Statistical comparison of southern resident killer whale (*Orcinus orca*) calls using fundamental frequency

KARI SHERMAN

OVAL Research Project
Colorado College Physics Dept.
902 N Cascade Ave, Colorado Springs, CO, 80946, U.S.A.
E-mail: k_sherman@coloradocollege.edu

ABSTRACT

Underwater vocalizations by the southern resident killer whale J pod of Puget Sound, WA were compared to detect differences in the fundamental acoustic frequency of several different calls. The call types S1, S3, and S4, were analyzed using a spectrogram computer program (Veirs 2006). Ten examples of each call were selected from a catalogue of manually identified calls. The call types were compared using mean, standard deviation, standard error of the mean, Student's t-test, histogram graphs, and skewness to look for a significant differences between the three calls' fundamental frequencies. The statistical analysis concluded that the calls were significantly different. The currently system of manual auditory call classification contains great uncertainty due the dependence on the human ear. With further research this process could be the beginning of an automated call detection and classification system based primarily on fundamental frequency of killer whale calls.

Key words: vocalization, sound, underwater acoustics, call type, fundamental frequency, harmonics, orca, killer whale, auditory classification.

INTRODUCTION

There are resident killer whale *(Orcinus orca)* pods off the west coast of Washington State and British Colombia. Each pod has a repertoire of about fifteen calls (Ford 1991). The calls in these repertoires exist for the life span of the orcas. Changes made in a call repertoire occur only due to certain circumstances. An orca call has a discrete recognizable structure with specific duration and frequency (Ford 1987). Presently the common method used to identify an orca call is manual human recognition by ear (Balcomb, 2006). This system is not only inefficient, in it inconsistent due the dependence on human ability and therefore human error.

Killer whales of North America

There are three types of killer whales known to swim the coast of North America. Off-shore killer whales are rarely observed because they do not usually come into the harbors. Transient killer whales swim the coast of British Colombia and the U.S.A. but stay in small groups and are present seldom. The resident killer whales in the inland waters of British Colombia and Washington State are the most understood of these three orcas due to the fact that they are most observed and spend most of their time in the coastal waters eating salmon (Balcomb, 2006).

The resident orcas are broken up into two groups, northern and southern residents, by the locations that they live (Ford, 1991). There are sixteen northern and three southern resident pods (Ford, 1991). Pods follow a matriarchal pattern with usually 1-7 cows per pod (Ford, 1991). To be called a pod means that the killer whales spend at least 50% their time together during the year (Foote, 2005).

An index of similarity between pods can also be calculated using the acoustic sounds that the pods make (Ford, 1991). A clan is a group of orca pods that all share a portion of their call repertoires. There are three northern clans and one southern clan. A pod of orcas can produce an average of 13 different calls (Ford, 1991). Two pods that make several of the same calls are acoustically similar to each other (Ford, 1991). The three southern resident killer whales pods are all of one clan and therefore are very similar in call repertoire (Ford, 1991). Pods J and K have an index degree of call sharing of .5 acoustic similarity (Ford, 1991). They share half of their calls. Pod L of the southern residents rates at an index degree of .3 and shares 30% of its calls with either of the two other southern resident pods (Ford, 1991). The acoustic similarity index is based on proportions of call repertoires shared between pods.

The call repertoire

A call repertoire is the quantitative name for the group of calls a pod can make. Of the southern residents, J pod's repertoire consists of 18 calls, K's of 10 calls and L's repertoire is 15 calls (Ford, 1987). There are only two calls that overlap through all three pods (Ford, 1991). The calls in each of these repertoires must be passed along through the pods by some process or exposure if they are not innately known (Foote, 2005).

The process in which young orcas learn calls through mother to calf exposure has been documented by Weib et al (2006).   There is a documented increase in family specific calls made by a pod for about ten days after a calf is born (Weib et al, 2006). Adult orcas intentionally expose the new calf to the family call. The call repertoire of an adult orca is less due to genealogy and more due to time and place of birth making an acoustic association the most reliable way to track genealogy between orcas. Orcas can change which sub-pod or family grouping they swim in, but the calls that they make are the same. The acoustic similarity of orcas is a very important tool to be able to accurately measure in order to better understand the whale life cycle.

Beyond the basic mimicry of the mother-calf relationship the calf can learn calls from other orcas too.   Foote (2005) discusses the transmissions of calls between orcas in three ways. The passage of acoustic call knowledge from parent to offspring is termed vertical transmission (Foote, 2005) and is probably the most common type. The label horizontal transmission applies to calls learned across killer whale generations from adults or grandparents to juveniles or calves within a single family (Foote, 2005). Oblique transmission is the term for calls learned by orcas from other orcas in the same pod but not of the same generation or direct family blood line (Foote, 2005).

Each pod's call repertoire has one call type that makes up about 50% of all the calls made by that pod (Foote et al, 2006). This one call type is the signature call for the pod (Weib et al, 2006). Orca call similarities are linked by familial relationships (Ford 1991). Call functions have been hypothesized by many researchers. It has been proposed that calls are made for the purpose of maintaining group cohesion and sharing information on foraging (Foote et al, 2006; Ford, 1991; Balcomb 2006). Calls are made to maintain contact between pod members while traveling and to distinguish between related and non-related orcas (Weib et al, 2006). Orca calls also are hypothetically used to alert other pods to a pod's presence long before a physical confrontation would occur (Balcomb, 2006).

Beyond each pod having a set of calls it makes and a single call for identification, the orcas in each pod have a unique way to produce calls (Balcomb, 2006). K. Balcomb (2006), whale researcher of San Juan Island, WA believes that it is possible to identify the individual whale that produces a call based on the individual style of the call. The ability to recognize the origin of the call by individual whale takes hundreds of hours of exposure to do accurately, but to what degree of accuracy is immeasurable. The accuracy is not quantifiable because the calls do not have an official accepted standard to be measured by.

Change / no change in calls

Killer whale call repertoires are the most stable collection of any animal besides humans (Ford, 1991). The oldest recordings of orca whales are from the decade of 1950. These recordings can be identified as J pod by present whale researcher K. Balcomb (2006) using manual auditory recognition. The recording demonstrates that J pod made the same calls in the same fashion over fifty years ago. K. Balcomb (2006) suggests that he can identify the individual whale that made the calls in the 1950 recording as a certain male alive today. Call types have been documented to be in the same call repertoire for at least twenty-eight years by a long time study under direction of J. Ford (1991) (Foote et al, 2006). These call repertoires are uncommonly stable, and yet changes over time do transpire.

A. Foote's (2005) thesis work, "Correlates of variability in killer whale stereotyped call repertoires," compares call usage and structure. Foote et al (2006) then compares the findings to Ford's collection based on research from 1977-1981. Foote (2005) identified three sources of vocal variation. The term "random copy errors" refers to a process of incorrect mimicking (Foote, 2005). Innovation of old call types is a process by which individuals alter an existing call so much it must be categorized as a separate call. The third, invention of new call types, can be literally interpreted (Foote, 2005). Foote's findings support that genes do not influence the adoption of certain call types; rather it is a learning process due to time and place of birth.

The discrete call

Orcas make many calls that can be grouped together into discrete call types based on frequency and duration of the call (Ford 1991). A selection of discrete calls will make up a repertoire. Each call type can be produced by one or more pods and by many orcas (Ford, 1991). There are multiple syllables in a single call and each can be quantified by an average length in milliseconds and an average frequency in hertz.

The auditory tone of a killer whale call is made by a system of rapidly emitted pulses (Ford, 1991). The rate of these pulses will determine the harmonic interval seen in a spectral analysis (Ford, 1991) (Fig. 1). A change in the repetition rate reflects a change in the frequency that represents a new syllable.

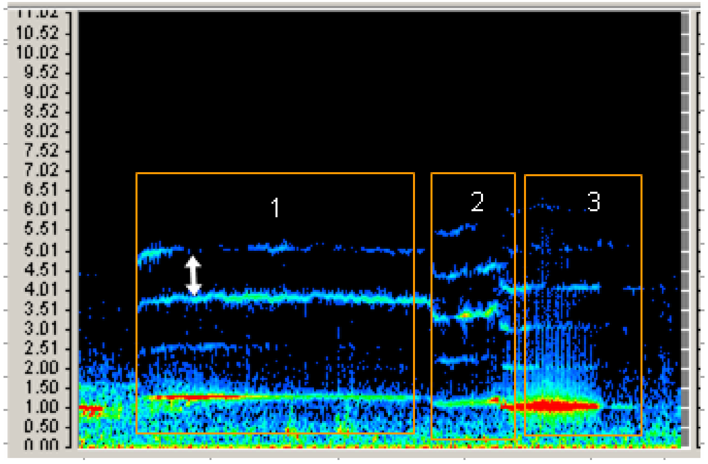


FIG. 1. Spectral analysis of a discrete call. Three syllables indicated with orange boxes and numbered 1,2, and 3. Harmonic interval is shown with a white arrow. The fundamental frequency is the bottom green and red line. Ambient noise in the background of the recording appears as dashed variable red and green dots. This call is S9 by J pod. It is 1.5 seconds in length. Identified by V. Veirs, Orca Sound Analyzer (2006).

Ford's (1987) "Catalogue of Underwater Calls Produced by Killer Whales (Orcinus orca) in British Colombia" was the first attempt to establish a comprehensive inclusive call collection. The catalogue made possible a dialogue among scientists using common observation standards with which to discuss vocalizations of killer whales. This catalogue also made possible the idea that researchers could identify a “normal” discrete call as opposed to an “aberrant” discrete call. Aberrant calls are distorted discrete calls that are recorded when killer whales engage in “excited” activities, such as socializing or feeding (Saulitis et al, 2005), or the birth of a calf (Weib et al, 2006). Abnormalities in frequency and duration can be detected to identify these excited behaviors.

Frequency of calls

Frequencies when played through musical instruments resonate at different harmonics (Hall, 1991). A fundamental frequency is the lowest note that an instrument can make (Hall, 1991), be it an object or a vocal cord. Harmonics are whole multiples of the fundamental frequency (Hall, 1991). Every killer whale discrete call has a quantifiable fundamental frequency that can be measured by calculating the average distance in Hz between the harmonics of the call. As seen in Fig. 1 it is possible to measure many harmonics in a single orca call. Some calls’ harmonics resonate at as high as 60 kHz (Balcomb, 2006), but this level of detection is hard to achieve due to the complicated situation of electronics, salt, and the dense medium of water that the sound travels through. There is relentlessly too much ambient noise in a recording to have devices sensitive to the 60kHz level (Veirs, 2006).

The Problem

Killer whale discrete calls have an average tonal frequency unique to each call (Ford, 1987). Higher frequency calls are more directional underwater than lower frequency calls (Miller 2006). Young small orcas may produce a higher frequency call than the average adult orca (Kneipp, 2006, unpublished). When there are calves less than two weeks old a pod will show excited behavior and make aberrant calls (Weib 2005). Unfortunately there has been no link identified between these four coincidental situations. This may be due to the fact that it is very hard to quantify data on individual calls.

The current method of manual auditory call recognition to identify and quantify killer whale calls (Balcomb, 2006) allows a significant amount of error. There is no mathematical standard to compare calls to, making the area of underwater orca vocalization research inconsistent. The identification method makes discrete call observation highly unreliable. Different individuals may make different judgments. There are few authorities on killer whales' vocalizations available to reliably perform such identification research. Reliable mathematical definitions of discrete calls are necessary to advance the field of underwater killer whale vocalization research. Once this is established, data will be universally comparable and aberrant calls will be detectable. The present opinion holds that discrete calls are context-independent (Ford, 1991). This may be due to the true nature of killer whale calls or it could be the result of erroneous data collection. There have been little to no solid connections established between call type usage and behavior (Ford, 1991). Foote et al (2006) suggest that group specific call type repertoires are essential for remote acoustic monitoring and thereby need to be established. More data is the field of orca vocalizations is needed to resolve these new hypotheses.

I propose a system of measurements based on the fundamental frequencies of discrete calls to begin the mathematical dictionary. I chose to measure the fundamental frequency of three discrete calls S1, S3 and S4 by the southern resident J pod of Puget

Sound, WA.. To determine the cohesiveness of each set of calls I used the standard error of the mean. To establish if the difference of means was significant I used Student's t-test.

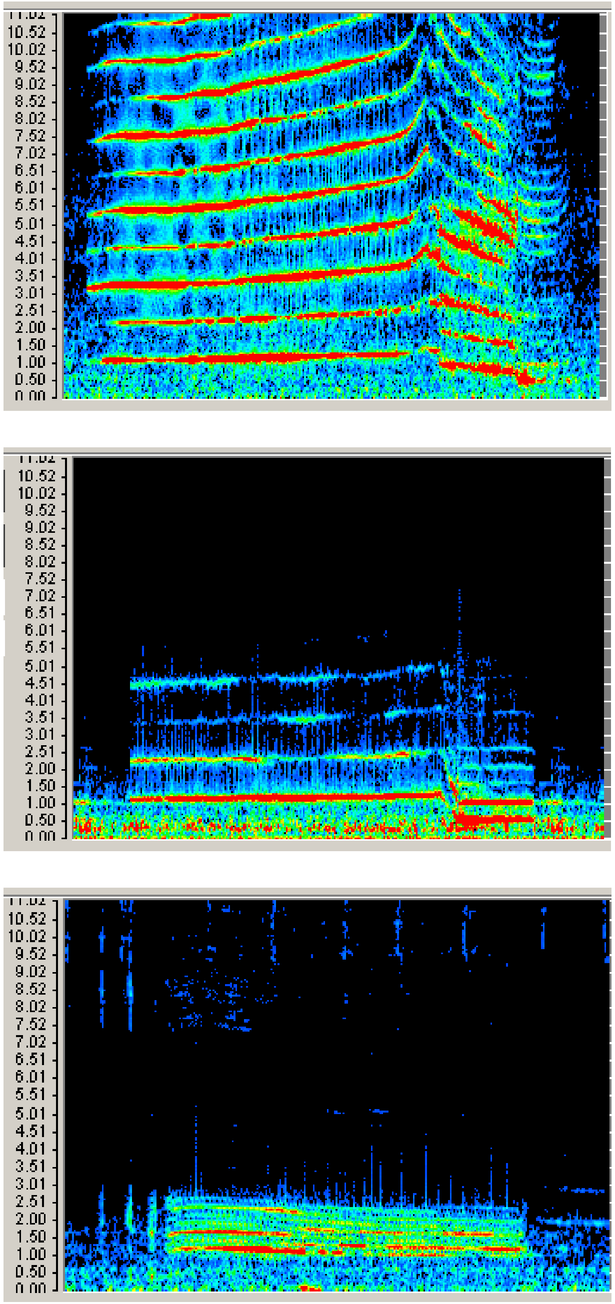


FIG. 2. The S1 call type (top) has a long first syllable, a downward slanting second syllable and a flat, short third syllable. The S3 (middle) has two syllables both relatively flat. The S4 (bottom) is a single long syllable. The frequency scale in KHz along the vertical axis is consistent for all three to show the relative size of each fundamental frequency.

Call Type

Three separate calls as identified by Ford (1987) were selected to be the study types. The S1, S3, and S4 call types are the most used calls of the southern resident J pod. S1 is the signature call of J pod comprising more than 60% of total call usage (Foote et al, 2006). It has three syllables (Fig. 2), the first dominating the duration of the call. S4 is the second most used call going from about ten to now near 20% of the total call usage in the last decade (Foote et al, 2006; Ford 1991). S4 has a notably smaller harmonic interval than S1 and S3. The S3 call overtook the S7 call since J. Ford's study (1991) on percent call usage, but only makes up about 5% of all calls made by J pod (Foote et al, 2006). S3 makes up 10% of call usage by K pod (Ford et al, 2006).

The S1 call is a long held frequency with two syllables at the end which drop twice in frequency (Fig. 1). The S3 call type is also a long held frequency and then a single drop at the end (Fig. 1). These drops in frequency represent syllables. The second syllable of a call will have a different fundamental frequency that will usually be smaller. For the S1 and S3 calls there is always a smaller fundamental frequency in the later syllables of the calls. The S4 call is a single syllable call that sometimes appears to have "feathers" on it when looked at in a frequency verse time graph (Fig. 3). It has the smallest fundamental frequency of the three calls. Often the beginning of the S4 call sounds like a "crank" because the pulses are almost discernable. Sometimes the S4 call sounds like a broken horn when the pulses are slightly faster.

METHODS

The predictions of this study are that there will be a separate fundamental frequency for each of the three call types analyzed. Standard error of the mean analysis will determine if the call types form three individual cohesive groups. Statistical t-tests on the fundamental frequency data sets prove if the call types are significantly different to the 95% confidence level. In an effort to minimize variables I choose calls that where the most frequently produced by J pod and are rarely, if ever, produced by K and L pods.

Measurements

Harmonics were measured by choosing a small portion of the call, about .05 seconds long, from the middle of the first syllable (Fig 3). The selected time span was chosen to include the maximum number of harmonics in the call. The harmonics were identified using a spectrogram analysis program (Veirs, 2006) and a frequency-power graph (Fig 5). The peaks of the frequency-power graph were measured to create the data sets for each call.

FIG. 3. White box represents the piece of this S4 call type selected for measurement the vertical axis is in KHz, horizontal axis represents time (Veirs, 2006). Notice the “feathers” in red below.

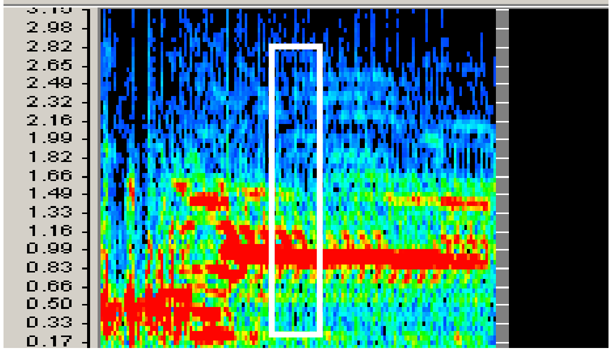
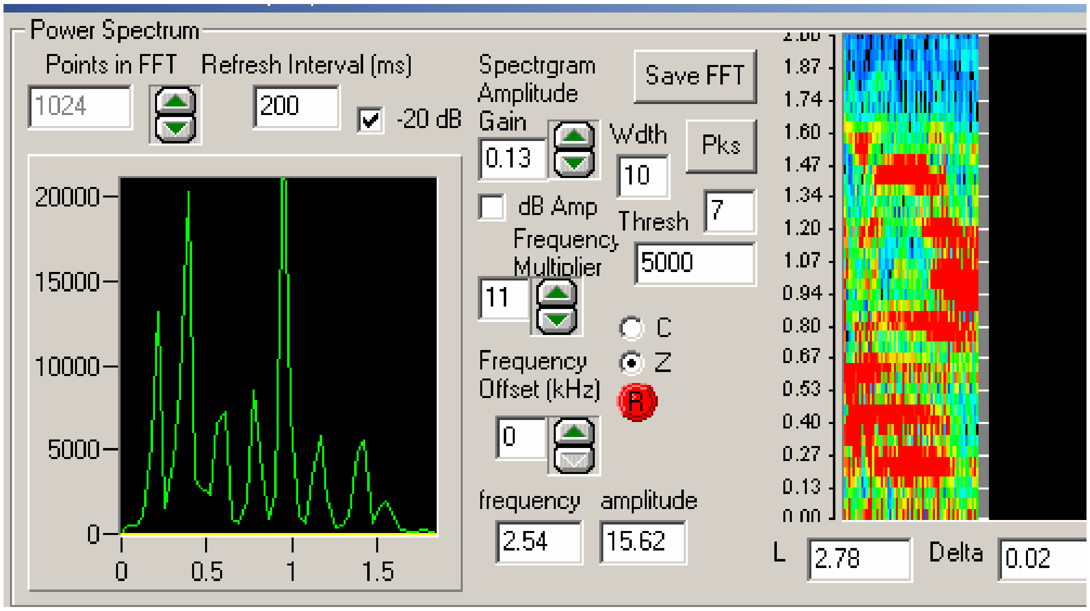


FIG. 4. The is the piece of Fig. 3 analyzed using the spectrogram to isolate the harmonics. To the right the enlarged portion of Fig. 3 is shown. There are 7 visible harmonics. To the left is the graphical representation of the frequency-power spectrogram. The peaks of the green graph represent the power put into each frequency of the call by the orca. The horizontal axis is measured in KHz, the vertical axis is amplitude measured in Amps. Picture from Veirs (2006).



Power Spectrum

Points in FFT: 1024

Refresh Interval (ms): 200

☑ -20 dB

Spectrgram Amplitude Gain: 0.13

Wdth: 10

Thresh: 7

☐ dB Amp

Frequency Multiplier: 11

5000

○ C

● Z

Frequency Offset (kHz): 0

frequency: 2.54

amplitude: 15.62

L: 2.78

Delta: 0.02

Procedure

The fundamental frequency was established for twenty-seven calls, ten S1, seven S3, and ten S4 call types. From the amplitude-frequency graph the peaks were measure with accuracy to two decimal places. For each call there was a range of 3 to 25 data points measured depending on the clarity and measurability of the harmonics. The difference from the first harmonic to the second harmonic was measured, from the second harmonic to the third harmonic, and so (see Table 1).

The mean fundamental frequency for each call was established from the mean of the differences between the harmonics. The standard deviation and standard error of the mean were calculated for each call to see how closely the harmonics varied around the mean fundamental frequency. Once a mean fundamental frequency was established for

S3 call type

Harmonics	Peaks	Difference	Mean	Std. Deviation	Std. Error
1	0.99	0.99	0.982	0.03675746	0.0116237
2	1.98	0.95			
3	2.93	1.03			
4	3.96	0.95			
5	4.91	0.95			
6	5.86	1.03			
7	6.89	0.95			
8	7.84	1.03			
9	8.87	0.95			
10	9.82	0.99			
11	10.81				

TABLE 1. Example Data Sheet for one call. For this particular call there were eleven measurable harmonics including the first harmonic. From these harmonics it was possible to measure the peaks form the amplitude-frequency graph(Vears, 2006). The difference between each peak was calculated by subtracting the first peak from the next higher peak and so on. The mean of the difference was found along with the standard deviation from the mean and the standard error of the mean for all thirty calls.

each call of a common call type I could establish the mean fundamental for the general call type to compare with another call type. T o establish the standard error of the mean fundamental frequency for the ten calls of each call type I took: sqrt ((sum of std of error^2)/N calls). See Table 2 in results.

Equipment consisted of several pre-recorded orca calls, Orca_Sound_Analyzer program, and Microsoft Excel. The pre-recorded calls were collected by V. Veirs with Colorado College, S. Veirs with Beam Reach marine science school, K. Balcomb with the Center for Whale Research, and by the Whale Museum on San Juan Island. These calls were all collected at Puget Sound, WA, near San Juan Island.

Orca_Sound_Analyzer is a computer program written in Visual Basic for analyzing orca calls. Any sound program with capabilities of measuring the amplitude and frequency of an orca call to equal precision would also work.

RESULTS

The frequencies of each harmonic for twenty seven calls were measured. The mean fundamental frequency was gathered for each call. One mean fundamental frequency was then created for each call type S1, S3, and S4 (Table 2). From this data the standard deviation of the mean and the standard error of the mean were calculated for each call type. Histogram graphs and skewness were used to determine the validity of each call type data set. Student’s t-test was then used to look for significant differences.

Call Type	Fundamental Frequency Mean(Hz)	Std. error of mean (Hz)	Range Low (Hz)	High(Hz)
S1	1074.765	19.3	1055.465	1094.065
S3	1092.67	20.1	1072.57	1112.77
S4	181.6	11.9	169.7	193.5

TABLE 2. The Standard error of the final mean fundamental frequency establishes that the three calls are only two distinct

From the ten calculated fundamentals for each call type a histogram graph was produced to test the normality of the data (Table 3). Skewness of the data was also calculated to be .1085 for S1, -1.062 for S3, and -.015 for S4 call types. Because the S3 call type histogram graph had such an abnormal curve, the data was deemed unreliable, and could not be used for analysis. Any comparison using the S3 data set would lead to error.

Using Student's t-test the two sets of data, S1 and S4, were compared. A t-test value of 42.4 (ref. p=2.81) yielded the result that there was a significant difference between the S1 and S4 call types with a 99% confidence level. Again, all comparisons to the S3 call type were discarded due to extreme abnormality.

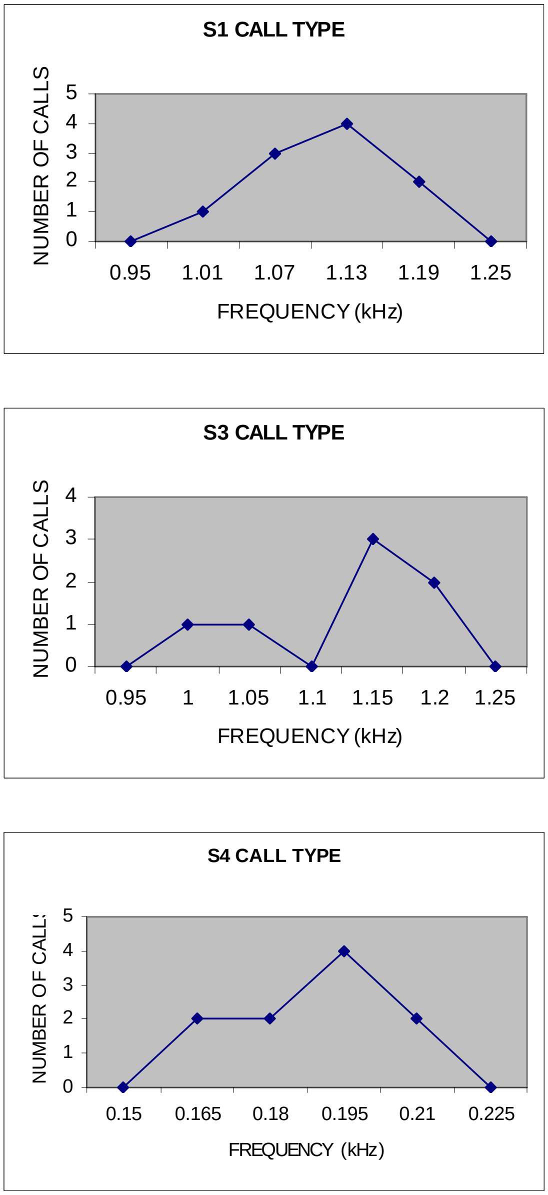


TABLE 3. Histogram graphs by call type. Shows distribution of mean fundamental frequency for each call type. The S3 call data does not form the normal bell shape.

DISCUSSION

There were two patterns, S1 and S4, in the original data that help to answer my original question, "can the fundamental frequency alone be used to identify the type of call that a killer whale makes?" The data gathered from S1 and S4 are significantly different to such an extent that they could be identified as separate calls based only on their fundamental frequency. This was an expected outcome that supports the hypothesis of this study. It should be noted that the fundamental harmonic was not always measurable for the analyzed calls. Rather, the interval between the harmonics was always measurable and from there it was possible to determine the fundamental frequency.

Error in the S3 data set may have come from two possible sources. K pod uses S3 at a 10% of total call usage rate whereas J pod only uses the S3 call type for 5% of total calls (Ford et al, 2006). Both S1 and S4 were used considerably more by J pod than K or L pods to a difference that they were only recorded at fewer than 2% of total call, combined for calls produced by K or L pods (Ford et al, 2006). It is highly probable that the S1 and S4 calls analyzed in this study were only gathered from J pod. It is possible that the S3 calls analyzed in this study were made by K pod and not J pod. Because pods have different dialects (Balcomb, 2006) this could have increased the variability in the S3 call type data set.

Ford (1987) measured the harmonic intervals at the beginning and end of the first syllable, where as I took one measurement from the middle of the syllable. Ford's (1987) method also included a much larger sample size than this study (Table 4). It is of interest that the mean calculated in this study is consistently greater than that of Ford (1987). This may be due to the equipment used to measure, sample size, or the pod which produced the calls.

Call Type	Mean (Hz)	Number	Ford's Start (Hz)	End (Hz)	Ford's Number
S1	1074.765	10	1020	1065	52
S3	1092.67	7	1068	1068	21
S4	181.6	10	159	159	29

TABLE 4. Comparison of Data to Ford (1987). It is evident that Ford had a much larger sample size. There is a large difference between the numbers of Ford (1987) and this study.

The data for this study had to be analyzed twice due to initial incorrect identification. The initial categorization was based entirely on manual auditory classification. When the calls were compared to visual images in Ford (1987) it was evident that the calls did not fall into proper categories and the data either had to be discarded or relocated from S1 to S3 and from S3 to S1 groupings. This emphasizes the need for stricter call classification guidelines. Confusion occurred at a 23% failure rate. There were seven out of thirty classifications performed incorrect using manual auditory recognition during the first analysis of the data. This study draws attention to all other studies that used manual auditory call recognition as their primary tool for identifying call type and questions the validity of those methods.

Conclusions

The contrast of the data from S1 and S4 call types support the hypothesis that calls can be identified based on fundamental frequency. The variability in S3 data could have been caused by multiple pods producing the call type, or due to improper call type identification. The harmonic intervals measured in this study are all larger than the intervals calculation by Ford (1987) y about 15 Hz. This could be due to a consistent disagreement in the location that was measured in each study. Because the calls are linear, the harmonic interval sampled from the middle of a syllable should not be larger than both start and end frequencies. More data collection is necessary to draw conclusions from this observation. Using the manual auditory call recognition method a high level of uncertainty must be allowed for in the data gathered. A better identification method would be useful improve the uncertainty in the call recognition process.

Further research

For future work in the area of call categorization by fundamental frequency I suggest mass analysis. A study of the scope and size achieved by Ford (1987) is necessary to create reputable guidelines for this field of underwater vocalization research. The automated killer whale classification system has great potential for the future. The fundamental frequency has the potential to be an essential part of the identification of a call type. Before behavior of whales can be associated with calls, the calls must have a scientifically definable nature, which will be found in the fundamental frequency.

ACKNOWLEDGEMENTS

I would like to thank Val and Scott Veirs for their help with this fun project. Thank you also to Leslie Veirs for her social relief and sumptuous cookies. Thank you to the Whale Museum and the Center for Whale Research on San Juan Island for the prerecorded calls.

LITERATURE CITED

Balcomb, K. 2006. At the Center for Whale Research. San Juan Island, WA. Personal interview on May 12.

Foote, A. D., R. W. Osborne, and A. R. Hoelzel. 2006. Acoustic studies of the Southern Resident Killer whale population: implications for remote acoustic monitoring and indications of vocal behavioral change due to vessel noise. The Symposium on Southern Resident Killer Whales. Seattle, WA.

Foote, A. D. 2005. Correlates of variability in killer whale stereotyped call repertoires. Thesis submitted to the University of Durham.

Ford, J. K. B. 1991. Vocal traditions among resident killer whales (*Orcinus orca)* in coastal waters of British Colombia. Can. J. Zool. 69: 1454-1483.

Ford, J. K. B. 1987. A catalogue of underwater calls produced by killer whales (*Orcinus orca)* in British Colombia. Canadian Data Report of Fisheries and Aquatic Sciences No. 633. Nanaimo, British Colombia.

Hall, D. E. 1991. Musical Acoustics. Brooks/Cole Publishing Company, Pacific Grove, Claifornia.

Kneipp, C. 2005. Killer whale calf vocal development: Understanding cultural transmission through acoustics. Beam Reach marine science and sustainability school.

Kobayasi, K. I., and K. Okanoya. 2003. Sex difference in amplitude regulation of distance calls in Bengalese Finches, *Lunchula striata* var. *domestica*. Animal Biology 53:173-182.

Miller, P. J. O. 2006. Diversity in sound pressure levels and estimates active space of resident killer whale vocalizations. Woods Hole Oceanographic Institute, Woods Hole, MA.

National Academies, The. 2003. Ocean noise and marine mammals. The National Academies Press. Washington D.C.

Veirs, V. 2006. May_8_065OrcaSoundAnalyzer.exe computer program written in visual basics for analyzing orca calls.

Weib, B. M., F. Ladich, P. Spong, H. Symonds. 2006 . Vocal behavior of resident killer whale matrilines with newborn calves: The role of family signatures. Acoustic Society of America.