes that are also highest in frequency. However, in Puerto Rico, the top two frequency classes only account for 61% of the units. This decrease in uniformity, although seemingly slight, will appear in other analyses as well. In addition, the small sample size and our current lack of knowledge about how typical these song structures are in the Caribbean populations makes it difficult to draw inferences about whale song structure for entire groups or populations. To underscore this point, statistical tests were not performed on this data. Descriptive statistics are presented here in order to illustrate the results of a typical HAL analysis of whale song.

Sequence Analysis

Some generalizations regarding the sequences can be made across all songs. It was very common for a song to consist of one or two "backbone" Classes, from which units occurred many times in a row (bearing in mind that although the units themselves might be different, the class is the same, so this backbone would not appear in a conditional probability analysis, where the units would simply appear as their identified selves). Within this backbone, single "drop in" units occurred at intervals that were too large and too variable for a sequential analysis to handle.

Differences in the composition of the backbone were evident, as were distinct patterns of Classes within the songs. Songs had "roots," or sequences which appeared consistently. These roots were fairly consistent across all songs from a region, with the exception of songs from Puerto Rico. Using three unit "roots," (as there were consistently more of these than any other length), Table 3 shows the proportion of similar roots across songs.

Table 3

Number of three unit sequence "roots" similar across songs in a region.

			Number of sequences shared by		
Region	Number of Songs	Number of 3 unit sequences	4 songs	3 songs	2 songs
Puerto Rico	4	58	1	8	10
Turks and Caicos	3	20	NA	12	3
Lesser Antilles (All)	4	46	6	5	NA
Lesser Antilles 1973	2	24	NA	NA	19
Lesser Antilles 1976	2	33	NA	NA	17

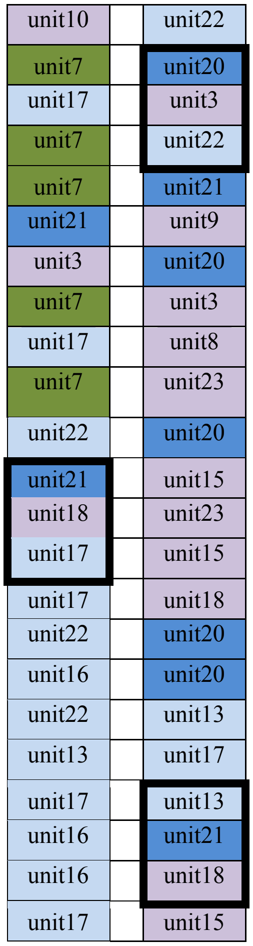
Puerto Rico. Songs from Puerto Rico were the least uniform of the geographical locations. Song96 was particularly anomalous, being composed primarily of long strings of Class Pink (P) interrupted by units from Class Light

Purple (LP), and shorter strings of Class Light Purple interrupted by units from Class Pink. Further discussion of songs from Puerto Rico will exclude this song.

There was no clear backbone in the songs from Puerto Rico – Song70 had a backbone of units from Class Light Blue (LB), Song75 had a backbone of units from Class Light Purple, and Song74 had a backbone composed of units from each. Song70 and Song74 were most alike, sharing the root sequences LB+DB+LP and DB+LP+LB (see Figure 9), in addition to having units from Class Dark Blue (DB) nearly always occur as single units (i.e., there were very few examples of more than one unit from Class Dark Blue in a row). Song75 contained units from Class Yellow (Y) in its patterns, and these patterns appeared to be more complex variants (e.g., additional repeats) on the patterns appearing in the other two songs. In addition, the backbone in song75 was much more difficult to distinguish.

Turks and Caicos. Songs recorded at Turks and Caicos were far more consistent than songs recorded in other locations. All three songs showed a backbone of Class Orange (O), a Class almost exclusive to this region, with drop-ins from Class Light Purple (although the backbone was more mixed between these Classes in song119). In all cases, any sequence of a length of more than three units (with the exception of one sequence in one song) was composed entirely of units from Class Orange and Class Light Purple. In addition, in sequences longer than three to four units, units from Class Light Purple were generally only found singly. Only in song119 were there a sufficient number of sequences with consecutive units from Class Light Purple to warrant mention.

Lesser Antilles. Songs recorded in the Lesser Antilles tended to share more characteristics when they were recorded the same year; however, overall trends across years were present as well. In 1973, the backbone of Lesser Antilles songs was from Class Light Purple; however, in 1976, the backbone was split between Class Light Purple and Class Light Blue. In 1973, the patterns for creating sequences were straightforward. The non-backbone classes (Dark Blue, Dark Purple (DP), and Yellow) never co-occurred; they only existed in sequences with units from Class Light Purple, where they either alternated (i.e., DB+LP+DB+LP), or occurred in doubles (LP+DP+DP+LP or DP+LP+LP+DP). When these patterns occurred in the songs from 1976, they were less organized – for example, instead of simply two units in the middle of a sequence such as DP+LP+LP+DP, the number present was much more variable. More elaborate sequences were suspected, such as DB+LB+G+LB+LB or LP+G+LP+LB; however, they could not be claimed to be entirely consistent as units would occasionally switch places or repeats would be added.



unit10		unit22
unit7		unit20
unit17		unit3
unit7		unit22
unit7		unit21
unit21		unit9
unit3		unit20
unit7		unit3
unit17		unit8
unit7		unit23
unit22		unit20
unit21		unit15
unit18		unit23
unit17		unit15
unit17		unit18
unit22		unit20
unit16		unit20
unit22		unit13
unit13		unit17
unit17		unit13
unit16		unit21
unit16		unit18
unit17		unit15

Figure 9. Sequences from song70 and song74 showing the LB+DB+LP and DB+LP+LB sequences framed in bold. (abbreviations stand for cell color; LB = Light Blue, DB = Dark Blue, LP = Light Purple)

Entropy Analysis

Examining the results of the analysis of entropy (as defined by the number of sequences of a particular length in a song divided by the number of sequences of the same length when the units in randomly generated songs) provided data that corresponds to that of the sequence analysis.

The entropy at each sequence length in a particular song can be graphed (Figure 10). Looking at these figures, the similarities and differences identified by the sequence analysis are evident.

In Puerto Rico, the two most similar songs, song70 and song74 have very similar patterns of change in entropy over time, while song75, the more complex of the songs, occupies a different trace; it finishes with a higher entropy (less organization) than the other two. This is indicative of larger, less organized sequences at longer lengths, and is congruent with the addition of units from Class Yellow (which appears in song75 but not song70 or song74).

In Turks and Caicos, there is a similar pattern. Although the songs in this region are generally more ordered, song102 and song118 are quite similar and their entropy graphs are indicative of this. Song119, as mentioned above, differs from the other two in that the backbone is less obvious and it is the only song in which multiple units from Class Light Purple could be found consecutively. Song119 has higher entropy at longer length sequences.

Finally, the entropy graphs for the Lesser Antilles are particularly interesting because they reflect the uniformity found within the sequence patterns. All four songs have very similar patterns of entropy measurements, most likely owing to similar rules for sequence construction. Additionally, the two songs recorded in 1973 are represented by the two lowest traces and the two songs recorded in 1976 are the two highest traces. This reflects the sequence information previously discussed. The 1976 songs are less organized than the 1973 songs; they have variable numbers of repeated units, and thus higher entropy.

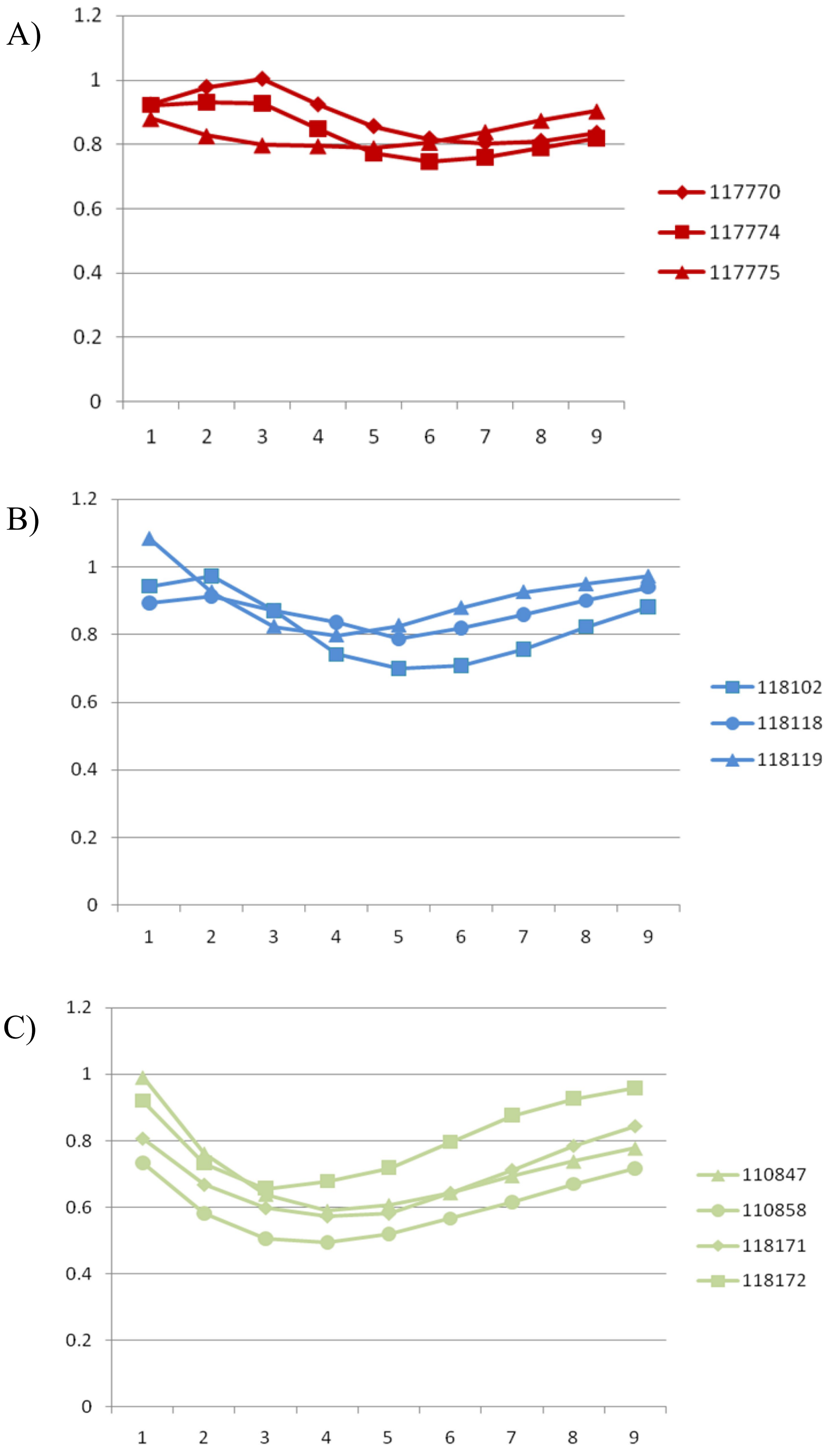


Figure 10. Entropy graphs for sequences of length 2 to 8 units, by song and region. **A)** Puerto Rico, **B)** Turks and Caicos, and **C)** Lesser Antilles (x-axis = entropy, y-axis = sequence length).

Discussion

The HAL model is meant to measure global co-occurrence within language. The Classes identified by the analysis come directly from the whale song input; HAL requires no teaching and does not mathematically alter the input in any way. The Classes identified can be explained on a simplistic level as substitutable grammatical or semantic categories; if a pattern of vocalizations is found to alternate between Class Yellow and Class Light Purple, it does not matter which units from within these Classes are used, as they all serve the same semantic or grammatical purpose. Just as English grammatical rules dictate the three classes "noun, verb, and direct object" to be ordered as such in a sentence, patterns found in humpback whale song may be obeying similar rules.

Frequency Analysis

Analysis of the frequency of occurrence of units in each cluster provides a very basic descriptor of the usage balance in the three geographic areas. This analysis provides information on both the Class and Unit levels; however, it is important to note that the idea of measuring and comparing frequency may be a vast oversimplification of song construction and may not, in the end, be particularly relevant.

Some Classes are almost entirely unique to a particular region, such as Class Orange is to Turks and Caicos. Monitoring the usage of Class Orange would be an excellent way to empirically test the transition of vocalizations between geographic areas as its progress could be tracked into nearby geographical locations.

On the other hand, the Class usage in Puerto Rico appears to be more diverse, supporting the idea that the Puerto Rican songs are at least in part the product of convergence of the songs from neighboring areas.

Important in frequency is the idea of substitutability. If the Classes shown by the HAL model are the result of semantic co-occurrence, then the units in a particular Class should have some degree of basic substitutability. For example, just as different regions of the United States use the words "soda" and "pop" to refer to the same beverage, it is possible that in the Lesser Antilles in 1973 unit20 and unit21 were used to convey the meaning inherent in Class Blue, while in 1976, unit1 and unit12 (also from Class Blue) were used to convey the same meaning.

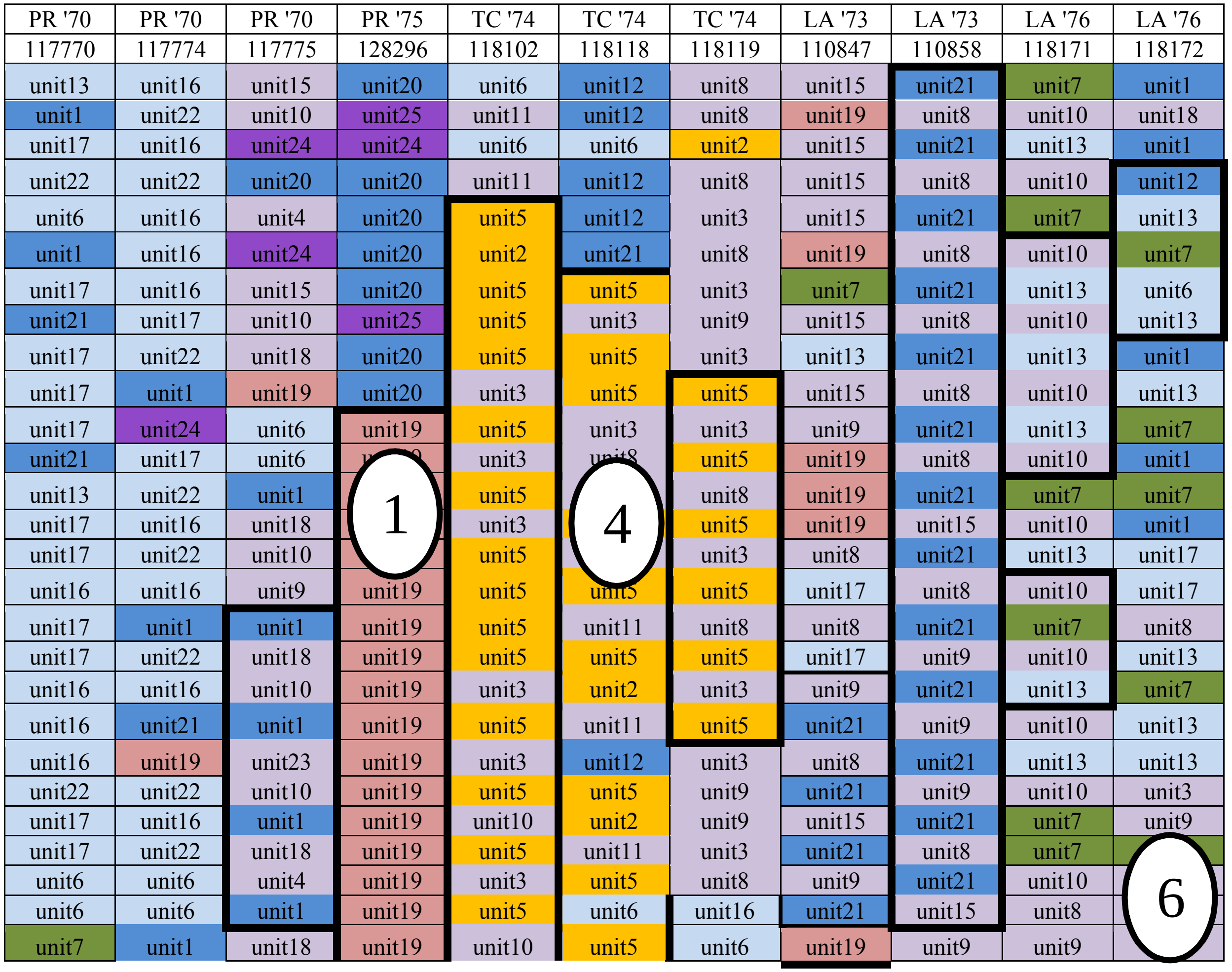
Sequence Analysis

The most insightful information from the HAL analysis is in regard to sequence. When unit identity is analyzed, for example in a conditional probability, the analysis, naturally, looks solely at the unit. In human language, this is the equivalent of only being able to recognize "Boy sees cat" as a grammatical pattern if it is repeated over and over without variation. "Boy sees dog" or "Girl sees cat" would most likely go unnoticed. The sequence analysis was able to identify sequences at the Class level which were unique to both geographical regions (or, in the case of the Lesser Antilles, time) and individual song. The specifics of the most consistent of these patterns have been discussed previously and are highlighted in Figure 11. Using the HAL methodology, the sentences “Boy sees cat,” “Boy sees dog,” and “Girl sees cat” would all be identified as grammatically and semantically similar; “boy” and “girl,” being nouns, would belong to one class and would be identified with the same color; “sees,” a verb, would be a different color; and “cat” and “dog,” both direct objects in this case, would be a third color – creating an identifiable pattern.

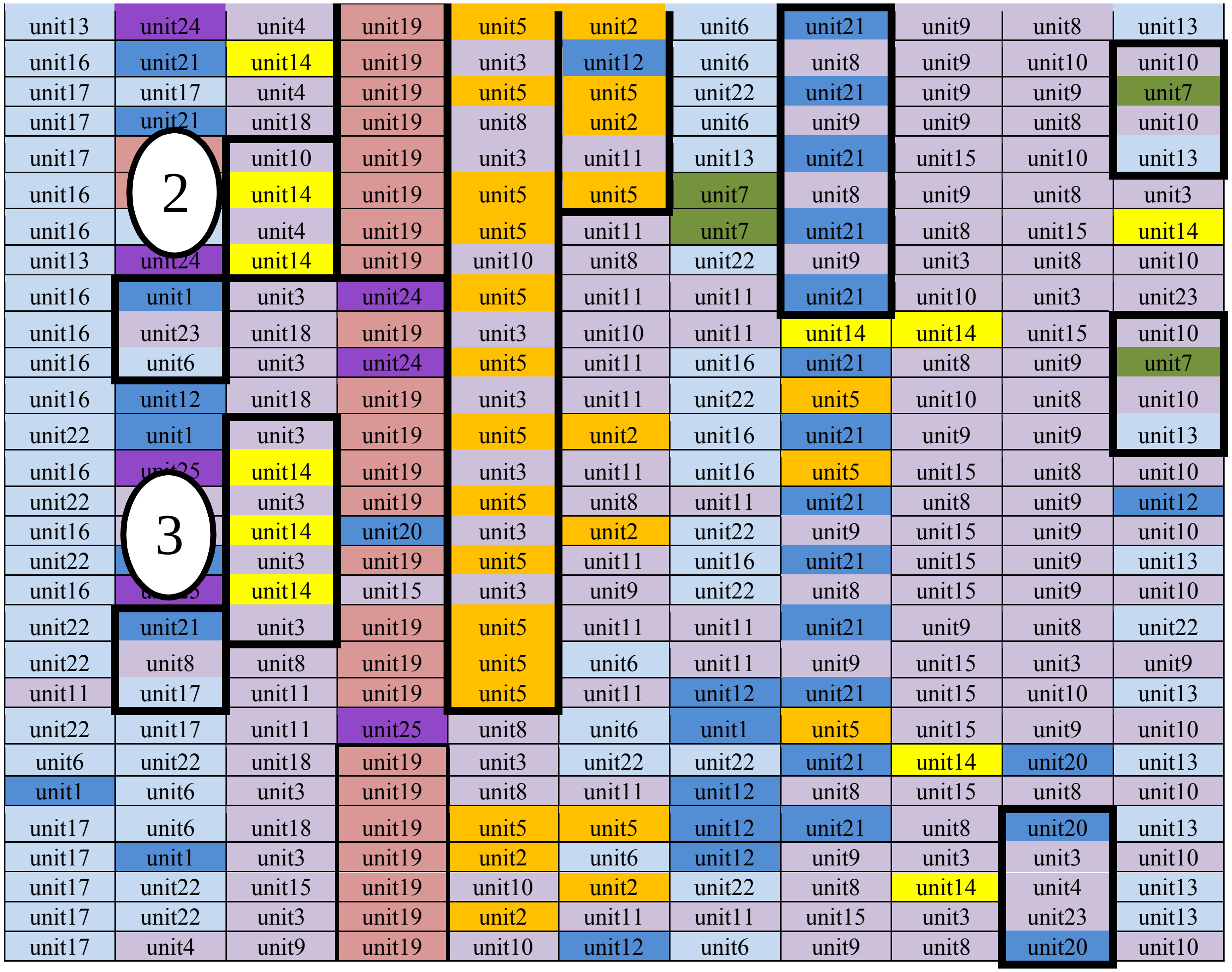
It is important to note that some patterns were not included in the HAL analysis. For example, the HAL analysis showed that no more than three or four units from Class Light Purple ever appeared consecutively in song102 or song118. Yet there is clearly a sequence of six units from Class Light Purple in song118. This contradiction is due to the fact that only sequences which occurred more than five times within a song were considered in the analysis. However, these exceptions were either unique or rare occurrences.

Puerto Rico. Songs recorded in Puerto Rico were the least uniform of the three regions. Song96 was particularly anomalous – it was drastically shorter and contained long repeats of a Class Pink, a class consisting of one unit and found infrequently in other songs. As a result, it was generally excluded from the analysis, although it should be noted that the variation within the Puerto Rican songs was not entirely due to song96.

Song “backbones” in Puerto Rico were either composed of units from Class Light Blue, Class Light Purple, or both. However, in all cases it is important to note the diversity of units here — because these backbones are composed of different units from the same Class, they could only be distinguished in a global analysis. Geographically, Puerto Rico is located between Turks and Caicos (to the north) and the Lesser Antilles (to the south). As a result, if song elements were passed between whales (and populations) over geographic regions, it is plausible that songs recorded in a central area would be a mixture. For example, song70 and song74 both display elements common to song171 (recorded in the Lesser Antilles) — a backbone of units from Class Light Blue and Class Light Purple; drop in units from Class Light Blue, Class Light Purple, and Class Dark Blue; and as well as root patterns composed of units from these Classes.



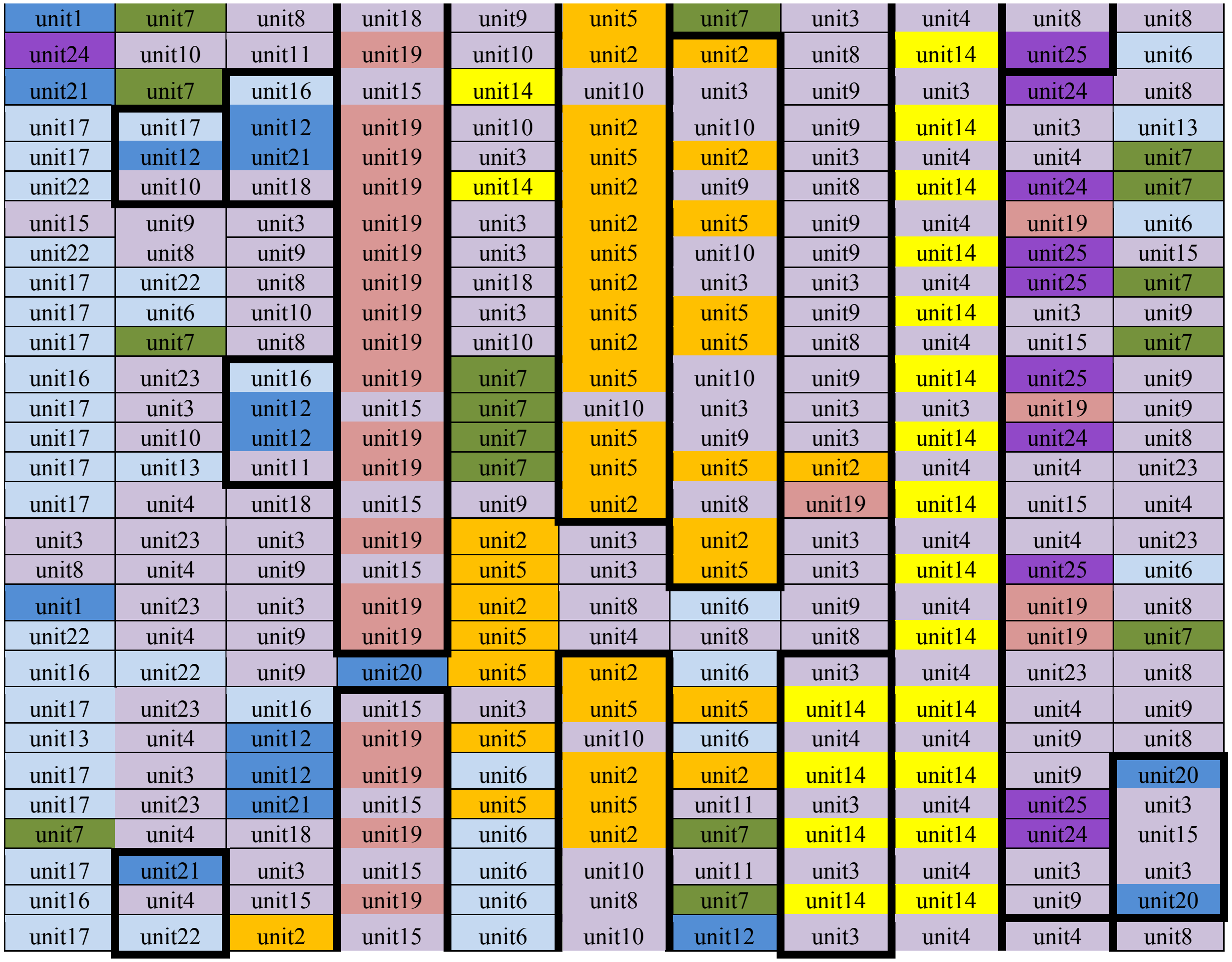
PR '70 117770	PR '70 117774	PR '70 117775	PR '75 128296	TC '74 118102	TC '74 118118	TC '74 118119	LA '73 110847	LA '73 110858	LA '76 118171	LA '76 118172
unit13	unit16	unit15	unit20	unit6	unit12	unit8	unit15	unit21	unit7	unit1
unit1	unit22	unit10	unit25	unit11	unit12	unit8	unit19	unit8	unit10	unit18
unit17	unit16	unit24	unit24	unit6	unit6	unit2	unit15	unit21	unit13	unit1
unit22	unit22	unit20	unit20	unit11	unit12	unit8	unit15	unit8	unit10	unit12
unit6	unit16	unit4	unit20	unit5	unit12	unit3	unit15	unit21	unit7	unit13
unit1	unit16	unit24	unit20	unit2	unit21	unit8	unit19	unit8	unit10	unit7
unit17	unit16	unit15	unit20	unit5	unit5	unit3	unit7	unit21	unit13	unit6
unit21	unit17	unit10	unit25	unit5	unit3	unit9	unit15	unit8	unit10	unit13
unit17	unit22	unit18	unit20	unit5	unit5	unit3	unit13	unit21	unit13	unit1
unit17	unit1	unit19	unit20	unit3	unit5	unit5	unit15	unit8	unit10	unit13
unit17	unit24	unit6	unit19	unit5	unit3	unit3	unit9	unit21	unit13	unit7
unit21	unit17	unit6	unit19	unit3	unit8	unit5	unit19	unit8	unit10	unit1
unit13	unit22	unit1	[illegible]	unit5	[illegible]	unit8	unit19	unit21	unit7	unit7
unit17	unit16	unit18	①	unit3	④	unit5	unit19	unit15	unit10	unit1
unit17	unit22	unit10	[illegible]	unit5	[illegible]	unit3	unit8	unit21	unit13	unit17
unit16	unit16	unit9	unit19	unit5	unit5	unit5	unit17	unit8	unit10	unit17
unit17	unit1	unit1	unit19	unit5	unit11	unit8	unit8	unit21	unit7	unit8
unit17	unit22	unit18	unit19	unit5	unit5	unit5	unit17	unit9	unit10	unit13
unit16	unit16	unit10	unit19	unit3	unit2	unit3	unit9	unit21	unit13	unit7
unit16	unit21	unit1	unit19	unit5	unit11	unit5	unit21	unit9	unit10	unit13
unit16	unit19	unit23	unit19	unit3	unit12	unit3	unit8	unit21	unit13	unit13
unit22	unit22	unit10	unit19	unit5	unit5	unit9	unit21	unit9	unit10	unit3
unit17	unit16	unit1	unit19	unit10	unit2	unit9	unit15	unit21	unit7	unit9
unit17	unit22	unit18	unit19	unit5	unit11	unit3	unit21	unit8	unit7	[illegible]
unit6	unit6	unit4	unit19	unit3	unit5	unit8	unit9	unit21	unit10	⑥
unit6	unit6	unit1	unit19	unit5	unit6	unit16	unit21	unit15	unit8	[illegible]
unit7	unit1	unit18	unit19	unit10	unit5	unit6	unit19	unit9	unit9	[illegible]



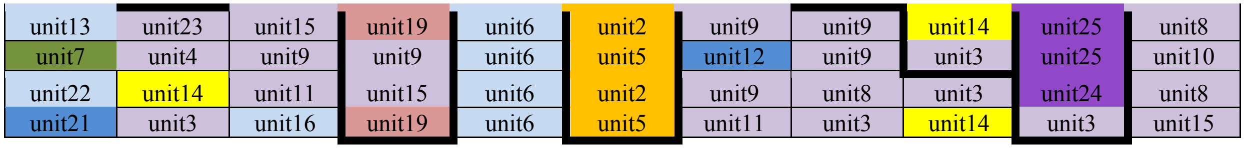
unit13	unit24	unit4	unit19	unit5	unit2	unit6	unit21	unit9	unit8	unit13
unit16	unit21	unit14	unit19	unit3	unit12	unit6	unit8	unit9	unit10	unit10
unit17	unit17	unit4	unit19	unit5	unit5	unit22	unit21	unit9	unit9	unit7
unit17	unit21	unit18	unit19	unit8	unit2	unit6	unit9	unit9	unit8	unit10
unit17	[illegible]	unit10	unit19	unit3	unit11	unit13	unit21	unit15	unit10	unit13
unit16	②	unit14	unit19	unit5	unit5	unit7	unit8	unit9	unit8	unit3
unit16	[illegible]	unit4	unit19	unit5	unit11	unit7	unit21	unit8	unit15	unit14
unit13	unit24	unit14	unit19	unit10	unit8	unit22	unit9	unit3	unit8	unit10
unit16	unit1	unit3	unit24	unit5	unit11	unit11	unit21	unit10	unit3	unit23
unit16	unit23	unit18	unit19	unit3	unit10	unit11	unit14	unit14	unit15	unit10
unit16	unit6	unit3	unit24	unit5	unit11	unit16	unit21	unit8	unit9	unit7
unit16	unit12	unit18	unit19	unit3	unit11	unit22	unit5	unit10	unit8	unit10
unit22	unit1	unit3	unit19	unit5	unit2	unit16	unit21	unit9	unit9	unit13
unit16	unit25	unit14	unit19	unit3	unit11	unit16	unit5	unit15	unit8	unit10
unit22	[illegible]	unit3	unit19	unit5	unit8	unit11	unit21	unit8	unit9	unit12
unit16	③	unit14	unit20	unit3	unit2	unit22	unit9	unit15	unit9	unit10
unit22	[illegible]	unit3	unit19	unit5	unit11	unit16	unit21	unit15	unit9	unit13
unit16	unit25	unit14	unit15	unit3	unit9	unit22	unit8	unit15	unit9	unit10
unit22	unit21	unit3	unit19	unit5	unit11	unit11	unit21	unit9	unit8	unit22
unit22	unit8	unit8	unit19	unit5	unit6	unit11	unit9	unit15	unit3	unit9
unit11	unit17	unit11	unit19	unit5	unit11	unit12	unit21	unit15	unit10	unit13
unit22	unit17	unit11	unit25	unit8	unit6	unit1	unit5	unit15	unit9	unit10
unit6	unit22	unit18	unit19	unit3	unit22	unit22	unit21	unit14	unit20	unit13
unit1	unit6	unit3	unit19	unit8	unit11	unit12	unit8	unit15	unit8	unit10
unit17	unit6	unit18	unit19	unit5	unit5	unit12	unit21	unit8	unit20	unit13
unit17	unit1	unit3	unit19	unit2	unit6	unit12	unit9	unit3	unit3	unit10
unit17	unit22	unit15	unit19	unit10	unit2	unit22	unit8	unit14	unit4	unit13
unit17	unit22	unit3	unit19	unit2	unit11	unit11	unit15	unit3	unit23	unit13
unit17	unit4	unit9	unit19	unit10	unit12	unit6	unit9	unit8	unit20	unit10

5

unit22	unit23	unit9	unit19	unit2	unit8	unit12	unit9	unit3	[illegible]	unit12
unit6	unit18	unit11	unit19	unit5	unit3	unit11	unit9	unit9	[illegible]	unit4
unit6	unit4	unit23	unit19	unit8	unit6	unit22	unit9	unit3	[illegible]	unit23
unit1	unit23	unit16	unit20	unit9	unit11	unit12	unit9	unit8	[illegible]	unit10
unit22	unit20	unit11	unit25	unit8	unit8	unit12	unit8	unit3	unit20	unit13
unit16	unit17	unit18	unit20	unit4	unit5	unit1	unit9	unit14	unit15	unit10
unit17	unit18	unit3	unit19	unit8	unit3	unit7	unit9	unit3	unit10	unit13
unit17	unit3	unit18	unit19	unit5	unit2	unit13	unit8	unit3	unit3	unit10
unit16	unit6	unit3	unit20	unit8	unit5	unit13	unit9	unit14	unit20	unit13
unit22	unit1	unit20	unit19	unit3	unit2	unit13	unit9	unit3	unit20	unit10
unit16	unit22	unit9	unit19	unit2	unit5	unit7	unit15	unit3	unit15	unit13
unit17	unit16	unit11	unit19	unit5	unit2	unit3	unit20	unit14	unit3	unit7
unit22	unit8	unit16	unit25	unit15	unit5	unit7	unit14	unit9	unit14	unit7
unit12	unit8	unit12	unit19	unit8	unit8	unit10	unit15	unit3	unit8	unit3
unit6	unit23	unit11	unit15	unit2	unit3	unit7	unit15	unit14	unit18	unit7
unit1	unit8	unit18	unit25	unit7	unit5	unit7	unit20	unit3	unit4	unit8
unit22	unit23	unit3	unit19	unit10	unit2	unit7	unit14	unit3	unit8	unit9
unit17	unit16	unit15	unit15	unit8	unit4	unit7	unit15	unit14	unit18	unit8
unit17	unit22	unit3	unit19	unit13	unit2	unit7	unit20	unit3	unit4	unit7
unit17	unit6	unit18	unit19	unit9	unit5	unit7	unit15	unit3	unit9	unit8
unit17	unit7	unit8	unit19	unit10	unit2	unit7	unit14	unit14	unit9	unit9
unit17	unit22	unit16	unit20	unit7	unit8	unit3	unit20	unit3	unit8	unit8
unit16	unit1	unit11	unit19	unit7	unit5	unit13	unit20	unit14	unit9	unit13
unit17	unit22	unit21	unit19	unit10	unit2	unit7	unit9	unit8	unit8	unit9
unit17	unit17	unit18	unit19	unit3	unit5	unit9	unit14	unit18	unit24	unit17
unit16	unit8	unit3	unit19	unit14	unit10	unit3	unit20	unit3	unit8	unit13
unit22	unit10	unit9	unit19	unit3	unit5	unit13	unit14	unit14	unit23	unit7
unit6	unit23	unit9	unit15	unit7	unit2	unit7	unit3	unit3	unit18	unit7
unit13	unit21	unit9	unit19	unit7	unit10	unit3	unit14	unit14	unit3	unit6



unit1	unit7	unit8	unit18	unit9	unit5	unit7	unit3	unit4	unit8	unit8
unit24	unit10	unit11	unit19	unit10	unit2	unit2	unit8	unit14	unit25	unit6
unit21	unit7	unit16	unit15	unit14	unit10	unit3	unit9	unit3	unit24	unit8
unit17	unit17	unit12	unit19	unit10	unit2	unit10	unit9	unit14	unit3	unit13
unit17	unit12	unit21	unit19	unit3	unit5	unit2	unit3	unit4	unit4	unit7
unit22	unit10	unit18	unit19	unit14	unit2	unit9	unit8	unit14	unit24	unit7
unit15	unit9	unit3	unit19	unit3	unit2	unit5	unit9	unit4	unit19	unit6
unit22	unit8	unit9	unit19	unit3	unit5	unit10	unit9	unit14	unit25	unit15
unit17	unit22	unit8	unit19	unit18	unit2	unit3	unit3	unit4	unit25	unit7
unit17	unit6	unit10	unit19	unit3	unit5	unit5	unit9	unit14	unit3	unit9
unit17	unit7	unit8	unit19	unit10	unit2	unit5	unit8	unit4	unit15	unit7
unit16	unit23	unit16	unit19	unit7	unit5	unit10	unit9	unit14	unit25	unit9
unit17	unit3	unit12	unit15	unit7	unit10	unit3	unit3	unit3	unit19	unit9
unit17	unit10	unit12	unit19	unit7	unit5	unit9	unit3	unit14	unit24	unit8
unit17	unit13	unit11	unit19	unit7	unit5	unit5	unit2	unit4	unit4	unit23
unit17	unit4	unit18	unit15	unit9	unit2	unit8	unit19	unit14	unit15	unit4
unit3	unit23	unit3	unit19	unit2	unit3	unit2	unit3	unit4	unit4	unit23
unit8	unit4	unit9	unit15	unit5	unit3	unit5	unit3	unit14	unit25	unit6
unit1	unit23	unit3	unit19	unit2	unit8	unit6	unit9	unit4	unit19	unit8
unit22	unit4	unit9	unit19	unit5	unit4	unit8	unit8	unit14	unit19	unit7
unit16	unit22	unit9	unit20	unit5	unit2	unit6	unit3	unit4	unit23	unit8
unit17	unit23	unit16	unit15	unit3	unit5	unit5	unit14	unit14	unit4	unit9
unit13	unit4	unit12	unit19	unit5	unit10	unit6	unit4	unit4	unit9	unit8
unit17	unit3	unit12	unit19	unit6	unit2	unit2	unit14	unit14	unit9	unit20
unit17	unit23	unit21	unit15	unit5	unit5	unit11	unit3	unit4	unit25	unit3
unit7	unit4	unit18	unit19	unit6	unit2	unit7	unit14	unit14	unit24	unit15
unit17	unit21	unit3	unit15	unit6	unit10	unit11	unit3	unit4	unit3	unit3
unit16	unit4	unit15	unit19	unit6	unit8	unit7	unit14	unit14	unit9	unit20
unit17	unit22	unit2	unit15	unit6	unit10	unit12	unit3	unit4	unit4	unit8



unit13	unit23	unit15	unit19	unit6	unit2	unit9	unit9	unit14	unit25	unit8
unit7	unit4	unit9	unit9	unit6	unit5	unit12	unit9	unit3	unit25	unit10
unit22	unit14	unit11	unit15	unit6	unit2	unit9	unit8	unit3	unit24	unit8
unit21	unit3	unit16	unit19	unit6	unit5	unit11	unit3	unit14	unit3	unit15

1 = long string of Class Pink almost entirely unique to song96

2, 3 = DB+LP+LB root sequence characteristic of songs from Puerto Rico. Note the actual units differ, i.e. the DB in example 2 is a unit 1, and the DB in example 2 is a unit 21.

4 = Class Orange backbone with Class Light Purple drop ins. Note that the Class Orange backbone is comprised of several different units, as are the Class Light Purple drop ins.

5 = DP+LP+LP+DP root sequence characteristic of songs from the Lesser Antilles

6 = LP+G+LP+LB; these and other, more elaborate variations appeared to be characteristic of songs from the Lesser Antilles recorded in 1976, however a specific pattern could not be identified

Figure 11. Examples of sequences found. Note that sequences are made from Classes, and units within a Class may vary. Colors correspond to clusters identified by HAL and discussed in text. Sequences of note are outlined in bold and numbers identify and explain sequences more specifically. Place of recording (PR = Puerto Rico, TC = Turks and Caicos, LA = Lesser Antilles), year of recording, and song number are noted for each song.

Turks and Caicos. Songs from Turks and Caicos were categorized by both their consistency and the presence of Class Orange, which was almost exclusively unique to this region (the highest occurrence in any other song was 1.6%, and in any other song it appeared in its frequency was less than 1%). For all intents and purposes, any sequence with more than three to four units was composed solely of units from Class Orange and Class Purple. In addition, longer sequences saw units from Class Purple only in single occurrences. Units from other Classes were recorded in Turks and Caicos, but only in shorter sequences. This lack of diversity may be evidence of a smaller "vocabulary" in the Turks and Caicos region; Class Orange is composed of only two units, and although all of the units from Class Purple appear at some time in the region, there was a distinct bias towards unit3, unit8, unit9, unit10, and unit11. Out of the nine units in Class Light Purple, these five accounted for 95.4% of the occurrences of the Class in the Turks and Caicos. This effectively created a vocabulary of seven units with which to build sequences of song at length greater than three to four units in this geographic region. For purposes of comparison, the song with the next smallest vocabulary was song74, which had 16 units and the only potential bias being towards unit16 and unit22 (Class Light Blue) — two units representing 75% of the occurrences of their five unit Class. A smaller vocabulary creates less lexical and/or semantic flexibility and thus less diverse songs. With this knowledge, one might hypothesize that this population of whales was relatively new and/or fairly isolated at the time these data were collected, which would consequently slow the process of change in song. Thinking back to the diversity of songs in Puerto Rico, a scenario such as a new or isolated population in the Turks and Caicos might explain why there appears to be very little similarity between the songs from these two regions, relative to the similarity between the songs of Puerto Rico and the Lesser Antilles. Horizontal cultural transmission such as this has been documented for quite some time in humpback whales (Winn et al., 1981), and recently was found to occur in an "outward ripple-like" pattern (Garland et al., 2011; Noad et al., 2000).

Lesser Antilles. Songs in the Lesser Antilles appeared to get more complex over time. Over the course of the three years between sets of recordings, the backbone of the songs changed from exclusively Class Light Purple to a combination of Class Light Blue and Class Light Purple. In addition, the root patterns became more complex and sometimes even unpredictable. It is possible that there is a pattern of divergence here. The songs recorded in 1973 are fairly rule based and similar, containing root patterns which alternate Classes in a x+y+x+y or x+y+y+x pattern. In 1976 the songs change more. As previously mentioned, song171 is increasingly similar to some of the songs recorded in Puerto Rico, as if the whale(s) singing this song influenced or was influenced by singers in this region. Song172 becomes much more diverse and much less predictable; almost as if a new song or variation thereof were in the process of being created.

Entropy Analysis

The results of the clustering are consistent with the overall entropy analysis. The variation in the Puerto Rican songs is evident; they show no agreement on the sequence length at which organization is the highest (song70 – 8 units, song74 – 7 units, song75 – 6 units). The entropy graphs of Turks and Caicos and the Lesser Antilles show that songs recorded in 1976 had more entropy than those recorded in 1973, consistent with the prediction that singers in the Lesser Antilles are branching out over time, the. In Turks and Caicos, one needs to bear in mind that the drastic decrease in the number of Classes used to create sequences occurred at sequences longer than three to four units. At five units, the decrease in the number of units causes increased predictability (decreased entropy). However, entropy increases from this point on. Just as prediction is a moot point when a series is random, so is it when a series is too repetitive; and even more so when the series is repetitive with one insertion at random intervals, as becomes the case in longer sequences in the Turks and Caicos with the backbone and drop-ins.

Evolutionary Advantages. Perhaps more interesting because of the lack of a common ancestor (or at least a recent one, as there is between humans and apes), is that similarities in the communication systems of humans and non-primates would indicate analogous situations in which language evolved in multiple phylogenetic lines (Weissengruber, Forstenpointner, Peters, Kubber-Heiss, & Fitch, 2002). From this we could hypothesize situations in which communication trait or ability would be selected for. For example, a finding that marine mammals use syntax and a grammar-like structure in their communications would provide evidence that at some point in the evolution of both species the environment supported the development of adaptive complex communication. It might then be possible to compare the evolutionary history of both species to determine what these conditions were and the mechanisms by which language developed. This comparison would speak to the importance of communication as a survival tactic, just as the analogous evolution of flight in birds, insects, and bats (mammals) speaks to the advantage of the development of flight.

HAL Validates the SOM Methodology. Because the HAL model builds the multi-dimensional space it works with solely from the input provided, it is constrained by the “garbage in, garbage out” theory. Meaning, if the input to HAL has no inherent organization, the multidimensional space created by the model – and therefore the output – will also be without organization (Burgess, 1998; Lund & Burgess, 1996). As a result of this characteristic, the existence of organized, interpretable, HAL output can provide additional support for the validity of the SOM as a method. In the work presented here, the HAL output can be interpreted, showing that the initial organization had structure. Because the “rules” by which HAL learns to classify units are based only on what can be gleaned from the input stream, the existence of organized output shows that HAL was indeed able to glean said rules, and therefore that the initial input (or SOM output) also had organization.

Conclusion

The analyses presented here are useful both as tests of a new method of identifying themes in whale song and as an empirical study of whale song in the

Caribbean. In terms of methodology, the combined use of SOM and HAL techniques for unit classification and sequence analysis, respectively, allows for fully automated and quantitative analysis of thematic structure. In contrast to subjective methods of identifying themes, the method described here can be strictly replicated without the possibility of disagreement over the identity of individual sounds or disagreement over whether specific sequences of songs are examples of a particular theme. This analysis of song thematic structure may also represent an empirical contribution because defining the difference between different themes in quantitative rather than subjective terms may reveal details of thematic structure that are not readily apparent from listening to a song or viewing its spectrogram.

A secondary goal of this study was to provide support for the existence of cultural transmission of song in humpback whales by showing dialects which were unique to regions, but which seemed to share particular traits with nearby regions or across periods of time. There is some evidence for this in the case of songs recorded in Puerto Rico, which is located geographically between the Lesser Antilles and Turks and Caicos – the other two locations in which data were recorded. Songs in Puerto Rico were much more diverse and a uniform, overall, region-specific pattern was harder to identify than in the other regions. However, the songs did share some characteristics of the songs recorded in the other two areas, most notably was a similarity between song70 and song74 and song171 (the latter of which was recorded in the Lesser Antilles).

Songs recorded in both the Lesser Antilles and in Turks and Caicos show signs of change by forms of dispersion (or lack thereof). The Turks and Caicos songs are quite uniform and conform to a fairly strict set of rules. In addition, the vocabulary (both number of unit types and number of Class types) used by whales in this area is very small. This may be indicative of a population that is isolated by geography to some extent, and which, as a result, is not exposed to songs from other populations which might in turn result in novelties in their own songs. In addition, this small vocabulary could be indicative of a newly developing population that is just establishing a distinct dialect and therefore would show a specific necessity to maintain a clear, easy to sing dialect with few elements in order to promote retention. In the Lesser Antilles, there is possible evidence of divergence of songs over time, as in 1973 the songs measured are far more similar than those measured in 1976. This is roughly the same pattern which would be expected in the future for the songs from the Turks and Caicos.

Additionally, the humpback whale results do provide support for global co-occurrence processes in the humpback whale song, as many of these patterns (plus the song backbones) would not have been identifiable without initially identifying Class. Classes, whether they are semantic or grammatical or represent a combination of the two concepts, do appear to be involved in the creation of humpback whale song and distinct patterns of Classes, which may be similar to patterns in human language, could be identified in all of the songs. For example, one could compare the English grammar pattern noun+verb+direct object to a pattern in the humpback song such as Class Light Blue + Class Dark Blue + Class Light Purple. In both cases, the categories (be they human grammar or whale Class) consist of a variety of individual elements, and any of these elements would serve the proper contextual purpose.

Limitations of These Studies

Much of the work presented here is novel. Because higher cognitive issues in non-human species are being dealt with, these results will naturally – and rightfully – be regarded with skepticism. However, this is the first time, to our knowledge, a concept such as contextual co-occurrence has been applied to animal vocalizations of any sort, and the results have been positive enough that they warrant further exploration.

In addition, we are aware of the inherent assumption throughout this paper that there is meaning to be found within humpback whale song in the first place. For now, this is a necessary assumption, as we cannot show evidence for the existence of something unless we formulate some very general idea – however incorrect it might be – of what it is we are actually looking for. To even begin to solve “2 + 2,” it is imperative to know that a number is being sought, as opposed to a letter, word, object, or person. In addition, because it has been established that the units in humpback whale song are ordered, it must logically follow that the order is, in some very broad sense, meaningful; whether this is cognitive, conscious, or instinctual is not something we attempt to address. It is our hope that, regardless of terminology, that the patterns, regularities, or syntax identified here can be used as a starting point to understanding purpose behind the vocalizations.

Methodologically, it is also important to one keep in mind the small sample size involved in this analysis. Although the actual songs contained large numbers of units, the number of Classes is relatively small, and several Classes only contain one unit. In addition, although the region comparison provides intriguing avenues to investigate, a larger number of songs from each region will be required for definitive conclusions in future analysis.

Lastly, statistical techniques for the confirmation of many the patterns identified herein are not yet available. The color coding technique was developed specifically for these experiments as a method of visualizing patterns created by Classes, as opposed to individual units. Techniques such as the entropy analysis will identify common patterns in the data, however, due to their categorical nature, there is little beyond qualitative description which can be used to describe the patterns once they have been found.

Future Studies

First and foremost, future studies should be concerned with confirmation of these results. Before anything can be definitively stated about the use of global co-occurrence in animal communication, many more successful studies must be presented and the parameters used in models must be successfully standardized. In addition, a universal technique for examining data such as these must be adopted – be it the color coded analysis developed and presented here, or another technique. This includes, as discussed above, successful statistical confirmation.

One important way to validate that the idea that global co-occurrence is a relevant construct in animal communication systems is by subjecting vocalization streams such as the ones tested here to other models — similar to the HAL model

— which operate by very different mechanisms, yet produce similar results. Two models which are closely related to HAL are Latent Semantic Analysis (LSA; Landauer, Foltz, & Laham, 1998) and the Simple Recurrent Network (SRN; Elman, 1990). Both models — like HAL — learn without human input or training, and both models have shown similar results with similar data sets (Li, Burgess, & Lund, 2000; Yan, Li, & Song, 2004). These models would be the most appropriate starting point for cross model comparative studies.

References

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. London: Sage.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *Emergence of language* (pp. 359-380). Hillsdale, NJ: Lawrence Erlbaum.

Au, W. W. L., Frankel, A., Helweg, D. A., & Cato, D. H. (2001). Against the humpback whale sonar hypothesis. *IEEE Journal of Oceanic Engineering, 26,* 295-300.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Brown, J. C., & Hodgins-Davis, A. (2006). Classification of vocalizations of killer whales using dynamic time warping. *The Journal of the Acoustical Society of America*, *119*, EL34-EL40.

Buck, J. R., & Tyack, P. L. (1993). A quantitative measure of similarity for tursiops truncatus signature whistles. *The Journal of the Acoustical Society of America*, *94*, 2497-2506.

Burgess, C. (1998). From simple associations to the building blocks of language: Modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, & Computers*, *30*, 188-198.

Burgess, C., & Lund, K. (2000). The dynamics of meaning in memory. In E. Dietrich & A. B. Markman (Eds.), *Conceptual and representational change in humans and machines* (pp. 117-156). Mahwah, NJ: Lawrence Erlbaum Associates.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, *7*, 300-307.

Chu, K. (1988). Dive times and ventilation patterns of singing humpback whales (*Megaptera novaeangliae*). *Canadian Journal of Zoology, 66*, 1322-1327.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*, 127-137.

Clemins, P. J., Johnson, M. T., Leong, K. M., & Savage, A. (2005). Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. *The Journal of the Acoustical Society of America*, *117*, 956-963.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, *142*, 397-421.

Darling, J. D. (1983). *Migrations, abundance and behaviour of 'Haiwaiian' humpback whales (Megaptera novaeangliae).* (Unpublished dodctoral dissertation). University of California Santz Cruz, Santa Cruz, California.

Darling, J. D., Jones, M. E., & Nicklin, C. P. (2006). Humpback whale songs: Do they organize males during the breeding season? *Behaviour, 143*, 1051-1101.

Deecke, V. B., Ford, J. K. B., & Spong, P. (1999). Quantifying complex patterns of bioacoustic variation: Use of a neural network to compare killer whale (*Orcinus orca*) dialects. *Journal of the Acoustical Society of America, 105*, 2499-2507.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.

Elman, J. L. (1993). Learning and development in neural networks - the importance of starting small. *Cognition*, *48*, 71-99.

Elman, J. L. (1995). Language as a dynamical system. In R. F. Port & T. van Gelder (Eds.), (pp. 195-223). Cambridge, MA: MIT Press.

Eriksen, N., Miller, L. A., Tougaard, J., & Helweg, D. A. (2005). Cultural change in the songs of humpback whales (*Megaptera novaeangliae*) from Tonga. *Behaviour*, *142*, 305-328.

Estes, K. G., Evans, C. S., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science*, *18*, 254-260.

Frazer, L. N., & Mercado, E. (2000). A sonar model for humpback whale song. *IEEE Journal of Oceanic Engineering, 25*, 160-182.

Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D.,...Noad, M. J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, *21*, 687-91. doi:10.1016/j.cub.2011.03.019

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*, 1204-1207. doi: 10.1038/nature04675.

Green, S. R., Mercado, E., Pack, A. A., & Herman, L. M. (2007). Characterizing patterns within humpback whale (*Megaptera novaeangliae*) songs. *Aquatic Mammals, 33*, 202-213.

Handel, S., Todd, S. K., & Zoidis, A. M. (2009). Rhythmic structure in humpback whale (*Megaptera novaeangliae*) songs: Preliminary implications for song production and perception. *The Journal of the Acoustical Society of America, 125*, EL225-EL230.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53-B64.

Helweg, D. A., Frankel, A. S., Mobley, J. R., & Herman, L. M. (1992). Humpback whale songs: Our current understanding. In J. A. Thomas, R. A. Kastelein, & A. Y. Supin (Eds.), *Marine mammal sensory systems* (pp. 459-484). New York: Plenum.

Herman, L. M., Kuczaj, S. A., & Holder, M. D. (1993). Responses to anomalous gestural sequences by a language-trained dolphin: Evidence for processing of semantic relations and syntactic information. *Journal of Experimental Psychology*, *122*, 184-94.

Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition*, *16*, 129-219. doi: 10.1016/0010-0277(84)90003-9

Herman, L. M., & Tavolga, W. N. (1980). The communication systems of cetaceans. In L. M. Herman (Ed.), *Cetacean behavior: Mechanisms and functions* (pp. 149-209). New York: Wiley Interscience.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language - (Reply to Fitch, Hauser, and Chomsky). *Cognition*, *97*, 211-225.

Janik, V. M. (1999). Pitfalls in the categorization of behaviour: A comparison of dolphin whistle classification methods. *Animal Behavior*, *57*, 133-143.

Kaufman, A. B. (2010). *Assessing animal vocal communication using the Hyperspace Analog to Language (HAL) model*. (Unpublished dodctoral dissertation). University of California Riverside, Riverside, California.

Kaufman, A. B., Colbert-White, E. N., & Burgess, C. (under review). Higher-order syntactical structures in an African Grey parrot's vocalizations: Evidence from the Hyperspace Analog to Language (HAL) model.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, *25*, 259-284.

Li, P., Burgess, C., & Lund, K. (2000). The acquisition of word meaning through global lexical co-occurrences. In E. V. Clark (Ed.), *Proceedings of the 30th Child Language Research Forum* (pp. 167-178). Stanford, CA: Center for the Study of Language and Information.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*, 203-208.

McCowan, B. (1995). A new quantitative technique for categorizing whistles using simulated signals and whistles from captive bottlenose dolphins (Delphinidae, *Tursiops truncatus*). *Ethology*, *100*, 177–193.

McCowan, B., Doyle, L. R., Kaufman, A. B., Hanser, S. F., & Burgess, C. (2008). Detection and estimation of complexity and contextual flexibility in nonhuman animal communication systems. In U. Griebel & K. Oller (Eds.), *Evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication* (pp. 281-304). Cambridge, MA: MIT Press.

McCowan, B., Doyle, L. R., Jenkins, J. M., & Hanser, S. F. (2005). The appropriate use of Zipf's law in animal communication studies. *Animal Behaviour*, *69*, F1-F7. doi: 10.1016/j.anbehav.2004.09.002

Melendez, K. V., Jones, D. L., & Feng, A. S. (2006). Classification of communication signals of the little brown bat. *The Journal of the Acoustical Society of America*, *120*, 1095-1102.

Mellinger, D. K. (2008). A neural network for classifying clicks of Blainville's beaked whales (*Mesoplodon densirostris*). *Canadian Acoustics, 36*, 55-59.

Menzel, E. W. (1973). Chimpanzee spatial memory organization. *Science*, *182*, 943-945.

Miksis-Olds, J. L., Buck, J. R., Noad, M. J., Cato, D. H., & Stokes, M. D. (2008). Information theory analysis of Australian humpback whale song. *The Journal of the Acoustical Society of America, 124*, 2385-2393.

Murray, S. O., Mercado, E., & Roitblat, H. L. (1998). The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. *Journal of the Acoustical Society of America*, *104*, 3626-3633.

Newman, M. M., Yeh, P. J., & Price, T. D. (2008). Song variation in a recently founded population of the dark-eyed junco (*Junco hyemalis*). *Ethology*, *114*, 164-173.

Noad, M. J., Cato, D. H., Bryden, M. M., Jenner, M. N., & Jenner, K. C. (2000). Cultural revolution in whale songs. *Nature*, *408*, 537. doi: 10.1038/35046199.

Nowak, M. A., & Komarova, N. L. (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences*, *5*, 288-295.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys use affixation to alter call meaning. *PLoS One*, *4*, e7808.

Parsons, E. C. M., Wright, A. J., & Gore, M. A. (2008). The nature of humpback whale (*Megaptera novaeangliae*) song. *Journal of Marine Animals and their Ecology, 1*(1), 22-31.

Payne, R., & McVay, S. (1971). Songs of humpback whales. *Science, 173*, 583-597.

Payne, K., & Payne, R. (1985). Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, *68*, 89-114. doi: 10.1111/j.1439-0310.1985.tb00118.x

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*, 201-236.

Placer, J., Slobodchikoff, C. N., Burns, J., Placer, J., & Middleton, R. (2006). Using self-organizing maps to recognize acoustic units associated with information content in animal vocalizations. *The Journal of the Acoustical Society of America, 119*, 3140-3146.

Premack, D. (2004). Is language the key to human intelligence? *Science*, *303*, 318-320.

Rendell, L. E., & Whitehead, H. (2001). Culture in whales and dolphins. *Behavioral and Brain Sciences, 24*, 309-382

Rickwood, P., & Taylor, A. (2008). Methods for automatically analyzing humpback song units. *The Journal of the Acoustical Society of America, 123*, 1763-1772.

Riesch, R., & Deecke, V. B. (2011). Whistle communication in mammal-eating killer whales (*Orcinus orca*): Further evidence for acoustic divergence between ecotypes. *Behavioral Ecology and Sociobiology*, *65*, 1377-1387. doi: 10.1007/s00265-011-1148-8

Riesch, R., Ford, J. K. B., & Thomsen, F. (2008). Stability and group specificity of stereotyped whistles in resident killer whales, *Orcinus orca*, off British Columbia. *Animal Behaviour*, *71*, 79-91. doi: 10.1016/j.anbehav.2005.03.026

Saffran, J., Hauser, M. D., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*, 479-500.

Savage-Rumbaugh, E. S. (1993). Language learnability in man, ape, and dolphin. In H. L. Roitblat, L. M. Herman, & P. E. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 457-473). Hillsdale, NJ: Lawrence Erlbaum.

Scarl, J. C., & Bradbury, J. W. (2009). Rapid vocal convergence in an Australian cockatoo, the galah *Eolophus roseicapillus*. *Animal Behaviour*, *77*, 1019-1026.

van der Schaar, M., Delory, E., Catala, A., & Andre, M. (2007). Neural network-based sperm whale click classification. *Journal of the Marine Biological Association of the UK*, *87*, 35-38.

Schusterman, R. J., & Gisiner, R. (1988). Artificial language comprehension in dolphins and sea lions - the essential cognitive skills. *Psychological Record*, *38*, 311-348.

Shannon, C. (1948). A mathematical theory of communication, I and II. *Bell System Technical Journal*, *27*, 379-423, 623-656.

Skowronski, M. D., & Harris, J. G. (2006). Acoustic detection and classification of microchiroptera using machine learning: Lessons learned from automatic speech recognition. *The Journal of the Acoustical Society of America*, *119*, 1817-1833.

Smith, J. N., Goldizen, A. W., Dunlop, R. A., & Noad, M. J. (2008). Songs of male humpback whales, *Megaptera novaeangliae*, are involved in intersexual interactions. *Animal Behavior, 76*, 467-477.

Sockman, K. W., Salvante, K. G., Racke, D. M., Campbell, C. R., & Whitman, B. A. (2009). Song competition changes the brain and behavior of a male songbird. *Journal of Experimental Biology*, *212*, 2411-2418. doi: 10.1242/jeb.028456

Suzuki, R., Buck, J. R., & Tyack, P. L. (2005). The use of Zipf's law in animal communication analysis. *Animal Behaviour*, *69*, F9-F17. doi: 10.1016/j.anbehav.2004.08.004

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

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236-244.

Weisman, R. G., & Ratcliffe, L. (2004). Relative pitch and the song of black-capped chickadees. *American Scientist*, *92*, 532-539.

Weissengruber, G. E., Forstenpointner, G., Peters, G., Kubber-Heiss, A., & Fitch, W. T. (2002). Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis sylvestris f. catus*). *Journal of Anatomy (London)*, *2001*, 195-209.

Williams, S. L., Brakke, K. E., & Savage-Rumbaugh, E. (1997). Comprehension skills of language-competent and nonlanguage-competent apes. *Language & Communication*, *17*, 301-317. doi: 10.1016/S0271-5309(97)00012-8

Winn, H. E., Thompson, T. J., Cummings, W. C., Hain, J., Hudnall, J., Hays, H.,...Steiner, W. W. (1981). Song of the humpback whale - Population comparisons. *Behavioral Ecology and Sociobiology, 8*, 41-46.

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

Yan, X., Li, X., & Song, D. (2005). A correlation analysis on LSA and HAL semantic space models. *Lecture Notes in Computer Science, 3314,* 711-717.

Yurk, H., Barrett-Lennard, L. G., Ford, J. K. B., & Matkin, C. O. (2002). Cultural transmission within maternal lineages: Vocal clans in resident killer whales in southern Alaska. *Animal Behaviour*, *63*, 1103-1119.

Zuberbühler, K., Cheney, D. L., & Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology*, *113*, 33-42. doi:10.1037/0735-7036.113.1.33