A Categorization of Mexican Free-Tailed Bat (*Tadarida brasiliensis*) Chirps

Gregory Backus

August 20, 2010

Abstract

Male Mexican Free-tailed Bats (*Tadarida brasiliensis*) attract mates and defend territory using multi-phrase songs that have a structured set of rules. A subjective view of their spectrograms shows similarity and dissimilarity between the chirps (a syllable within the song) of different males. We developed a rigorous algorithm to characterize the shapes of these chirps. The discrete Fourier transform allowed us to focus on frequency information while a four level Daubechies 2 wavelet decomposition allowed us to focus on time. For comparison, we compressed large data vectors into a single data point using multidimensional scaling. Segmenting chirps, to further emphasize ranges of time and frequency, gave categorizations that most closely resemble the subjective groupings.

1 Introduction

Even though it is common throughout a large number of animal taxa to use an acoustic display for both mate attraction and territory defense [5], most terestrial mammals instead use visual and olfactory displays in sexual selection [2]. Thus, the exceptions found in bats [3] should be of particular interest to ethologists. In vocal animal communication, one differentiates a song from other acoustic signals if it consists of complex, hierarchical, individual units [4]. Recent evidence from Bohn et al.[1] showed that the song of the Brazilian free-tailed bat (*Tadarida brasiliensis*) did, in fact, have stereotyped and hierarchically organized phrases which follow a set of rules.

The study by Bohn et al [1] subjectively found three separate, recognizable phrase types in the songs of *T. brasiliensis*. The longest phrase type, usually found at the beginning of songs, was referred to as a chirp and characterized by its repetition of separate type "A" and "B" syllables. Secondly, the trill, often in the middle of a song, was a phrase type characterized by short, discrete bursts. Lastly, buzzes, while similar to trills, were significantly longer with more repeated syllables and were often found at the end of songs. A preliminary observation of their spectrograms found the type "B" syllable from the chirps of different bats was found to be one of the most distinctive syllables. In order to understand why these syllables fall into different categories and what their significance is, it would be beneficial to have an algorithm for categorizing the chirps used in each song of each bat.

2 Background

2.1 Energy

In signal processing, calculating the energy of the signal can compress a large amount of information into one point of data. This is done with the equation,

$$E = \frac{1}{n^2} \sum_{i=1}^{n} |y_i|^2,$$

where each sample in the signal is represented with y_i and the total signal length is n. To clarify, the term "energy" used in signal processing should not be confused with the same term used in physics. While the two may be related, we only use energy as a method of information compression between similarly manipulated signals and, thus, its physical characteristics are not relevant to this study.

2.2 Fourier Analysis

Fourier analysis is one of the most common mathematical techniques used in signal processing. The Fourier transform is an equation that converts a function, usually represented in the time domain, into one that is represented, instead, in the frequency domain. However, because digital recording equipment cannot take a continuous sample of an audio signal, it must be converted using a specific Fourier transform known as the discrete Fourier transform. This is done with the equation,

$$\hat{y}_k = \sum_{j=0}^{n-1} y_j e^{(-2\pi ijk)/n}$$

where $y = \{y_j\}_{j=-\infty}^{\infty}$ is a n-periodic sequence of complex numbers (the original signal), and $\hat{y} = \{\hat{y}_k\}_{k=0}^{n-1}$ is the transformed signal sequence. Because y is n-periodic, $y_j = y_{j+n}$ for any integer j, and thus, we are only concerned with the subsequence $\{y_0, y_1, ..., y_{n-1}\}$. Therefore, a non-periodic audio recording can be represented as a discrete, finite sequence, as only one period is needed for the discrete Fourier analysis.

Because n^2 calculations are required for the discrete Fourier transform, there is an algorithm used to reduce the number of calculations involved to $5n\log_2 n$, known as the fast Fourier transform. However, this algorithm can only be used when n is even and, because it can be iterated, it is most efficient when $n = 2^m$ for some positive integer m. If we let $n = 2N$, then the discrete Fourier transform can be separated into the sum of the even indices plus the sum of the odd indices, or

$$\hat{y}_k = \sum_{j=0}^{n-1} y_j e^{(-2\pi ijk)/n} = \sum_{j=0}^{N-1} y_{2j} e^{2j(-2\pi ik)/n} + \sum_{j=0}^{N-1} y_{2j+1} e^{(2j+1)(-2\pi ik)/n}$$

and, therefore,

$$\hat{y}_k = \sum_{j=0}^{N-1} y_{2j} e^{2j(-2\pi ik)/n} + e^{(-2\pi ik)/n} \Big(\sum_{j=0}^{N-1} y_{2j+1} e^{(2j)(-2\pi ik)/n} \Big).$$

Because 2^m nodes will be rarely be recorded exactly, the signal can become 2^m-periodic by setting $y_q = 0$ for all integers $n \leq q < 2^m$ for the least value of m where $2^m > 0$.

Through this transformation, it is possible to quantify the frequencies that most strongly characterize a bat's chirp. Using the Matlab command `fft()`, we can take a signal vector as an input and give the transformed vector of complex numbers as an output. Thus, the magnitude of each element in the vector must be taken for proper visualization purposes. Figure 1 shows the recorded audio signal of a type "B" chirp and its transformed counterpart.

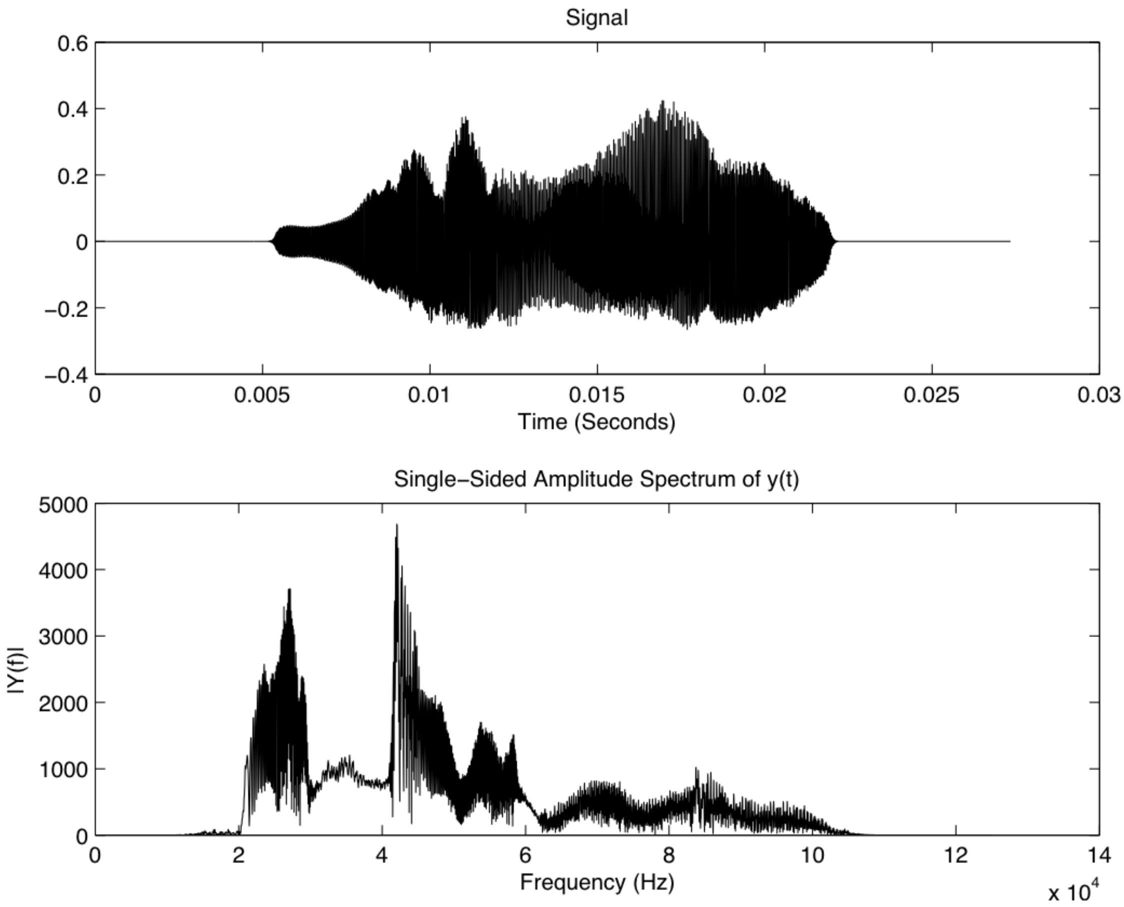

Figure 1: *The original signal from one type "B" chirp and its fast Fourier transform.*

2.3 Wavelet Analysis

While frequency data is generally very important in signal processing, it is often also critical to consider the time at which these frequencies occur. Because the discrete Fourier transform gives the signal in only the frequency domain, we lose the time data that was present in the original signal. However, wavelet analysis, can be used when both time and frequency data are relevant. Through the wavelet transform, a signal can be decomposed into multiple sequences of coefficients, separated by frequency ranges and still preserving the time domain.

Two functions, the scaling function ϕ and the wavelet function ψ, are used to generate a family of functions that can express a deconstructed signal. After applying a filtering algorithm to a signal f of length N, it can be represented with two sequences of coefficients. One sequence, the first level detail coefficients (d_1), is of length $N/2$ and represents the highest octave of frequencies present in the signal, while the other, the first level approximation coefficients (a_1), is of length $N/2$ and represents all the lower frequencies remaining in the signal. Afterward, we find that $f = d_1 + a_1$, and very little, if any, data is lost in this transform. Conveniently, this process can be iterated on the approximation coefficient to give us data on the next highest octave with the sequence d_2, and the frequencies below it with the sequence a_2. This can be iterated as long as the length of the approximation coefficients can be divided. Thus, after a decomposition, $f = \sum_{j=1}^{k} d_j + a_k$.

Because a wavelet often needs to have certain properties or shapes, there are many varieties of wavelets from which we can choose. We focus on Daubechies wavelets, a family of orthogonal wavelets, and more specificically, Daubechies 2 wavelet, as this is both continuous and can detect the sharper spikes of discrete data. Figure 2 shows a graphical representation of bat chirp and its four level Daubechies 2 decomposition.

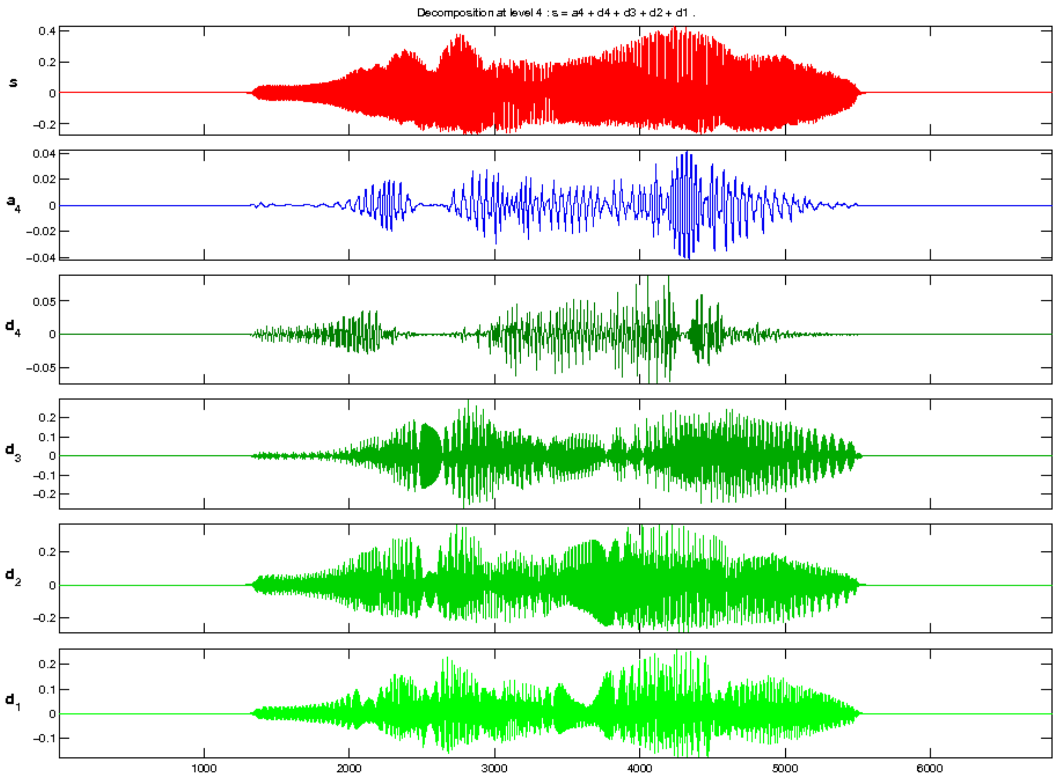

Figure 2: *The original signal of a bat chirp, its level 4 approximation, and its first 4 sequences of detail coefficients.*

2.4 Spectrogram Analysis

A specific extension of the discrete Fourier transform that also considers both time and frequency is known as the short-time Fourier transform. Using a windowing function, $w(m)$, the transform is defined as

$$X(n, \theta) = \sum_{j=-\infty}^{\infty} x[j]w[n-j]e^{-\theta ji}$$

where $n \in \mathbb{N}$ is the window size, and $\theta \in \mathbb{R}$ is frequency. The output of this algorithm is a discrete Fourier transform of several smaller time segments throughout the signal. The magnitude of each of discrete Fourier transform can be plotted vertically at each point in time, giving a three dimensional visualization called a spectrogram. Using a range of colors, it is possible to represent spectrograms on a two dimensional graph like shown in Figure 3.

Unfortunately, the short-time Fourier transform outputs large amounts of information, making it difficult to categorize. While we may be interested in the overall "shape" of these spectrograms, they are not conveniently made for comparison. Thus, other techniques, such as the Fourier transform and Wavelet analysis are preferred in this study.

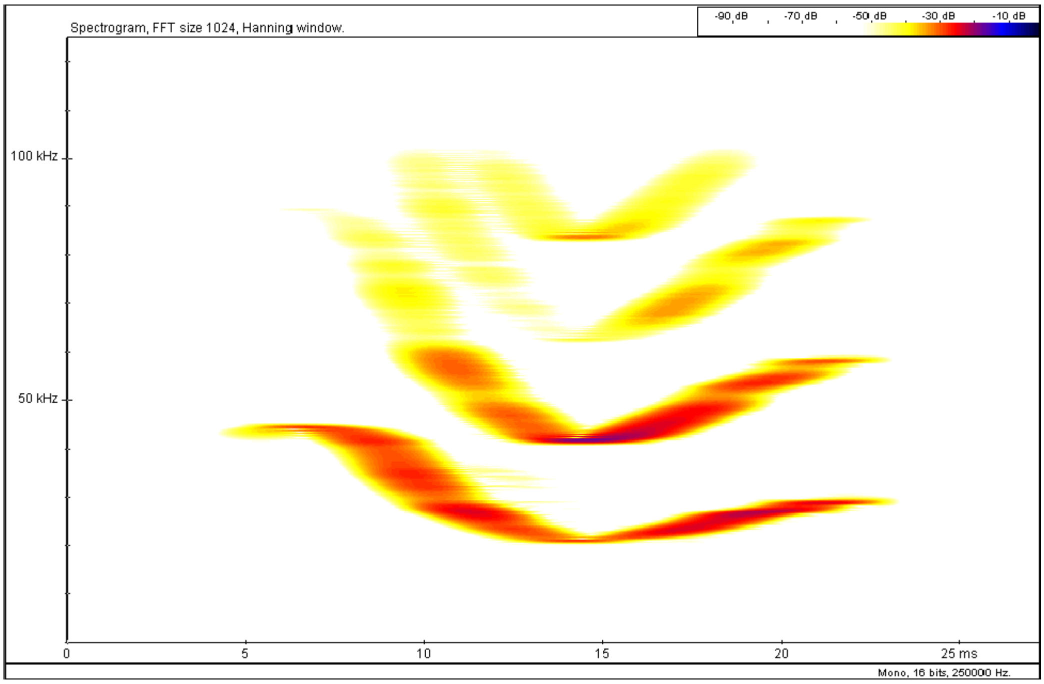

Figure 3: *Spectrogram of of a bat chirp.*

2.5 Multidimensional Scaling

In order to make a comparison within and between the different bat chirps, a data compression method can be used to visualize the similarity or dissimilarity between n objects in N dimensions. Multidimensional scaling represents a set of N-dimensional distances with much lower dimensional distances, or $d_{ij} \cong \delta_{ij}$, where d_{ij} is the estimated distance and δ_{ij} is the actual N dimensional distance, for all objects i, j. The N-dimensional distance between two objects i and j can be calculated using the norm,

$$\delta_{ij} = [(y_i - y_j)'(y_i - y_j)]^{1/2} \tag{1}$$

where y_i and y_j are column vectors and $(y_i - y_j)'$ is the transpose of $(y_i - y_j)$. This can be used to find the distance between all objects in the distance matrix,

$$D = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,I} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,I} \\ \vdots & \vdots & & \vdots \\ \delta_{I,1} & \delta_{I,2} & \dots & \delta_{I,I} \end{pmatrix}.$$

Assuming that we want to express the distance in k dimensions, this square matrix is used to calculate another square matrix B, who's two greatest eigenvalues and corresponding eigenvectors are used to create a $k \times n$ matrix that approximates the distances in k dimensions. When this is plotted, assuming that the k greatest eigenvalues represent a large portion of the information, the similarity between n objects can be visualized in k dimensions.

3 Data

3.1 Study Animals

The bats used in this study were either collected from a colony in College Station, Texas or from one in Austin, Texas and were kept in a vivarium at Texas A&M in accordance with the NIH guidelines for experiments involving vertebrate animals. Full songs were recorded within this vivarium for a previous study, and it was from these which we obtained the 25 “type B” chirps we analyzed for each bat. In the initial stage of our study, we used only the chirps from a total of 29 different Brazilian Free-tailed Bats (three of which were recorded in two sets over two years and one of which was recorded in three sets over three years). In the the second stage, we used 10 “type B” chirps from an additional four bats.

3.2 Preparatory Editing

In SIGNAL 4.04, we used a gate detection algorithm to determine the beginning and end of these syllables. After cutting these to the proper length, we tapered both the beginning and the end of the signal so that these syllables could be properly isolated. Next, we added ten milliseconds of zeros on both the beginning and end of the signal. Because the sample rates varied between a number of signals, we resampled each to 250 kilohertz (kHz) in SIGNAL 4.04. We then applied a lowpass filter at 80 kHz and a highpass filter at 1 kHz below the fundamental frequency. To ensure that later comparisons weren't biased by signal length, we normalized them so that the root square mean (RMS) for each was one volt. Lastly, each signal was exported as a Waveform Audio File Format (".wav") for later comparison.

4 Strategies

4.1 Pure Signal

Because each signal initially contained a large, varying number of data points, we applied a segmenting technique in order to compress each signal into as little data as possible. This was beneficial because not only did it give us less data, but it also gave us an equal number of data points in each signal to compare. Ideally, by determining which time segments have the greatest amount of energy, we should be able to have some numerical estimate of the signal's shape.

To do this, we used Matlab to cut each signal into ten segments of equal length. After calculating the energy in each of the 10 segments, we exported a 10-dimensional vector of energies for each chirp. In Stata 11.1, we compiled all chirp vectors into one large matrix and performed a multidimensional scaling.

Taking only the first dimension, we attempted then to estimate the shape of each signal in a single data point. Comparing this one dimensional representation between all chirps, the effectiveness of this segmenting technique could be assessed. While our comparisons were surprisingly accurate for such simple segmenting, many signals that do not appear to be similar in subjective categorizations did appear to be similar according to this estimation (Figure 4). Therefore, this method was not satisfactory.

4.2 Fast Fourier Transform

Since the inaccuracy in the first strategy might have come from the fact that the signals were only divided into time segments, properly categorizing the chirps might require segmenting by frequency as well. Thus, after dividing the signal into ten equal segments, we further divided these segments into fifteen frequency segments. This was done by first using the fast Fourier transform function in Matlab on each segment. Then, in the frequency domain, it was possible to separate the segments into 10 kHz ranges. Calculating the energy in each of these 150 segments per chirp, we exported 150-dimension vectors.

After all vectors were compiled in Stata, we performed a multidimensional scaling to compress the signals into one dimension. When these were plotted, the desired subjective categories were even more even less evident (Figure 5). This is likely because either the

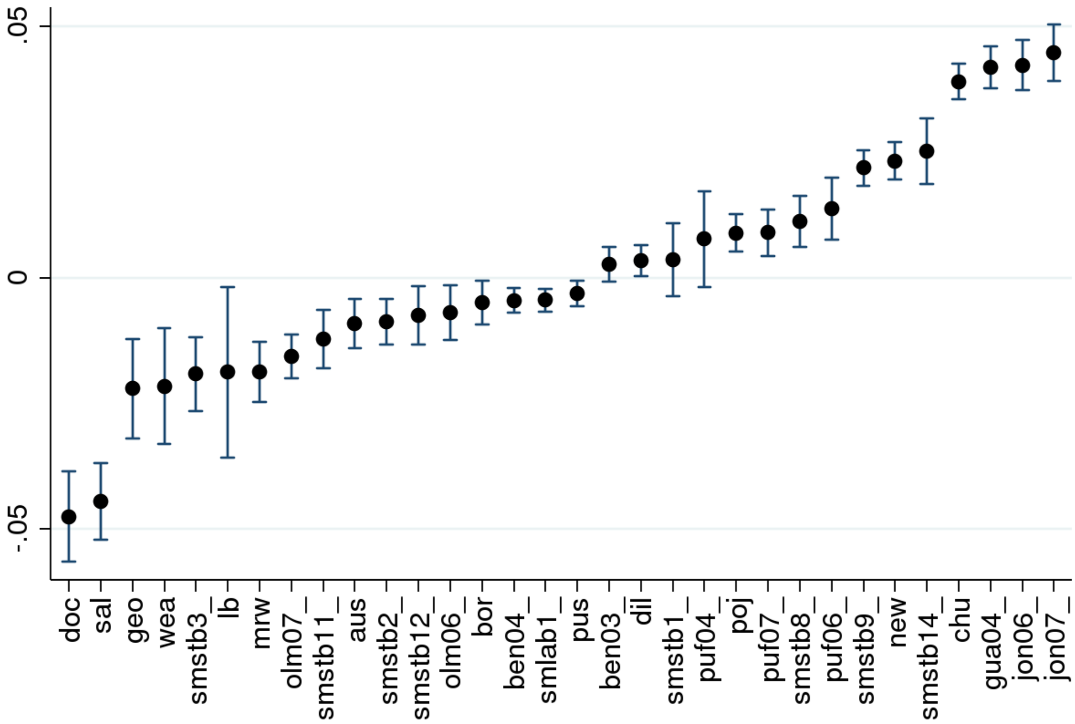

Figure 4: *Each "type B" chirp was divided into ten equally sized segments. A multidimensional scaling was performed on the calculated energies in each segment. The average values and standard errors for each bat are shown.*

10 kHz frequency segments were too thin, or the lowest and most of the highest frequency segments contained little energy information.

4.3 Wavelet Decomposition

Wavelet decomposition is another signal processing technique that can divide the chirps into frequency segments. In this strategy, after we divided a signal into ten segments of equal length, we performed a four level Daubechies 2 wavelet decomposition on each in Matlab. This divided each tenth segment into four octave sequences of detail coefficients and a sequence of fourth level approximation coefficients. Once we calculated the energy in each of these segments, we exported a 50-dimensional vector for each chirp.

After compiling all vectors into Stata, we performed a multidimensional scaling on all the data. Plotting the first dimension, we found that, while categories weren't necessarily distinct from one another, they did appear to be organized in a manner that reflects the shape of their spectrograms (Figure 6). This first dimension allowed us to categorize these signals into four different shapes that were observed in their spectrograms.

The success in this strategy most likely lies in the fact that it divides the signal by octave. In the chirps of these bats, more unique information is found in lower frequencies. Thus, because wavelet decomposition divides each signal by octaves, which occur more often at lower frequencies, our method using wavelet decomposition is able to focus on the most important information. If we were able to determine which frequencies are actually the most important in these chirps, we might be able to widen the gap between these categories.

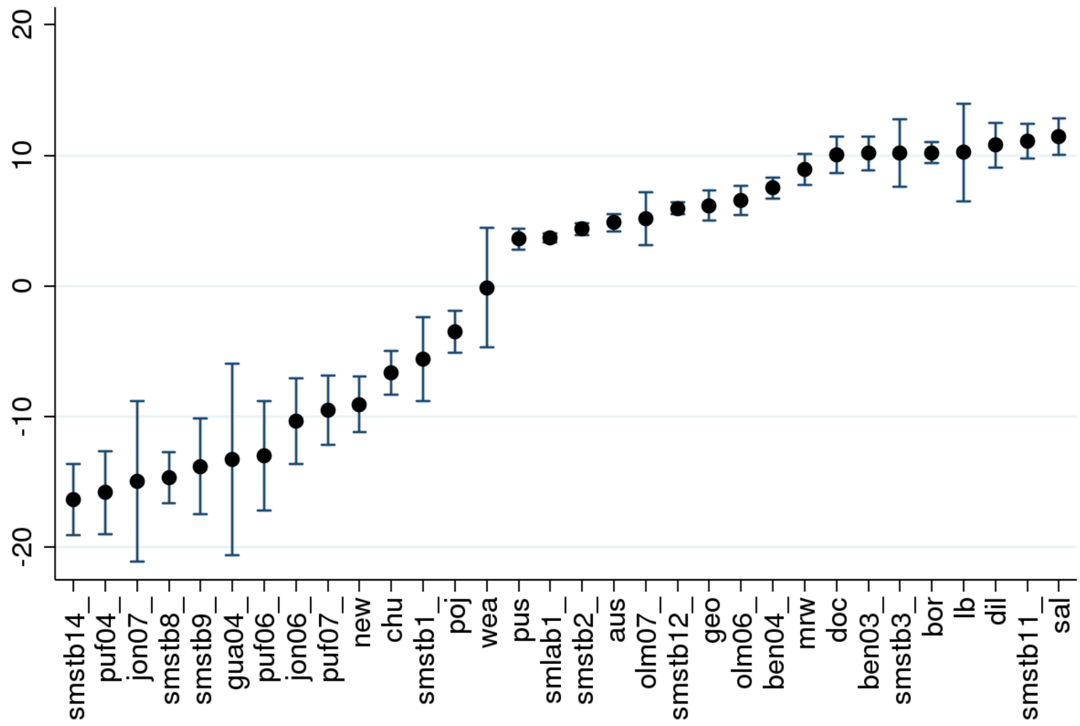

Figure 5: *Each "type B" chirp was divided into ten equally sized segments. A fast Fourier transform was used to divide each of these into 15 frequency segments. Next, a multidimensional scaling was performed on the calculated energies in each segment. The average values and standard errors for each bat are shown.*

References

[1] Kirsten M. Bohn, Barbara Schmidt-French, Christine Schwartz, Michael Smotherman, and George D. Pollak. Versatility and stereotypy of free-tailed bat songs. *PLoS ONE*, 4(8):e6746, 08 2009.

[2] J. W. Bradbury and Sandra Lee Vehrencamp. *Principles of Animal Communication*. Sineauer, Massachusetts, 1998.

[3] Susan M. Davidson and Gerald S. Wilkinson. Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Animal Behaviour*, 67(5):883–891, 2004.

[4] Peter Marler and Hans Willem Slabbekoorn. *Nature's Music: The Science of Birdsong*. Academic Press, London, 2004.

[5] Andrea Megela Simmons, Arthur N. Popper, and Richard R. Fay. *Acoustic Communication*. Springer, New York, 2003.

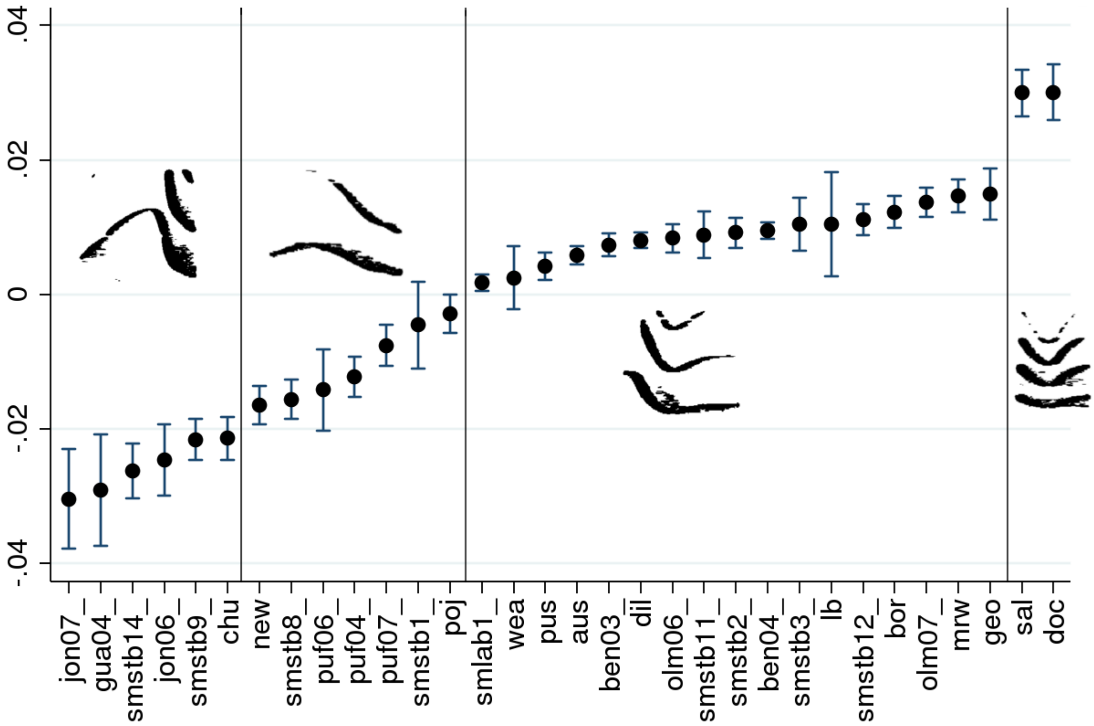

Figure 6: *Each "type B" chirp was divided into ten equally sized segments. A wavelet decomposition was used to divide each of these into four detail segments and one approximation segment. Next, a multidimensional scaling was performed on the calculated energies in each segment. The average values and standard errors for each bat are shown.*