Do killer whales use pulsed calls as a communication system? Differentiating calls based on characteristic frequencies and harmonics

Dave Cade
Beam Reach Marine Science and Sustainability School
dave102@beamreach.org;

Introduction-

Much work has been done in recording, categorizing, describing and analyzing the sounds made by killer whales (e.g. Grebner et al. 2009). One of the best studied populations is the resident killer whale population of the NE Pacific, specifically those whales that summer along the Washington and British Columbia coasts. Killer whales (as well as all other members of delphinidae) rely heavily on vocalizations for foraging, navigating and communication (Wieland et al. 2009). Their vocalizations can be categorized into three types: echolocation clicks, tonal whistles and pulsed calls (Ford 1989) that are potentially made with different anatomical structures. It is well established that clicks and pulsed calls result from the movement of air through a structure unique to odontocetes called the *museau de singe* by whalers but more recently called the phonic lips. Traits like click repetition rate are determined by the periodic opening and closing of the lips. Whistles are produced similarly, but it is likely that they are regulated as well by nasal plugs and their nodes in conjunction with the ligaments of the blowhole (Berta et al. 2006).

The Southern Resident population can be divided into sub-groups (clans, pods and matrilines) based on geographic range, behavior (such as beach rubbing) and association, but also by call types that are characteristic of each group (Ford 1987). While whistles have proven to be more difficult to categorize (although Riesch et al. did have interesting results in 2005), non-echolocation pulsed calls that occur in discrete, repetitive patterns have been thoroughly categorized in both Northern and Southern resident populations. The calls of the Southern Resident population (the proposed focal population of this study) have been categorized into 26 discrete call types with 9 subtypes (Ford 1987).

The purpose of these calls has not been conclusively determined. Attempts have been made to glean information about specific behaviors from calls made by orcas with mixed results (Weiss et al. 2007, Grebner et al. 2009, Parijs et al. 2004). Ford (1991) and others postulate a communicative purpose at the very least relating to the use of calls to maintain contact among pod members. It is a basic assumption in many papers that vocalizations are used in communication and thus research is focused on what it is that they are communicating (e.g. Thomsen et al. 2002). Indeed, it is common when analyzing sounds that animals make for researchers to assume that any loud sound is a signal and thus to infer a communicative purpose (Bradbury and Vehrencamp 1998). Thus far, however, clear, unequivocal evidence that calls are used in killer whale communication remains elusive. The vast majority of scholars agree that animal communication "involves the provision of information by a sender to a receiver, and the subsequent use of that information by the receiver in deciding how to respond" (ibid, p.2). This definition, however, is unsatisfactory on its own because it begs the question of what is meant by information. Bradbury and Vehrencamp (1998) go on to discuss "true communication" as a signal sent that must benefit both the sender and the receiver in some fashion. It is not the purpose of this study to derive what the benefit of these pulsed calls may be to the sender and receiver, but to show that they have evolved in a way so as to maximize the chance that the message of the sender is received as intended. Put another way, the purpose of this study is to search for evidence that killer whale calls have evolved over time in a way that maximizes their effectiveness as part of an acoustic communication system.

Many current policies for killer whale management have been driven by the evident need of killer whales to communicate with each other vocally in an increasingly noisy environment (for a discussion of anthropogenic noise, see Holt 2008). Although anthropogenic factors are relatively recent additions to ambient noise on an evolutionary scale, the need for killer whales to be understood by other members of their pod has always been present and has likely pushed the development of call sequences into patterns that are readily differentiated. This need parallels the communicative needs of all species, including humans. In 1995, a group of Japanese students at the Transnational College of LEX published what they meant to be a readable account of Fourier analysis. Through their research into the property of waves, they developed a method using analysis of formant frequencies (see discussion of formants below) to describe how the Japanese language has naturally acquired a certain symmetry (figure 1) with regards to the five Japanese vowel sounds. That is, they suggest that:

> [o]ver the millennia, human beings have unconsciously isolated a few sounds that are the best suited for most clearly vocalizing language. The result is [a] symmetrical arrangement [whereby] each vowel has been selected to maximize this symmetry of sounds. (Transnational College of LEX 2002, p. 161)

My own analysis (see methods section for description of process) suggests that other human languages also demonstrate this symmetry of vowel sounds (figure 2). If, as

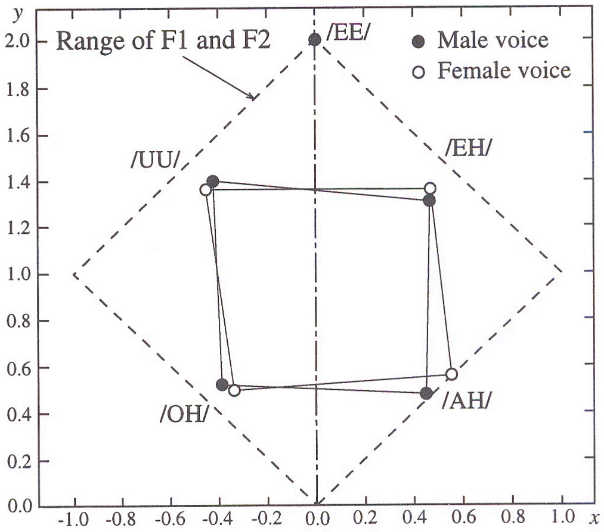

Figure 1- The symmetrical distribution of the five Japanese vowel sounds. The sound /ee/ was shown to have the largest separation of formant values (F1 and F2). The offset from the center of the /ee/ sound and the distance between the F1 and F2 relative to /ee/'s distance were measured for the other four vowels (see discussion in methods section). A logarithmic scale was used since human hearing is based on the octave. The x-axis is offset and the y-axis is relative distance. The dashed line represents the theoretical possible boundaries of the formants. Note that the five sounds are well separated within the boundaries of the diamond. (Transnational College of LEX 2008)

suggested above, this symmetry is a result of the unconscious morphing of sounds over time (as language developed) until there were discrete, replicable units (vowels) that were separate enough so as to be readily distinguishable, then it is also likely that vocal communication systems utilized by other animals would have evolved over time to be as aurally distinguishable as possible, and a similar symmetry would be present. If such a property could be demonstrated in killer whale calls, it would strongly suggest that their vocalizations developed and evolved as a communication system. This study investigates the possibility that the calls of Southern Resident killer whales (SRKWs) demonstrate a pattern similar to that which has been found in the communication systems of humans.

Contributing to the idea that killer whale lexicon has had a chance to evolve over time in an way analogous to language evolution are studies pointing to its changes in the short term (Wieland 2009) and pointing to the development of

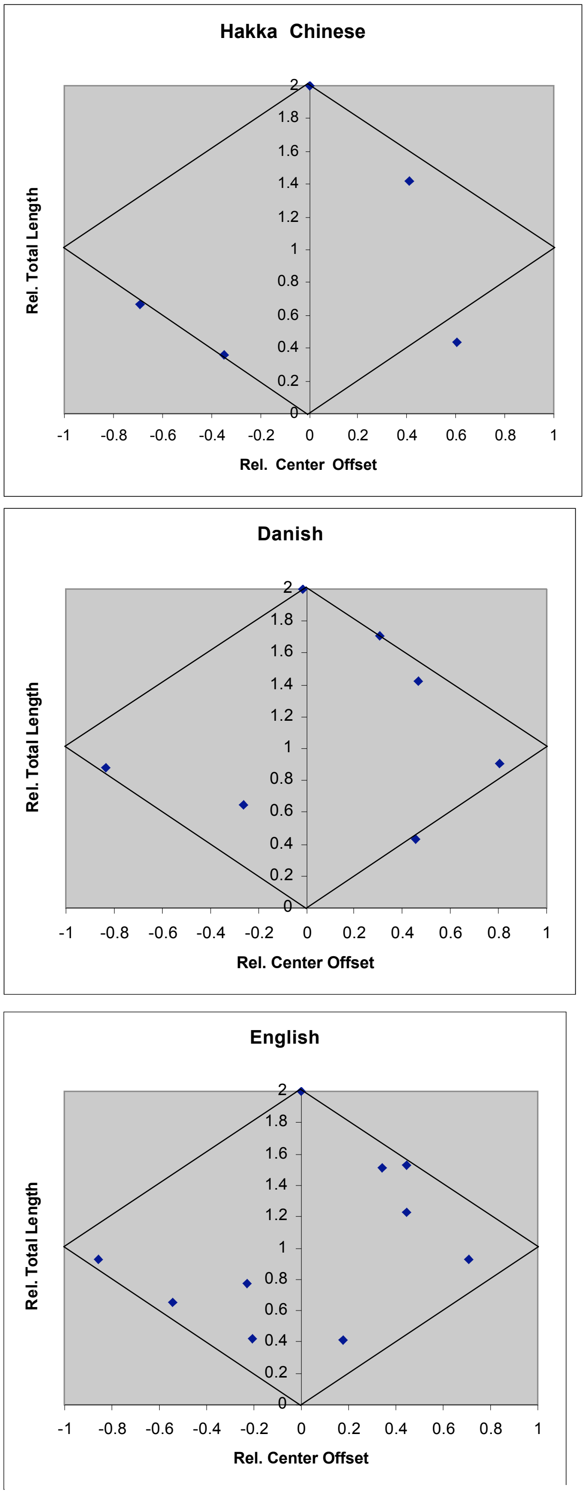

Figure 2- The relative spacing of the formant structure of the vowels of three different languages for adult males. Each point represents a distinct vowel sound. Note the wide spacing of the points that are (with one significant exception) still within the confines of the boundary. Data for top chart from Cheung 2007, other chart data from Lindau 1978.

vocabulary by mimicry of other whales. Vocal mimicry, the building block for developing language over successive generations, has been shown to exist in few non-human mammals, but does indeed exist in cetaceans (Fitch 2000). There is also evidence that young killer whales learn to make specific calls from more mature killer whales instead of genetically inheriting the ability (National Marine Fisheries Services 2008, Ford 1991).

In human languages vowels are determined by the structure of their formant frequencies. That is, formants (specifically the first two formants) are regions that, if removed, change the overall character of the sound of the vowel. More formally, a formant is an identifiable region of high energy in the frequency spectrum of a vocalized sound and is assumed to result from the resonance of the vocal tract (O'Shaughnessy 2008). It is well established that the perception of the sound of a stationary vowel is determined by the first two formants (Kollmeier et al. 2008) and this fact drove the analyses shown in figures 1 and 2. There is precedent for using formant frequencies in other mammals. In red deer, changes in formant frequencies have been used to successfully classify individual animals based on their common roar. The sounds produced by red deer are modulated because of resonances in the supralaryngeal vocal tract. Additionally, variation in the movement of the larynx, mandibles tongue and lips causes individual differences in the formant structure of the calls (Reby et al. 2006). The problem is that killer whales (and all odontocetes) use a completely different set of anatomical structures (as discussed above) to produce pulsed calls. Despite this, many methods developed for analysis of human speech and shown to be applicable to terrestrial mammals have also been used effectively to categorize many types of delphinid vocalizations. Cepstral analysis (which is also used in the identification of formants) has been used effectively to classify odontocetes by species (Roch et al. 2007). Because formant analysis depends on the resonances of structures of the vocal tract, however, it does not appear that killer whale calls can be classified by formant analysis alone. It has also been noted that the difficulties of using formant analysis for killer whale calls could be due to the high value of the fundamental frequency of the calls (Miller et al. 2007). Indeed, despite what appeared to be initially promising results, the first two identified formants have shown not to be determining characteristics of pieces of killer whale calls.

However, calls have long been successfully classified by their harmonic structure (e.g. Ford 1987) and it is properties related to this structure that will be analyzed in this study. Results to date have shown that all calls contain one frequency which, if removed from the spectrum, result in a fundamental change to the sound of the call.[1] Tracking this frequency across different call types in combination with an analysis of the harmonic interval will provide two variables that are fundamental to the character of killer whale calls in a way analogous to the characterization of human speech (specifically vowel sounds) by the first two formants (the two variables in that case). Statistical methods will then be employed to determine if these calls indeed have a maximally differentiated structure.

My hypothesis is that each syllable of the calls commonly used by SRKWs will have a structure uniquely determined by a characteristic frequency and harmonic interval. Additionally, each syllable will be highly distinguishable from each other in a way analogous to how the formant structure of vowels can be shown to be highly distinguishable. If this hypothesis proves true, it will demonstrate that killer whale calls have undergone evolutionary changes similar to human language. Additionally, if some but not all calls prove to be easily distinguishable, this approach could be used to determine which calls are used for communication and which serve some other purpose. If killer whale vocalizations can be shown to have properties that maximize

---

[1] As a sidebar, it will be interesting to investigate whether the characteristic frequencies of the call that are found match at all with the pulse repetition rate. It has been calculated that the pulse repetition rate usually matches with one of the harmonic bands (Deecke et al. 1999)

communicability, it would indicate strongly that these calls do indeed serve such a purpose. Finally, a potential byproduct of this analysis could be that calls that are not shared between pods might be able to be typed and linked to each other for future determination of a shared communicative purpose. That is, for example, if call A is only used by L-pod and call B is only used by J-pod, but the two calls occupy the same space on the x-y plane, it is possible that these calls serve the same purpose in the killer whale lexicon.

Methods-

Southern Resident killer whale calls will be recorded during the months of September and October 2010 in the San Juan Islands region of the Salish Sea. Most recordings will be done in Haro Strait because of its convenience to our base of operations, but we expect to make forays into many of the surrounding straits. We will use a towed LabCore 40 four-hydrophone array which has a peak frequency response at 5 kHz and a single Cetacean Research Technology (CRT) C54 XRS/266 high-frequency hydrophone with a flat response curve from 1-30 kHz. The hydrophones will be towed at a depth of 1.85 m by attaching each cable to a finned weight. The LabCore array will be towed off the port stern and the CRT off the starboard stern. Only hydrophones 1, 3, and 4 will be used in the array to account for the CRT channel in our 4-channel recording devices (a pair of 2-channel SoundDevices 702 recorders). Recordings will be continuous digital files sampled at 192 kHz with a 16-bit depth rate. The files will then be down-sampled to 48 kHz in processing and spliced into 1-minute, 4-channel wav files.

General speed of movement during recordings will be 2.5 knots through the water. Our vessel is propelled by electric motors which minimizes background noise. For most recordings, the array will be towed and thus horizontally deployed behind the boat. At times, however, the boat may be forced to stop or slow down to comply with the “Be Whale Wise” guidelines and Washington state laws. In such a scenario, the hydrophones will lie vertical in the water, but this will not affect the quality of the recordings.

Calls will be typed based on Ford’s 1987 call catalog and digital audio examples put together by Nora Carlson in the spring 2010 Beam Reach session. Each call type will be broken up in to syllables based on abrupt frequency changes. These designations are more specific than Ford’s original determination of “parts” for Southern Resident calls since his parts sometimes include abrupt frequency shifts, peaks and dips in the spectrum. To be included in analysis, a syllable must have a flat (slope flatter than ± .1) section of length > .1 ms.

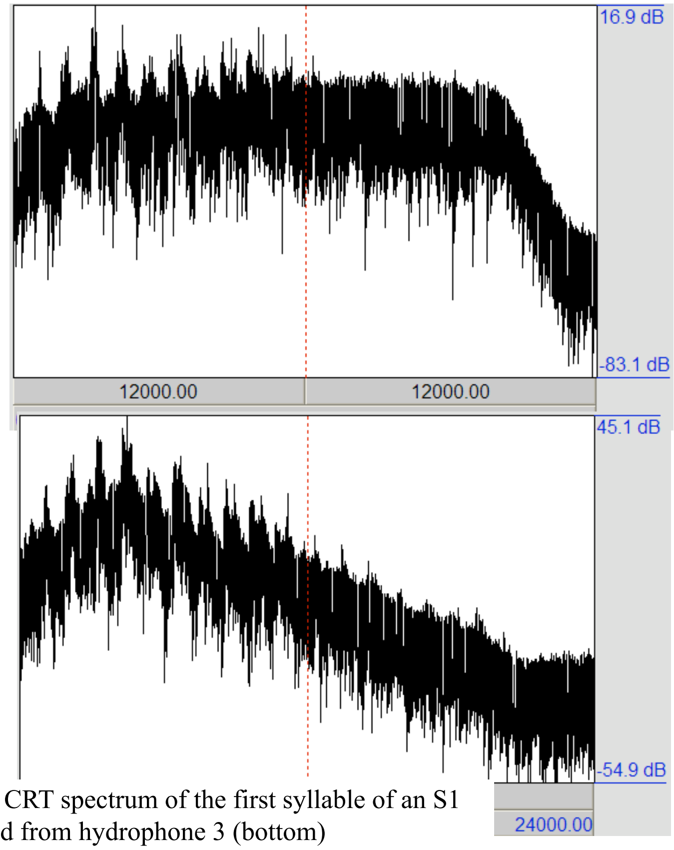

Figure 3- CRT spectrum of the first syllable of an S1 call (top) and from hydrophone 3 (bottom)

The CRT will be used for analyzing the calls when possible because it has a flatter overall representation of background noise across a larger frequency bandwidth (figure 3). The LabCore array will be used only when the received level of the call on the CRT is not at a high enough level to note the differences in aural character. Note in the figure that the peak which registers the highest amplitude in the CRT spectrum is the 3rd peak, whereas in hydrophone 3 (H3) it is the 4th peak. Most recordings made before September 20th will not be analyzed on the CRT because of low gain settings at that time.

Individual syllables will be analyzed (using the spectrum produced by PRAAT) for features that are *characteristic* of the sound, much like how the frequency values of the first two formants determine the sound of human language vowels. Another technique that shows promise and is used commonly to describe individual orca sounds as well as characterize groups of mammal sounds involves cepstral analysis. This will not be attempted here but has potential for future application.

The key piece of this analysis is looking for elements of a call that are *characteristic.* Some progress has been made in identifying frequency peaks that have this property. That is, a peak is characteristic if, when it is stopped along with a 1 kHz band, the sound quality changes. This change is usually marked by a change in timbre, frequency and volume. The frequency of the spectral peak around which the band was removed can be noted (using the peak calculator in Audacity or the formant calculator in PRAAT, which takes an average as the peak shifts). If the peak appears to be bimodal, the average frequency of the two highest peaks will be recorded. The lowest frequency is noted as F1, the next lowest F2 etc. In the call pictured in figure 3, there is only one characteristic peak. That is, there is only one peak which, if removed, changes the character of the sound. On both hydrophones, the characteristic peak is similar (3329 for H3, 3372 for the CRT), but on H3 it is not the peak with the highest amplitude. This discrepancy again suggests use of the CRT data as much as possible.

Because some calls (like the ones above) have proven to have only one characteristic peak, and yet that peak alone is not sufficient to dictate the timbre, frequency and volume of the call, there is another facet of the call that needs to be analyzed. Average harmonic interval (often called the side band interval (SBI) in the literature, see Ford 1987) over the visible spectral peaks will be examined as a measure of the other characteristic components of the call, namely the fundamental frequency. To determine this measure, the smallest and largest easily distinguishable frequencies of a given set of harmonic bands will be measured. Then their difference will be divided by one less than the total number of bands to give an average interval. While these results should be consistent with previously published data about individual calls (e.g. Ford 1987) in instances where my identified syllables overlap with published call parts, there is evidence that killer whale calls change over time (Wieland 2009), so it is important to calculate this interval for the data I am using.

Together, these two measures (characteristic frequency and harmonic interval) should be determinant of the character of the call syllable. That is, if a whale were to change either element, the call would sound fundamentally different. Once these two measures are procured for a variety of calls, then analysis can be performed in a way analogous to that performed on the formant structure of vowels. For a detailed description of the following methodology, see Transnational College of LEX 2002.

With human language, the general procedure is to have a number of speakers of the standard grammatical form of the language produce the same set of sounds into a recording device. The F1 and F2 values are extracted from a spectrogram, and then average results are reported for a variety of categories, usually male, female and children. In all languages examined in preparation for this study (Lindau 1978, Watrous 1991, Cheung 2007, Johnson and Martin 2001, Liu and Ng 2007, Hillenbrand 1995, Kiefte et al 2010, Fant et al. 1969, Peterson and Barney 1951) there is a vowel

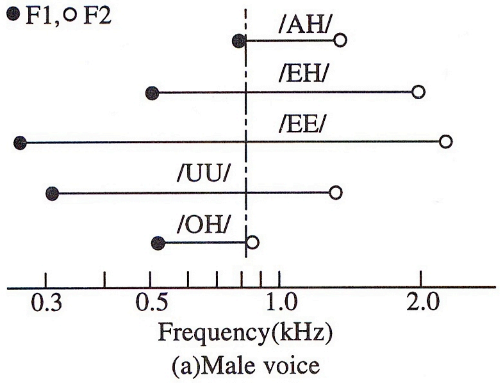

Figure 4
(Transnational College of Lex 2002)

sound that has a smaller F1 and a larger F2 than all other vowels, leading to the conclusion that it operates in the largest possible space for F1 and F2. In languages with vowel nomenclature with which this author was familiar, that sound was always akin to the long “ee” sound as in “beat” or “reed.” If F1 and F2 are plotted for each vowel (figure 4) the result is akin to those reported by the Transnational College of LEX. Note that the x-axis in this scale is the natural logarithm of the frequency values (for reasons discussed in the figure 1 caption). The dashed center line is the center of the logged frequency of F1 and F2 of the longest vowel. It is interesting to note that only rarely in all languages examined does F1 move to the right of this line or F2 to the left. This center line is thus taken as an upper and lower bound, respectively, for F1 and F2. The center of each of the other vowels is calculated and the positive or negative distance to the center of the long vowel is termed the “offset.” A ratio of this offset to the longest possible offset (which occurs for a center halfway between the long vowel’s center and end points) is calculated. This is the x-axis in figures 1 and 2. The y-axis is the length of each vowel relative to the longest vowel length. The diamond encasing the points in the figures is the theoretical boundaries of these values given the assumptions (described above) of the locations of F1 and F2 relative to the longest vowel.

With this new set of values, statistical tests can be performed using R that measure the spread of these data. It is obvious in figure 1 that the five Japanese values are spread very far apart and thus very distinguishable in aural character. The results in figure 2 are a little less obvious, but still demonstrate large spacing. While a final statistical test to measure how spread out the points are has yet to be chosen, results have been encouraging using the coefficient of variation (C.V.) to describe the difference in the mean distances of each point to each other point in the formant diamond. Lower C.V.s exist in the spread out vowel formations, and higher C.V.s are present if the data chosen are random within the confines of the formant diamond.

Finally, it is important to note for this methodology that the exact procedure for killer whales will be slightly different than that laid out above given the different variables involved (characteristic frequency and harmonic interval versus formant values). Because both analyses are derived from two measures of frequency in an audible spectrum, however, a similar methodology is indeed appropriate.

Literature Cited-

Berta, A., J. Sumich and K. Kovacs. 2006. Marine Mammals: Evolutionary Biology. 2nd ed. Elsevier Inc., Burlington, MA.

Bradbury, J and S. Vehrencamp. 1998. Principles of Animal Communication. Sinauer Associates, Inc., Sunderland, MA.

Carlson, Nora. 2010. Southern Resident killer whale call catalog. Complied during Beam Reach spring session 2010.

Cheung, Y. M. 2007. An acoustical analysis of the vowels, diphthongs and triphthongs in Hakka Chinese. 16th International Congress of Phonetic Sciences Publication. Saarbrücken, Germany. www.icphs2007.de: 841-844

Deecke, V., J. Ford and P. Spong. 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(4):2499-2507.

Fant, G., G. Henningsson and U. Stålhammar. 1969. Formant frequency of Swedish vowels. Department of Speech, Music and Hearing: Quarterly Progress and Status Report 10(4):26-31.

Fitch, W. T. 2000. The evolution of speech: a comparative review. Trends in Cognitive Sciences 4(7).

Ford, J. 1987. A catalogue of underwater calls produced by killer whales *[Orcinus orca]* in British Columbia. Canadian Data Report of Fisheries and Aquatic Sciences 633.

Ford, J. 1989. Acoustic behavior of resident killer whales *(Orcinus orca)* off Vancouver Island, British Columbia. Canadian Journal of Zoology 67:727-745.

Ford, J. 1991. Vocal traditions among resident killer whales *(Orcinus orca)* in coastal waters of British Columbia. Canadian Journal of Zoology 69:1451-1483.

Grebner, D., D. Bradley, D. Capone, S. Parks, J. Miksis-Olds and J. Ford. 2009. Killer whale discrete pulsed call variation. Journal of the Acoustical Society of America 126(4):2272.

Hillenbrand J., L. Getty, M. Clark and K. Wheeler. 1995. Acoustic characteristics of American English vowels. Journal of the Acoustical Society of America 97(5):3099-3111.

Holt, M. 2008. Sound exposure and southern resident killer whales *(Orcinus orca):* A review of current knowledge and data gaps. U.S. Dept. of Commerce, NOAA Tech. Memo. NMFS-NWFSC-89, 59 p.

Johnson, K. and J. Martin. 2001. Acoustic vowel reduction in creek: effect of distintive length and position in the word. Phonetica 58(1-2):81-102.

Kiefte, M., T. Enright and L. Marshall. 2010. The role of formant amplitude in the perception of /i/ and /u/. Journal of the Acoustical Society of America 127(4):2611-2621.

Kollmeier, B., T. Brand and B. Meyer. 2008. Perception of speech and sound. The Springer Handbook of Speech Processing. Benesty, J., M. Sandhi and Y. Huan, eds. Springer-Verlag, Berlin, Germany.

Lindau, M. 1978. Vowel features. Language 54(3):541-563.

Liu, H. and M. Ng. 2009. Formant characteristics of vowels produced by Mandarin esophageal speakers. Journal of Voice 23(2):255-260.

Miller, P., F. Samarra and A. Perthuison. 2007. Journal of the Acoustical Society of America 121(6):3932-3937.

National Marine Fisheries Service. 2008. Recovery Plan for Southern Resident Killer Whales (*Orcinus orca*). National Marine Fisheries Service, Northwest Region, Seattle, Washington.

O'Shaughnessy, D. 2008. Formant estimation and tracking. The Springer Handbook of Speech Processing. Benesty, J, M. Sandhi and Y. Huan, eds. Springer-Verlag, Berlin, Germany.

Parijs, S., T. Leyssen and T. Similä. 2004. Sounds produced by Norwegian killer whales (Orcinus orca) during capture. Journal of the Acoustical Society of America 116(1):556-560.

Peterson, G. and H. Barney. 1952. Control methods used in a study of vowels. Journal of the Acoustical Society of America 24(2):175-184.

Reby, D., R. André-Obrecht, A. Galinier, J. Farinas and B. Cargnelutti. 2006. Cepstral coefficients and hidden Markov models reveal idiosyncratic voice characteristics in red deer (*Cervus elaphus*) stags. Journal of the Acoustical Society of America. 120(6):4080-4089.

Riesch, R., J. Ford and F. Thomsen. 2005. Stability and group specificity of stereotyped whistles in resident killer whales, Orcinus orca, off British Columbia. Animal Behavior 71(1):79-91.

Roch, M., M. Soldevilla, J. Bortenshae, E. Henderson and J. Hildebrand. 2007. Gaussian mixture model classification of odontocetes in the Southern California Bight and the Gulf of California. Journal of the Acoustical Society of America 121(3):1737-1748.

Thomsen, F., D. Franck and J. Ford. 2002. On the communicative significance of whistles in wild killer whales (Orcinus orca). Die Naturwissenschaften 89(9) 404-407.

Transnational College of LEX. 2002. Who is Fourier? A Mathematical Adventure. A. Gleason, trans. Language Research Foundation, Boston, MA.

Watrous, R. 1991. Current status of Peterson-Barney vowel formant data. Journal of the Acoustical Society of America 89(5): 2459-2460.

Weiss B., H. Symonds, P. Spong and F. Ladich. 2007. Intra and Inter-group vocal behavior in resident killer whales, Orcinus orca. Journal of the Acoustical Society of America 122(6):3710-3716.

Wieland, M., A. Jones and S. Renn. 2009. Changing durations of southern resident killer whale *(Orcinus orca)* discrete calls between two periods spanning 28 years. Marine Mammal Science 26(1):195-201.