# Spectral Coherence Analysis of Killer Whale (*Orcinus orca)* Vocalizations

Miloslav Nosal[1], Eric M. Nosal[2], and Eva – Marie Nosal[3]
[1,2]*Department of Mathematics and Statistics and Department of Computer Science,*
*University of Calgary, 2500 University Drive, Calgary, AB, Canada, T2N 1N4 and STATSCON, Stat. Consulting*
[3]*SOEST, Department of Geology and Geophysics, University of Hawaii, 1680 East-West Road,*
*POST 813, Honolulu, HI 96822, USA*

nosal@ucalgary.ca

*Keywords:* Killer whale, orcinus orca, clan, pod, matriline, pulsed call, tonal call, orca dialect, digital signal analysis, spectral density function, power spectrum, spectrogram, coherence function, harmonic overtones, whistle,

**Abstract**

In this paper, we investigate individual characteristic dialects of various clans of resident and transient killer whales frequenting coastal waters of Vancouver Island, British Columbia. Digital signal processing methods of *spectral* analysis and *coherence* analysis are used to analyze clan specific digitally recorded orca vocalizations. These methods are used to characterize dialects of various clans in terms of frequency domain analysis. It is demonstrated that the sequence of harmonic overtones in orca dominant tonal whistles forms a clear sequence of overtones same as those found in classical acoustic musical instruments and can be identified with standard musical notes. Furthermore it is shown that these spectral methods can be used as a characterization and identification of clan-specific orca dialects.

## 1. Introduction

Underwater acoustic vocalization of killer whales (*Orcinus orca)* encountered in the waters off the coast of Vancouver Island, British Columbia, consists of a rich repertoire of pulsed and tonal calls. A majority of these calls are used for the purpose of interaction and communication among members of a group during various activities such as hunting, feeding, socializing, relaxing etc. It has been established in the literature that orcas associate themselves into family groups, called *pods*, and that several related pods create a larger, genetically similar group of individuals, called a *clan*. Furthermore, it has been established that all individuals within a clan emit similar calls, learned by an offspring from its mother and other family members, but that different clans use different sets of sounds. The BBC(2000) series *Predators* documents a mother killer whale teaching its calf sophisticated hunting techniques using contrived vocal communications. The learning process takes many years and co-operative hunting maneuvers coordinated by sophisticated vocalizations continue for many decades. It has been hypothesized that familial groups of killer whales develop highly specific sets of sounds. These sets of sounds, referred to as *dialects,* are uniquely characteristic for each clan (Ford 1983, 1989, 1991, 2000, Thomsen 2001).

A great majority of sounds produced by a pod of killer whales when hunting, feeding, socializing, resting etc. can be clearly classified into two distinct categories. Fast repeated pulsed sounds or clicks of a percussive character with a quick attack and a short delay are referred to as *echolocation*. These sounds are used primarily for underwater location purposes in a similar way as those used in sonar technology. Another group of sounds has a clearly recognizable tonal character similar to a sound of a *whistle*. These sounds have clearly identifiable musical tonal properties such as a pitch and a tonal color, which can be recognized by a musically trained human ear. Among whistles is a subset of dominant, very loud sounds which are very frequently and regularly repeated during pod's activities. These sounds, which we will call *dominant tonal whistles*, have very specific characteristics. They have a very clear pitch definition in terms of *harmonic overtones (HO)* and a clear tonal color defined by the relative amplitudes of these harmonic overtones, just like sounds produced by regular acoustic instruments. These whistles have a relatively constant pitch over a certain period of time which can be clearly defined within the standard musical notation system such as C#7, F#7 or G#7 etc. The starting pitch may be eventually modulated or slide up or down in a distinct manner. In this paper we concentrate mainly on spectral analysis of these dominant tonal whistles.

Some results have been published trying to analyze orca vocalizations using digital signal analysis methods such as *power spectrum* and *spectrogram* (Ford 1983, 1989, 1991, Thomsen 2001). In this paper we investigate further extension of these methods and use of *spectral coherency* in order to identify and characterize specific dialects of various clans of killer whales frequenting the coastal waters of Vancouver Island, British Columbia. Using detailed numerical analysis of recorded killer whale vocalizations combined with graphical methods and musical concepts we provide new methodology for animal vocal communication analysis.

## 2. Materials and methods

There are many pods and clans of killer whales frequenting protected waters of Johnstone Straight and Georgia Straight between Vancouver Island and mainland British Columbia. Two most distinct groups are the *resident* and *transient* whales. Residents spend in these waters great part of the year and feed mainly on various migrating salmon species. The transients are predators roaming vast expanses of the ocean and feeding primarily on marine mammals. There is a fairly detailed description of these killer whale groups and their classification in the available literature (Ford 1983, 1989, 1991, 2000, Thomsen 2001). Some killer whale sounds are also available on CDs (Ford 1992, Spong 2001).

In this paper, we investigate individual dialects of clans A, G, R and J of Vancouver Island resident killer whales using *spectral* analysis and *coherence* analysis. These analyses are based on actual vocalizations of these clans recorded live in Johnstone Straight and Georgia Straight. Recordings were made using underwater hydrophones and SONY and converted to standard digital *WAV* file formats with 16 bit depth and 44.1 kHz sampling rate. We are primarily using orca vocalizations from the CD track 1 labeled *Communication and Dialects of Killer Whales* (Ford 1992) in addition to own digital recordings. This CD, titled *Blackfish Sound, Underwater Communication of Killer Whales,* is available from the Vancouver Aquarium Marine Science Centre. The first track of this CD was imported into *Sound Forge Version 7.0.* Using a studio audio monitoring system, dominant tonal whistles of individual clans were tentatively identified by a musically trained ear (endowed with perfect pitch) and labeled with respect to the time-location on the track. For example a sound slip frequently used in this paper is labeled A2-06-037, meaning that it is located 2 minutes and 6.037 seconds from the beginning of this track, and relates to clan A. Preliminary Sound Forge spectral analysis demonstrated that no relevant audio information below 1 kHz frequency is imbedded in the available signals. For this reason, all sound clips were filtered to remove any background noise below 1 kHz and normalized to – 1 dB.

From the available sound clips of each clan, representative shorter sound bits were selected and for each of them a fast Fourier transform, spectral function, power spectral density and spectrogram were computed. For a preliminary assessment, corresponding tools from the Sound Forge Version 7 libraries were utilized. However, despite the considerable power of Sound Forge spectral analysis tools, these methods were not sufficiently powerful and sublime to capture real nature of orca sounds. The final mathematical and statistical analyses were performed using *Signal Processing Toolbox* available within the *MATLAB Version 12.* The corresponding MATLAB program which was used to perform the required analyses is attached below. Additional analyses were performad using *SONAR Version 2.2.* While the spectrogram and power spectrum density allow a characterization of an individual killer whale call in terms of its frequency domain, they do not allow a rigorous quantitative comparisons between various calls or sets of calls (dialects). For this reason, coherence functions between selected calls were computed. These functions allow comparisons of various calls in terms of frequency content and therefore allow one to establish similarity or dissimilarity between calls based on rigorous quantitative procedures. Using numerical evaluation of spectral and coherence functions and their high resolution graphical interpretations, identification and characterization of dialects of resident clans A, G, R, and J of investigated killer whales were performed. Additional coherence analyses compared calls of resident clans to transient killer whales in the same general marine environment.

## 3. Spectral coherence analysis

The goal of spectral estimation is to describe the distribution (over the entire frequency range ) of the acoustic power contained in a digital signal, based on a finite set of data. Estimation of power spectra is useful in a variety of applications, including the detection of signals buried in wide-band noise. The power spectrum of a stationary random process $x$ is mathematically related to the correlation sequence by the discrete-time Fourier transform. In terms of normalized frequency as follows:

$$S_{xx}(\omega) = \sum_{m=-\infty}^{\infty} R_{xx}(m) \exp\{-i\omega m\}$$

where *i* is the imaginary unit and the correlation function $R_{xx}(m)$ is a one-variable analogue of the true cross-correlation function defined for any two stationary random processes *x* and *y* as

$$R_{xy}(m) = E[x_{n+m} y_n]$$

If the time series *x* represents series of samples from a digitally sampled acoustic sound with sampling frequency $f_s$, then we can use a physical frequency $f$ (in Hertz) and the relation

$$\omega = \frac{2\pi f}{f_s}$$

Using the physical frequency, we can write the power spectrum as

$$S_{xx}(f) = \sum_{m=-\infty}^{\infty} R_{xx}(m) \exp\{-2\pi i f m / f_s\}$$

The average power of the digital signal *x* (represented by a numerical sequence x) over the entire Nyquist interval is given by

$$R_{xx}(0) = \int_{-\pi}^{\pi} \frac{S_{xx}(\omega)}{2\pi} d\omega = \int_{-f_s/2}^{f_s/2} \frac{S_{xx}(f)}{f_s} df$$

The intergrands from the above expressions

$$P_{xx}(\omega) = \frac{S_{xx}(\omega)}{2\pi}$$

and

$$P_{xx}(f) = \frac{S_{xx}(f)}{f_s}$$

are defined as the power *spectral density* (PSD) of the stationary digital random signal $x=x_n$.
The average power of a signal over a particular frequency band, $[\omega_1, \omega_2]$, $0 \le \omega_1 < \omega_2 \le \pi$, can be found by integrating the PSD over that band.

$$\overline{P}_{xx}[\omega_1, \omega_2] = \int_{\omega_1}^{\omega_2} P_{xx}(\omega) d\omega + \int_{-\omega_2}^{-\omega_1} P_{xx}(\omega) d\omega$$

$P_{xx}(f)$ represents the power content of a signal in an infinitesimally narrow frequency band centered at *f*, which is why we call it *the power spectral density*. The units of the PSD are power (e.g., watts) per unit of frequency.

The PSD is a special case of the *cross spectral density* (CSD) function, defined between two signals *x* and *y* as

$$S_{xy}(\omega) = \sum_{m=-\infty}^{\infty} R_{xy}(m) \exp\{-i\omega m\}$$

The magnitude-squared ***coherence function*** between two digital signals related to a frequency *f* is given by

$$C_{xy}(f) = \frac{\left|S_{xy}(f)\right|^2}{S_{xx}(f)S_{yy}(f)}$$

This ratio is a real number between 0 and 1 that measures the correlation between digital signals $x = x_n$ and $y = y_n$ at the frequency *f*. If the signals have a similar behavior at a given frequency *f*, then the coherence at that frequency will be large and/or close to one. If the signals have very dissimilar properties at a frequency f, then the coherence will be zero or close to zero. The coherence is a function of the the power spectra of *x* and *y* and cross-spectrum of *x* and *y* as follows:

$$C_{xy}(f) = \frac{\left|P_{xy}(f)\right|^2}{P_{xx}(f)P_{yy}(f)}$$

MATLAB function cohere has been used to compute the coherence function in all numerical examples.

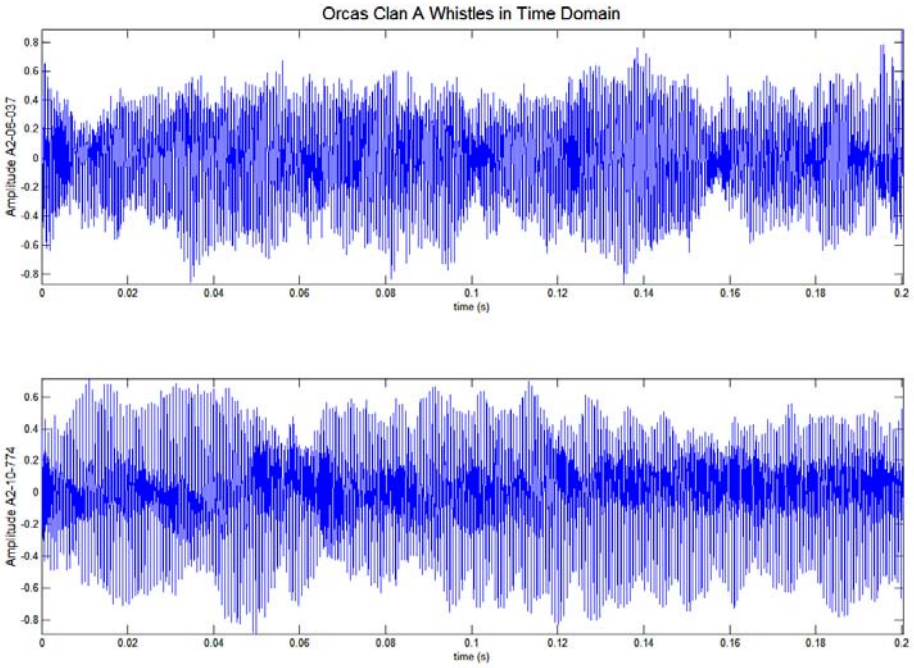
Using the coherence function, we can compare two digital signals and see whether they are similar or not. This is particularly true about acoustic musical sounds which are in a digital format. Musical sounds are often composed of several sinusoidal waveforms with clearly defined frequencies. The lowest frequency, which usually carries the highest energy and power, is referred to as the *fundamental tone.* The fundamental frequency defines the pitch of the tone. The higher frequencies are referred to as *harmonic overtones (aliquot tones)* and their frequencies are multiples of the fundamental frequency by natural numbers n = 1 (fundamental), 2 (octave), 3 (fifth above octave), 4 (second octave) 5 (third above the second octave) etc. Harmonic overtones define the tonal quality of the sound or color of the sound. Thus a violin and a clarinet playing the same note will have the same fundamental note frequency but different composition of the harmonic overtones. Two musical tones of the same pitch will have the same fundamental frequency. Two musical tones of the same sound color will have the same harmonic overtones. Spectral density and a coherence functions are the prime tools for sound analysis. The spectral density function will have *peaks* at frequencies corresponding to harmonic overtones. If two tones are identical in pitch and color (or at least very similar) their spectral densities will have peaks at the same frequencies and of a similar height. Furthermore, the coherence function will have peaks which correspond to c-l0cated peaks of the spectral densities. These ideas are applied in section 4 for comparative analysis of clan A orca sounds.

**4. Results**

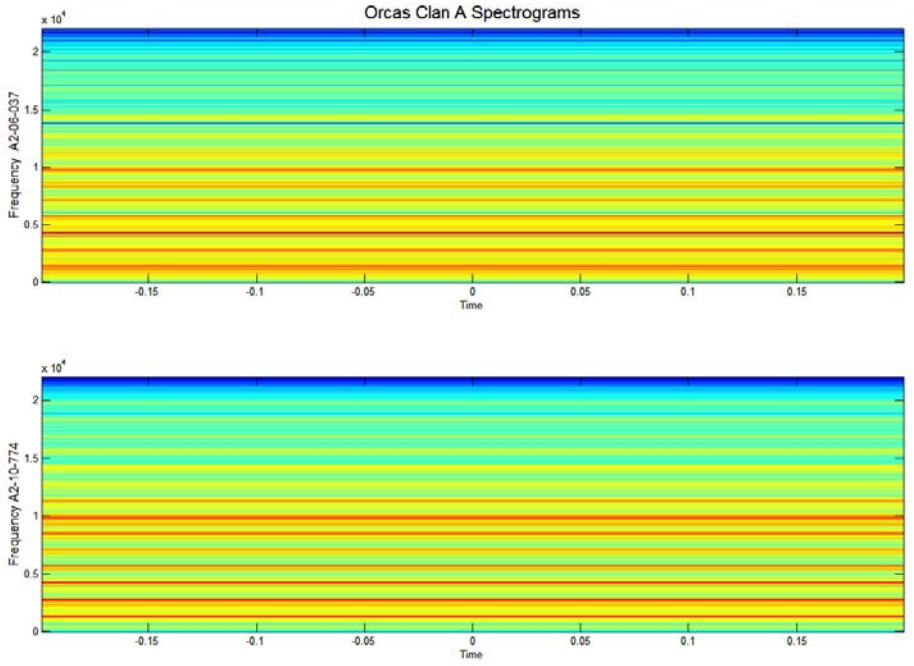
Digital sounds on track 1 of the CD titled *Blackfish Sound, Underwater Communication of Killer Whales*, track (titled *Communication and Dialects of Killer Whales* (Ford 1992)) represent examples of individual dialects of clans A, G, R and J of Vancouver Island resident killer whales. For the purpose of this analysis, sound clips of dominant tonal whistles of various clans were selected and analyzed using methods of the above section 3. The corresponding selected clips were as follows (see section 2 for the notation definition):

**Clan A digital clips:**
A3-06-037 A2-10774 A2-46-905 A7-52-572
**Clan G digital clips:**
G9-15-699 G9-20-622 G9-37-248
**Clan R digital clips:**
R10-11-427 R10-15-327 R10-35-112
**Clan T digital clips:**
T11-43-285 T11-46-629 T11-57-484

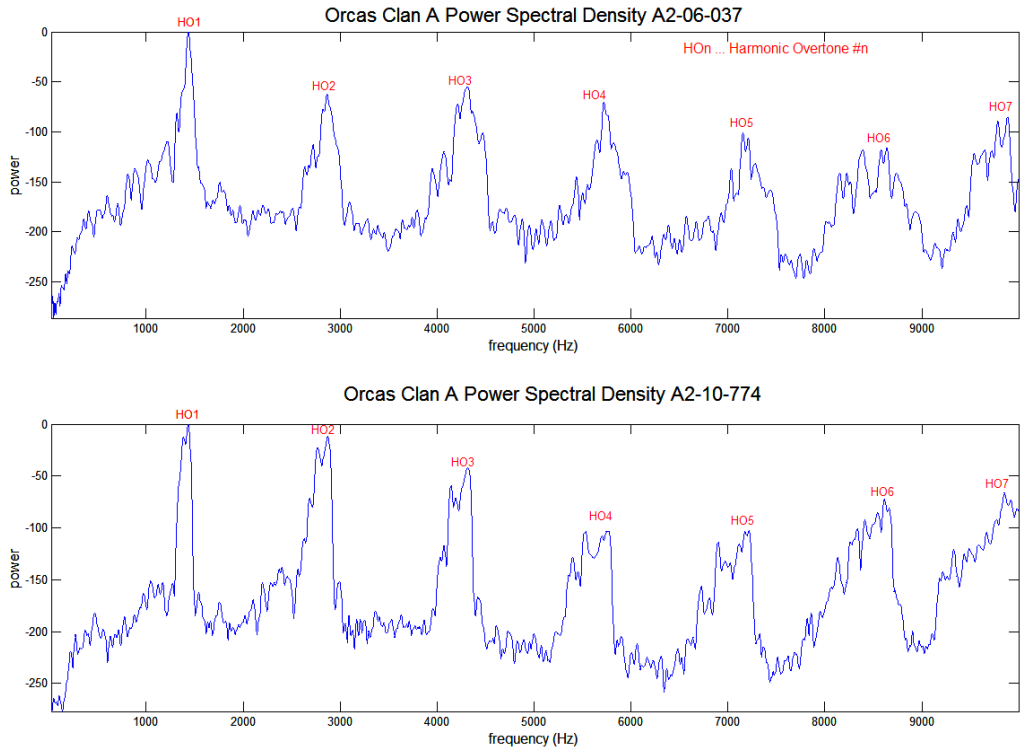
We will demonstrate the spectral and coherence methodology applications to killer whale dominant tonal whistles comparisons using two sound clips of the clan A, namely A3-06-037 and A2-10774. Aural comparisons of these short clips of these two whistles confirm a clear musical character, same pitch and very similar tonal coloration. Following are these two clips in the time domain representation:

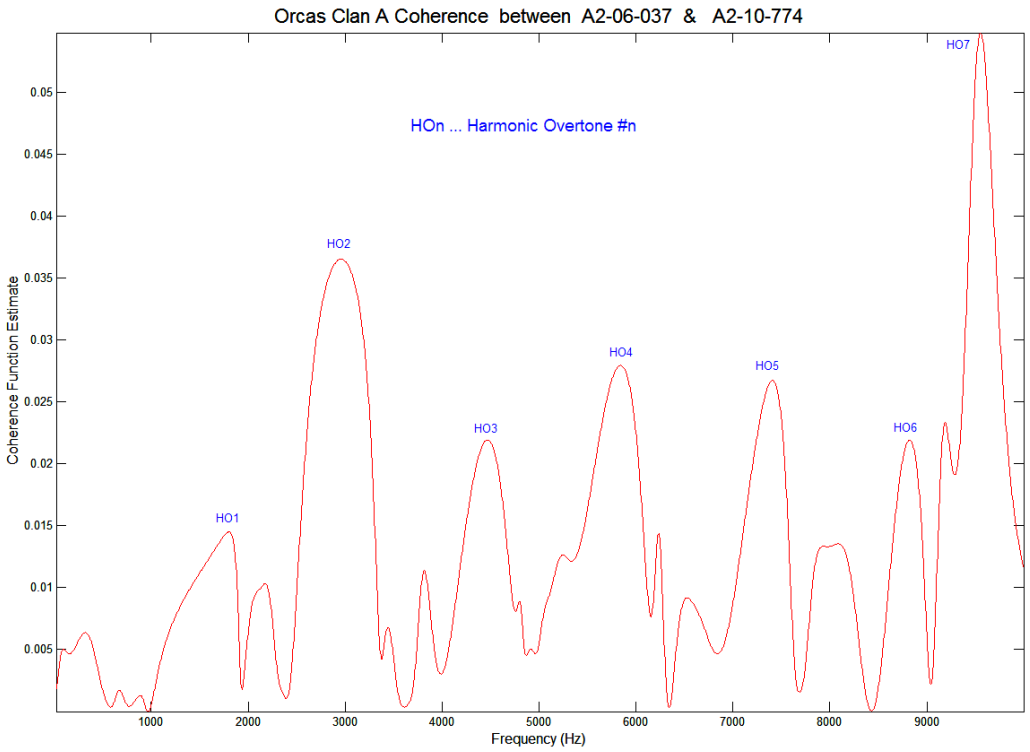



The above given time domain graphs of the whistles correspond to the following spectrograms.



Following are the spectral density functions and the coherence function of these two clan A orca calls:



## 5. Conclusions

From the graphs of the power spectral density functions we can clearly see that both calls have a clearly musical character with the fundamental frequency about 1440 Hz corresponding to the note F#7 and 6 harmonic overtones below 10 kHz. In the above graphs these harmonic overtones have been denoted as HO1, HO2, … HO7. Furthermore, we can see that all these frequencies are approximately the same for both clan A whistles and furthermore that they clearly coincide with the coherence function peaks. Following is a numerical table of these frequencies:

| Sound and function ID | Fundamental HO1 = F#7 | HO2 | HO3 | HO4 | HO5 | Ho6 | HO7 |
|---|---|---|---|---|---|---|---|
| Spectrum A2-06-337 | 1440 | 2868 | 4320 | 5720 | 7160 | 8635 | 9880 |
| Spectrum A2-10-774 | 1435 | 2875 | 4315 | 5722 | 7215 | 8615 | 9855 |
| Coherence of the above | 1790 | 2950 | 4460 | 5840 | 7400 | 8820 | 9555 |

Clan A typical dominant tonal whistles harmonic overtones (Hz)

- The sequence of harmonic overtones forms a clear sequence of aliquote tones found in classical acoustic musical instruments. Considering f = 1440 Hz (about F#7) as the fundamental frequency, the ideal sequence of harmonic overtones should be as follows: 1440 – 2880 – 4320 – 5760 – 7200 – 8640 – 10080. The analysis of actual killer whale whistles comes incredibly close to these ideal harmonic overtones and any deviations are apparently results of analytical errors.

- Coherence peaks come very close to peaks for the spectral densities. They seem to be systematically biased up. It can be seen that this fact is a consequence of smoothing a high level of random noise imbedded in these signals using Hanning window of length of 170 samples and overlap of 160 samples. Using different windows produces slightly different bias but the underlying relationship of the peaks remains the same.

- Similar results have been obtained when comparing other clips from the same clans. When comparing sounds from different clans, spectral peaks are always highly dissimilar and the coherence function does not produce any indication of similarity.

- The above points illustrate the power of spectral and coherence analysis when comparing dominant tonal whistles of killer whales from the same clan. It should be pointed out that each clan has several characteristic dominant tonal whistles and one has to make sure to compare corresponding whistles. Furthermore, this analysis can identify similarity or dissimilarity of characteristic clan calls of a constant pitch only. After the introductory part of such call is modulated, varied or slides in pitch these methods do not provide appropriate tools for comparisons. However, it is conjectured that individual clan dialects can be fully characterized by dominant tonal whistles of constant pitch. There has never been found a case of two different clans producing a dominant tonal call of the same pitch and color.

## 6. MATLAB program for spectral coherence analysis

```
function [power1, power2, f, coh, coh_mean, F, sound1, sound2] =
cohwork1(song1, song2, NFFT, f_cutoff, log_opt)

% This script imports two wave files. For each file, a plot of time vs.
amplitude is displayed.
% Each signal is transformed into the frequency domain. The spectrograms are
plotted.
% Each signal is then converted to a power spectrum and plotted.
% The power sepctra are compared using coherence (MATLAB function 'cohere').
%
```

```
% USAGE: [power1, power2, f, coh, coh_mean, F] = cohwork(song1, song2, NFFT,
f_cutoff, log_opt);
%
% OUTPUT:
% power1 = power spectra for first signal (vector of amplitudes)
% power2 = power spectra for second signal
% f = vector of frequencies for power spectra
% coh = coherence
% coh_mean = mean of coherence
% F = vector of frequencies for coh
%
% INPUT:
% sound1 = name of first wav file
% sound2 = name of second wav file
% NFFT = length of Fourier transform used for spectrograms, power spectra, &
coherence
%        (currently only default windows are used)
% f_cutoff = [fmin fmax]
%          = vector with lowest and highest frequencies to be considered
% log_opt = 1 (plots on a log scale) or 0 (plots on a linear scale)
%
% EXAMPLE:
% [power1, power2, f, coh, coh_mean, F, sound1, sound2]  = ...
%            cohwork1('A2_06_037begfil.wav', 'A2_10_774begfil.wav',8800, [20
10000],0);
%
% emn 03/10/04

[sound1, fs1, nbits1] = wavread(song1);
[sound2, fs2, nbits2] = wavread(song2);

% need the length to be the same for both signals and
% even (for when we truncate in the frequency domain)
leng = floor(min(length(sound1), length(sound2))/2)*2;

sound1 = sound1(1:leng);
sound2 = sound2(1:leng);


% ******************* TIME DOMAIN  PLOTS
**************************************

% assume that both sampling frequencies are the same (44100 Hz)
t = (1:leng)/fs1;

figure(1)

subplot(2,1,1)
plot(t, sound1);
axis([0 t(end) min(sound1) max(sound1)])
title('Time Domain')
xlabel('time (s)'); ylabel('amplitude');
subplot(2,1,2)
plot(t, sound2)
axis([0 t(end) min(sound2) max(sound2)])
xlabel('time (s)'); ylabel('amplitude');
```

```
% ******************** SPECTROGRAM PLOTS
**************************************

figure(2)

subplot(2,1,1)
specgram(sound1, NFFT, fs1);
title('Spectrograms')
subplot(2,1,2)
specgram(sound2, NFFT, fs2);

% *************** POWER SPECTRA COMPUTATIONS AND PLOTS
************************

[power1,f] = pwelch(sound1, [], [], NFFT, fs1);
[power2,f] = pwelch(sound2, [], [], NFFT, fs2);

fmin = max(find(f < f_cutoff(1)));
fmax = min(find(f > f_cutoff(2)));

% put on a log scale with max at 0 dB
power1 = 20*log(power1/max(power1));
power2 = 20*log(power2/max(power2));

f = f(fmin+1:fmax-1);
power1 = power1(fmin+1:fmax-1);
power2 = power2(fmin+1:fmax-1);

figure(3)
subplot(2,1,1)
if log_opt == 1; semilogx(f,power1); end
if log_opt == 0; plot(f,power1); end
axis([f(1) f(end) min(power1) max(power1)])
xlabel('frequency (Hz)'); ylabel('power');
title('Power Spectra')
subplot(2,1,2)
if log_opt == 1; semilogx(f,power2); end
if log_opt == 0; plot(f,power2); end
axis([f(1) f(end) min(power2) max(power2)])
xlabel('frequency (Hz)'); ylabel('power');

% ********************** COHERENCE COMPUTATIONS AND PLOTS
************************

figure(4)
[coh,F] = cohere(sound1,sound2, NFFT,fs1,170,160,'none');
Fmin = max(find(F < f_cutoff(1)));
Fmax = min(find(F > f_cutoff(2)));
F = F(Fmin+1:Fmax-1);
coh = coh(Fmin+1:Fmax-1);
if log_opt == 1; semilogx(F, coh); end
if log_opt == 0; plot(F, coh); end
axis([F(1) F(end) min(coh) max(coh)]);
title(strcat('Coherence',' between ', song1, '  &     ', song2))
xlabel('Frequency (Hz)'); ylabel('Coherence Function Estimate');
coh_mean = mean(coh)
```

**References**


Altes, R.A., (1980), Detection, estimation, and classification with spectograms, Journal Acoust. Soc. Am. 67(4), 1232 - 1246

BBC One (2000), Predators, Part 7: Natural Born Killers, BBC Worldwide Ltd.

Bloomfield, P., (1076), Fourier ANalusis of Time Series, John WIley & Sons, New York, 258 pages

Brillinger, D.R., 1975) Time Series: Data Analysis and Theory, Holt, Rinehart & Winston, New York, 500 pages

Ford, J.K.B., and Fisher, H.D., (1983), Group-specific dialects of killer whales (*Orcinus orca*) in British Columbia, In *Communication and Behavior of Whales*, edited by R. Payne, Vol. 76, 129 – 161.

Ford, J.K.B., (1989), Acoustic behaviour of resident killer whales (*Orcinus Orca*) off Vancouver Island, British Columbia, Can. J. Zool, 67, 727 – 745.

Ford, J.K.B., (1991), Vocal traditions among resident killer whales (*Orcinus Orca*) in coastal waters of British Columbia, Can. J. Zool 69, 1454 – 1483.

Ford, J.K.B., (1992), Communication and Dialects of Killer Whales, CD track 1 on *Blackfish Sound, Underwater Communication of Killer Whales*, Vancouver Aquarium Marine Science Centre

Ford, J.K.B., Ellis, G.M., and Balcomb, K.C., (2000) *Killer Whales,* 2nd edition, UBC Press, Vancouver.

MathWorks, (2000), MATLAB, Version 6.0 Release 12, Signal Processing Toolbox, User's Guide, The MathWorks Inc.

Mellinger, D.K., (2000), Recognizing transient low-frequency whale sounds by spectrogram correlation, Journal Acoust. Soc. Am. 107(6).

SONY, (2003), Sound Forge Version 7.0, Sony Pictures Digital Inc.

Spong, P., Symmonds, H., (2001), Orcas of the Inside Passage, CD, OrcaLab.

Thomsen, F. and Franck, D., (2001), Characteristics of whistles from the acoustic repertoire of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia, J. Acoust. Soc. Am. 109(3).