Deciphering the Dolphin Language
Appendix U [1]

by **James T. Fulton**

U.1 Introduction

The goal of this appendix is to report the evidence related to the eternal question of whether the dolphin, and particularly the bottlenose dolphin, *Tursiops truncatus*, unless otherwise indicated, employs what is traditionally described as language in English. If it logically does, the secondary goal is to define the next steps in deciphering that language.

This appendix is subsidiary to "Hearing: A 21st Century Paradigm" and the website Processes in Biological Hearing.

There is a limitation in the vocabulary used in the linguistic community revolving around the multiple uses of the word language. This work will separate the word *language used as a noun* and the word *language used as an adjective*. As an adjective, it will be used to delineate a language protocol (neural software) that is coupled with a vocalization capability (physiological hardware or plant) that together support language (used as a noun) to describe both intraspecies communications and interspecies communications (under more complex conditions).

- - - -

Human academicians of one linguistic school have long argued that only humans employ language; all lower animals only communicate. In a 1980 text[2], Herman & Tavolga (Chap. 4) reviewed the literature of this linguistic school from the communications perspective and Herman (Chap. 8) reviewed the similar literature related to the cognitive capabilities of the dolphin. Herman noted (page 409), "The recent demonstration s of language-learning capabilities in *Pongidae* (with a string of citations and using 1980 cladogram) reveal that most and possibly all of the design features of human language are present in the learned linguistic transactions of the apes. Nevertheless, many linguists continue to defend a distinction between human and animal language skill by assuming, for example, that the demonstrated capacities of the apes, though impressive, are likely to be restricted to the elementary 'Stage-1 Level' of two-word sentences or by defining new criteria for language that are not met by apes (with a string of citations). It seems likely, as Thorpe (1972) suggests, that linguists will continually be able to retreat to new definitions of language in the face of demonstrations that animals have met prior criteria."

> The assertion that only humans employ language reminds one of the similar position that only humans employed color capability in vision; all other animals were monochromatic; today we use the color vision of dogs as models of human color vision and we know all animals (and insects) enjoy color vision.

More recently Chomsky has taken up the cudgel to defend the above linguistic school by introducing a new criteria based on requirement that a species employing language must demonstrate use of the recursive form within its syntax. Hauser reviewed this situation in 2001 and went on to discuss the concept of recursion. Recursion in linguistics will be addressed in **Section U.2.1.4**. Suffice to say here that the recursive technique occurs in two distinct forms and dolphins have been shown to interpret the serial form correctly. Herman et al. have shown that dolphins can in fact understand recursive sentences with a high degree of statistical reliability, even when presented in a totally acoustic "foreign," i.e. synthetic, language. As a result, the noted linguistic school must now retreat once again by insisting that to exhibit true language, a species must exhibit the capability of using the nested form of recursion; a recursive form that humans have considerable difficulty with.

Recently linguists have focused on this so-called unique recursive syntax capability, to differentiate between dolphin *communications* and human *language*. Such a distinction does not appear to be sustainable based on the evidence. There is in fact evidence that the dolphin may have a more flexible vocalization capability than humans. This expanded vocalization capability clearly expands the physiological plant of the dolphin. This capability could easily provide a distinct syntax in the language protocol of dolphins not found in the human language protocol. This equivalent syntax could easily result in greater communications scope than provided by the so-called unique recursive capability of the human language protocol.

---

[1] June 12, 2014

[2] Herman, L. *ed*. (1980) Cetacean Behavior: Mechanisms and functions. NY: John Wiley & Sons *Reprinted in* 1988 by Krieger Publishing, Malabar Fl.

Christian[3] in a TED talk given in 2011 came to less precise but different conclusion, "Powerful message, but I do disagree about our species being unique because we have language. Other species (whales, other primates, elephants, etc.) have language and pass knowledge down through the generations. What may make us unique is WRITTEN language."

As Herman noted (page 410), "Accordingly, the best path to follow at this point may be to bypass additional dispute on definitions in favor of further empirical study of language capabilities of animal subjects."

- - - -

Herman & Tavolga noted (1980, pg 176) three approaches to determining whether animals employ language,
1. Cryptanalysis.
2. Behavioral analysis, also known as ethology.
3. The serendipitous method.

The serendipitous method is assumed and a human and dolphin are placed together for a long time in the hope that a common language will arise between them. Although this procedure usually results in a close social bond, it has never resulted in an interspecies language. One should note the human voice range is below the optimum hearing range for dolphins and the majority of the subtle features of dolphin vocalizations occur at frequencies above the auditory range of humans.

While occasionally noted that deciphering the potential language of the dolphin requires a multi-discipline investigation, no such team of specialists from the required disciplines has yet appeared. This appendix assembles data and conjecture from a wide variety of teams in order to evaluate the potential for dolphin language (as a noun) based on its use of all the elements of a recognizable language protocol. From this assemblage, an initial plan of attack to expose the language of the dolphin is defined. The plan calls for extensive computerized evaluation of the phonemes, morphemes and rule-based syntax of the dolphin. This part of the plan extends beyond the resources of the author.

A critical tool in defining the phonemes involves the discipline of information theory. It is important to note the dolphin language is vocal, not written. This vocal speech involves a phonetic symbol set quite distinct from the alphanumeric symbol set of written languages. Its phonetic symbol set may be analogous to that of humans but this question can not currently be addressed because the dolphin phonetic symbol set is currently an *open* (incompletely defined) set. As a result, attempts to employ the concept of entropy must be examined warily. Functions like the entropy, of any order, of a language can only be calculated if the symbol set is *closed* (completely defined).

The usual test for the existence of a phoneme is the minimal pair test; is there a difference in meaning between pairs such as 'din' versus 'tin'?

- - - -

Considerable progress has occurred during the last few years in determining whether or not dolphins, and particularly the bottlenose dolphins, *Tursiops truncatus*, are able to communicate with each other. It is well recognized that this species has the brain power (based on physical size, complexity, and brain weight to total body weight) and inquisitiveness to support language. Its ecological niche also supports a need for language communications; it feeds in groups patrolling large estuaries where visual communications is limited. It is a tribal animal with various tribes residing in specific areas for very long (evolutionary) times. This relative isolation of tribes with only occasional communications between tribes has led to identifiable differences in the sound patterns used by different tribes.

Adams in his "The Hitchhiker;s Guide to the Galaxy," set the stage for the following discussion:

> "Man had always assumed that he was more intelligent than dolphins because he had achieve so much—the wheel, New York, wars and so on—whilst all dolphins had ever done was muck about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man—for precisely the same reasons."

U.1.1 Physiological framework related to dolphin communications EXPAND

While the physical anatomy of the dolphins is well documented, the literature is much poorer regarding the

[3]Christian, D. (2011) http://www.youtube.com/watch?v=yqc9zX04DXs

physiology of the dolphins, particularly in regard to their vocalization facilities, their hearing facilities and their neural systems, including cognitive capabilities.

After the following global discussion, the more detailed physiology of the dolphin will be considered. This physiology goes beyond the assertions of Connor & Smolker (1996, pg 644) and expands on the mechanisms and capabilities of the vocalizations found in dolphins.

U.1.1.1 The evolution of the cetaceans

A paper by Manger in 2006 generated a large and vocal opposition. His assertion was that the unique development of the members of Cetacea was due to changes in the temperature of the oceans over th eons. The first response with many authors occurred in 2007[4] and included a major survey of the archeology, phylogeny, evolution and cognitive and communications traits of *Cetacea*. ” **Figure U.1.1-1** shows the archeological time line relating to Cetacea and the temperature of the oceans (a key factor in Manger’s assertions).

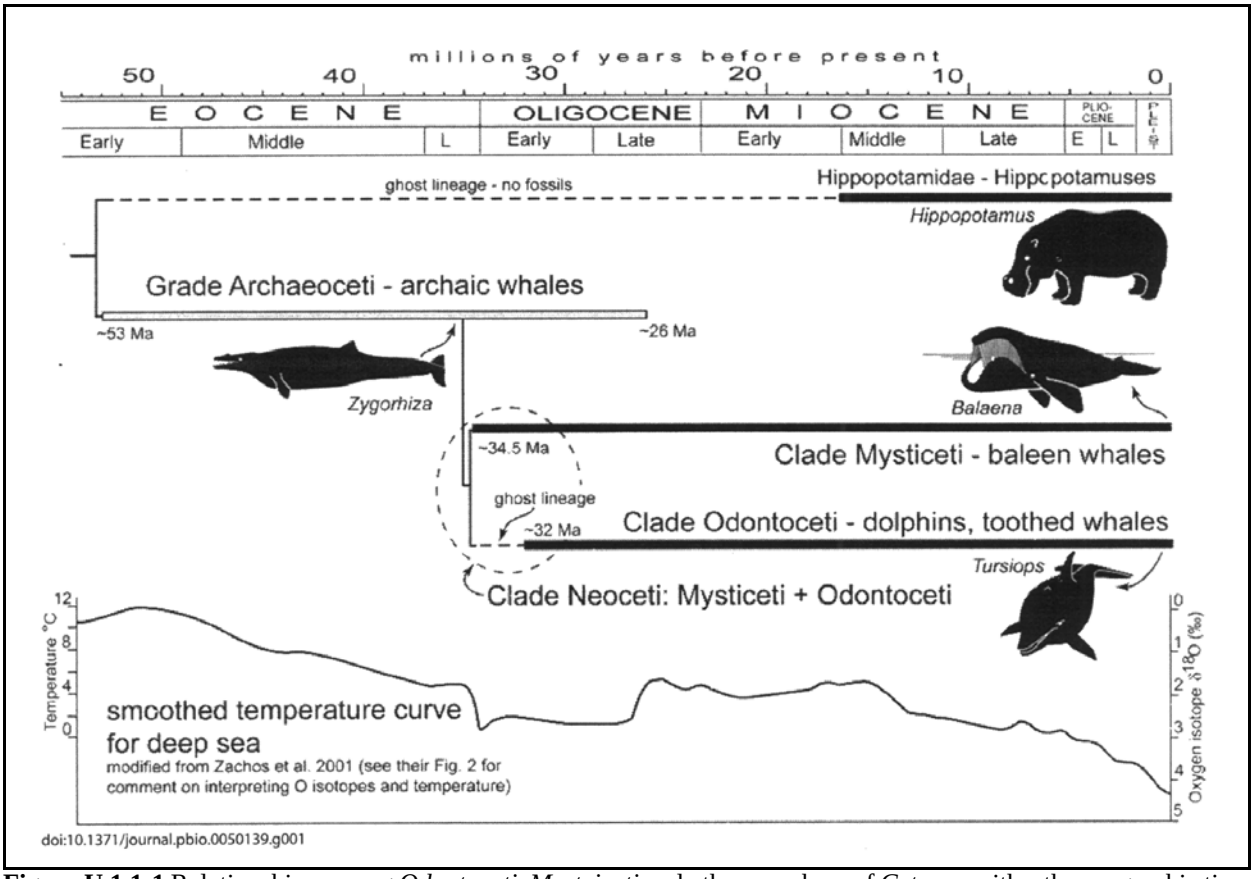

Figure U.1.1-1 Relationships among *Odontoceti, Mysteiceti* and other members of *Cetacea* with a thermographic time line. See text. From Marino et al., 2007

The team does make some detailed observations of the development of the *Cetacea* brain independent of any association with the primate brain, and the appropriate differences in cytoarchitecture to be expected. They indicate the primate and *Cetacea* lines separated at least 55 million years ago. The paper also suggests the presence of many elements of language protocol among *Tursiop*s and probably among several other species of the family. They specifically cite Rehn et al. Rehn et al[5]. develop the character of the vocalizations of the killer whales (also toothed whales, *Odontoceti*) off Vancouver Island, British Columbia. They note a variety of whistles and burst-pulse forms used in assorted sequences during close range socialising. They noted the “first systematic attempt to bring order

[4]Marino, L. Connor, R. Fodyce, R. et al (2007) Cetaceans have complex brains for complex cognition *PloS Biol* vol 5(6): e139.doi:10.1371/journal.pbio.0050139

[5]Rehn, N. Teichert, S. & Thomsen, F. (2004) Structural and temporal emission patterns of variable pulsed calls in free-ranging killer whales (Orcinus orca)

into the many forms of variable calls in wild killer whales, Teichert (citing their Appendix) and Thomsen et al. (2001a) categorized more than 2000 variable calls into six structurally related classes that could be distinguished by contour characteristics and carrier frequency." The Teichert data clearly demonstrates truncation of killer whale vocalizations when recording at a maximum of 20 kHz as well as vocalization complexity similar to that of the bottlenose dolphin. The Thomsen et al. paper demonstrates many vocalizations of the same form in the killer whale as in the bottlenose dolphin (**Section U.2.3.5**).

In a second 24 page paper with many different authors[6] in 2008, many interesting observations (that to a large extent are present in the 2007 paper) are summarized.

"Bottlenose dolphins are adept vocal learners, a trait rare among mammals (Caldwell, Caldwell & Tyack, 1990; Deecke et al., 2000; Hooper et al., 2006; Janik & Slater, 1997; McCowan & Reiss, 1995b; Reiss & McCowan, 1993; Richards et al., 1984). In fact, a large proportion of vocal variation within cetacean species is likely the result of vocal learning (Rendell & Whitehead, 2001). There is evidence for individual-level variation in the whistle repertoires of dolphins (McCowan & Reiss, 2005a). Bottlenose dolphins produce individually distinctive whistles that they apparently use to identify conspecifics and may also be employed as a cohesion call (Janik & Slater, 1998; McCowan & Reiss, 1995a & 2001). There are different viewpoints as to whether these whistles are actually different whistle types, known as ''signature whistles'' (Sayigh et al., 1999; Janik et al., 2006; Tyack, 1997) or a shared rise-type whistle that is individually distinctive, similar to the contact calls of many other species (McCowan & Reiss, 1995a & 2001). In either case there is little disagreement that these are individually distinctive calls that predominate within a much larger whistle repertoire (Janik & Slater, 1998; McCowan & Reiss, 1995a). Manger (2006) refers to these whistles to suggest that dolphins may simply be learning that ''increasing pod coherence in response to a specific vocalisation decreases their chance of being predated'' (p. 318). This suggestion demonstrates a lack of understanding of whistle variation in dolphins. There is population-level variation in whistleparameters (Ding, Wursig & Evans, 1995) and two cases of population-specific (i.e. again - not species typical) nonwhistle call types – 'bray' calls associated with salmon capture off eastern Scotland (Janik, 2000) and 'pop' calls associated with male aggression and herding of females in Shark Bay, western Australia (Connor & Smolker, 1996). The claim that dolphin calls are 'species-specific' needs interpretation; there is abundant within-species variation in dolphin vocal output. Other research that has explored dolphin vocal repertoires and complexity of dolphin communication (McCowan & Reiss, 1995a; McCowan, Hanser & Doyle, 1999 & 2002) has shown that dolphins produce many different whistle types and that there is some evidence that the sequential order of whistles is an important feature of their whistled communication.

Much other information is also included in the Marino (2008) paper, including an archeological record of dolphin development.

In closing, Marino et al. stress the importance to scientific inquiry of the declarative, "form follows function" rather than "function follows form." Form follows function is a major feature in both neural and anatomical dolphin physiology.

U.1.1.2 The neuroanatomy of the dolphin

The neuroanatomy of the dolphin follows the general framework of all advanced mammals (those with crenelated cerebrums in particular). The top block diagram of the mammalian neural system shown in **Section 16.1.1** of "The Neuron and Neural System" is appropriate for the discussions below. Of particular focus are the stage 1, 2, 4 & 5 of the hearing modality and the vocal modality of the stage 7 skeletal-motor system. It will be useful to delineate the vocal modality that is most identified with speech and language from the facial, somatomotor and skeleto-motor modalities that are more closely identified with "body language."

U.1.1.3 The cytoarchitecture of the dolphin

The brain of Tursiops truncatus is larger than in humans, and even on a relative size basis, it is comparable or larger than the human brain.

As noted in the Marino et al[7]. papers, the highly crenelated brain of *Tursiops truncatus* is an example of evolutionary convergence with the brains of the primates. While following independent paths for the last 55 million years (and

---

[6]Marino, L. Butti, C. Connor, R. et al. (2008) A claim in search of evidence: reply to Manger's thermogenesis hypothesis of cetacean brain structure *Biol Rev* vol 83, pp 417–440

[7]Marino, L. Connor, R. Fodyce, R. et al (2007) Cetaceans have complex brains for complex cognition *PloS Biol* vol 5(6): e139.doi:10.1371/journal.pbio.0050139

possibly 95 million years), beginning with a much simpler common ancestral brain. They now look much alike superficially although they exhibit few fissures and lobes in common. It has a highly crenelated surface but the sulci and gyri are unrelated to their position in the brain of homo sapien. As a result the gross structure of the *Cetacean* and primate share little in common. This has made identifying the functions related to the specific cytoarchitecture of the dolphin brain by analogy to the primate brain difficult.

Marino et al. have provided detailed information on the cytoarchitecture of the dolphin[8], *Delphinus delphis*. While the gross organization is familiar to those who have studied the higher primates, the lobes, sulci and gyri are uniquely Cetacean. They note, "The adult common dolphin brain is very similar, at least at the gross morphological level, to that of the adult bottlenose dolphin brain (Marino et al., 2001b). There is a great resemblance between the cortical surface features of the common dolphin brain and the bottlenose dolphin brain. Furthermore, although a comprehensive comparative analysis of sulcal and gyral formations between the bottlenose and common dolphin remains to be conducted, it appears that there is substantial similarity at the level of the location and morphology of cortical gyri and sulci. Also, at the subcortical level the proportions of most of the structures appear to be similar across the two species. As in other cetacean brains, olfactory structures are absent and auditory structures, such as the inferior colliculus, are proportionately very large in the common dolphin brain. Just as in the bottlenose dolphin brain, the enlargement of auditory processing structures in the common dolphin brain is not accompanied by reduced visual structures. Consistent with the behavioral and electrophysiological evidence for a high degree of hemispheric independence, the corpus callosum is small relative to the massive hemispheres. This finding is also consistent with quantitative evidence in other odontocete species."

Jacobs and associates have provided extensive information on the cytoarchitecture of the dolphin[9].

As an aside, it appears the left and right hemispheres of the Cetacean brain operate more independently than among the primates. The various species of Cetacea are able to induce sleep in one hemisphere while the other remains awake, a feature required since Cetacean breathing is under the control of the awake brain more than in terrestrial mammals.

U.1.1.4 The communications framework applicable to dolphin speech

This Appendix will explore the capabilities of the dolphin based on a generic communications system, **Figure U.1.1-2**, used by Shannon and many others coupled to a block diagram of the neurological/motor systems of the dolphin. Stages 1 and 2, the hearing modality of the peripheral neural system of the receiving dolphin, are bit sgiwb explicitly in this figure. They are combined with the stage 4 elements.. The caption describes the functional relationship between the upper and lower elements of the figure.

---

[8]Marino, L. Sudheimer, K. Pabst, A. et al. (2002) Neuroanatomy of the Common Dolphin (Delphinus delphis) as Revealed by Magnetic Resonance Imaging (MRI) *Anatom Rec* vol 268, pp 411–429

[9]Jacobs, M. Galaburda, A. Mcfarland, W. & Morgane, P. (1984). The insular formation of the dolphin brain: quantitative cytoarchitectonic study of the insular component of the limbic lobe. Journal of Comparative *Neurology* vol 225, pp 396–432.

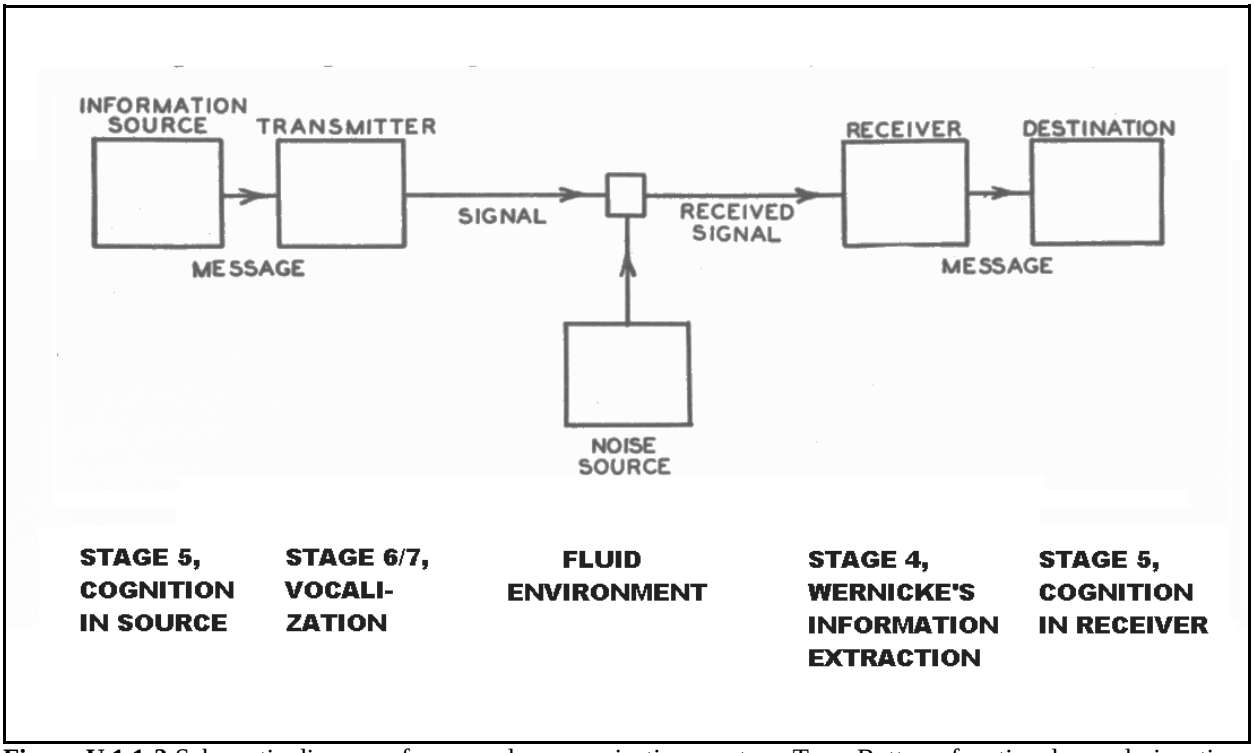

Figure U.1.1-2 Schematic diagram of a general communications system, Top. Bottom, functional area designations within the physiological system and the environment. The information source at top corresponds to the stage 5 cognitive engines of the vocalizing dolphin. The term message in the top view is designated an instruction in the lower view. The tansmitter in the upper view corresponds to the stage 6/7 vocalization mechanisms of the dolphin. The term signal in the upper view is designated vocalization in the lower view. The term received signal in the upper view is described as the perturbed auditory stimulus in the lower view. The receiver consists of the hearing mechanism and Wernicke's area of stage 4 information extraction within the cerebrum. The destination corresponds to the stage 5 cognitive engines of the cerebrum of the receiving dolphin. Modified from Shannon, 1948.

At a more detailed level, Figure S.2.1-1 and Section S.2.1 of Appendix S describe the operation of the overall neural system involved in speech. It also shows the signal paths used in mimicry in animals (including humans). These paths are labeled the cerebellum oriented reflex path, R_C, and the striatum/pallidum oriented reflex path, $R_{S\text{-}P}$.

The linguistic community has entertained great debates over an extended period of time on precisely what language is and is not. Appendix S on "language" of this work provides a major contribution to this discussion that the reader may want to review at this point, http://neuronresearch.net/hearing/pdf/AppendS_Language.pdf . Relevant material is also found in Appendix L on "Dolphin Echolocation" and in Section 8.1.4 on spectrographic techniques within Chapter 8 on information recovery within the neural system.

Humans are currently carrying out an argument concerning language that was duplicated years ago about the unique capability of humans to use tools. We now recognize that a wide variety of animals employ tools effectively, including the dolphins. Others include the otters, octopuses, monkeys and even the "bird brained" Goffins' cockatoo. The cockatoo named Figaro has even been observed fashioning a tool by prying a 2.5 inch splinter from the wood rail of its cage and using the splinter to coral a cashew beyond its reach[10]. ***It appears only a matter of time until humans accept the fact that the highly stylized communications between dolphins employs defined labels and other elements of a lexicon and therefore constitutes a language***.

As early as the 1980's, the neurobiology and anthropology communities have been at odds with the linguistic

[10]McGowan, K. (2014) Crafty Cockatoo. Discover Magazine, March, page 14

community with regard to non-human language. Noback noted[11] that "the vervet monkeys of Kenya, use semantic communications in their alarm calls. These monkeys use three different alarm calls for three different predators and respond differently and appropriately according to whether the call (live or tape-recorded) means a leopard, a martial eagle, or a python. To a leopard alarm the monkeys run into the trees, to an eagle alarm the monkeys look up, and to a snake alarm the monkeys look down." Semantic communications sounds like a synonym for language.

In English, there is no common term for speech-based communications other than language. In the current context, it is important to distinguish between speech-based communications and the language protocol (lexicon, semantics, etc.) supporting that communications. The language protocol has been described as the "faculty of language." Ling noted this situation, "Firstly, an issue of ambiguous terminology must be addressed. The term 'language' may variously refer to a particular language such as English or German, or to the hereditary human traits of language faculty. To avoid misinterpretation, I will use the German word 'Sprache' (plural: 'Sprachen') to refer the former sense, and use 'Faculty' to refer to the latter.

It will become important to separate vocalization from other somatomotor functions (including facial expression). These elements of "body language" should not be confused with the term language as used in vocalization and/or speech.

- - - -

The precise definitions of homology and analogy were also in question during the 1980's. Campbell provided a comprehensive discussion of these terms. The critical question is whether these terms apply to similar structures traceable to a common ancester or whether they refer to a common functional capability. Campbell notes[12] that "the wings of bats, pterodactyls, and birds are homologous as *limbs* but are not homologous as *wings* because their common ancestor lacked wings." On the other hand, these wings are clearly analogous as means of aerial locomotion, flight.

Ploog followed the above investigators in the late 1980's in a broad discussion of "vocal communications" versus human speech but limited his discussion to homology among the primates[13].

As shown by their spectrograms, the vocal system of dolphins shows very similar *functional* performance characteristics to that of the human in spite of major *morphological* differences.

This appendix will take the broad view that *speech and language are associated with the function of communication and not with the morphology of the sound generation equipment of the source.* It will take the position that broadcasting is a one-way use of language in contrast to conversation which implies a two-way use of language. Many forms of communication are restricted to one-way signaling but rely upon language and all of its internal features.

Based on the above background and the analyses to follow, this appendix propounds the null hypotheses that;

Dolphins, and very likely other members of Cetacea, employ a highly developed acoustic communications system that employs all of the critical elements and features associated with the term language.

Man has attempted to communicate with the bottlenose dolphin since ancient times using man's communications tools. This has always required the dolphin to decode, or adopt, man's language. The ultimate goal of this work is to decode the dolphin's language.

Decoding an unknown language is a difficult but not impossible task. It has been done casually by sailors over the span of recorded time upon reaching a new population. Anthropologists have then followed the sailors and provided detailed analyses of the languages encountered. The process of decoding a new language is very similar to the decrypting of an encrypted message. This type of decryption was raised to a high art during the Second World War where messages in a foreign language were encrypted prior to radio transmission. Kahn has provided a masterful book on the decryption of human coded messages by going back to the fundamental characteristics of a language,

---

[11]Noback, C. (1980) Neurobiological aspects in the phylogenetic acquisition of speech *In* Armstrong, E. & Falk, D. *eds*. Primate Brain Evolution: Methods and Concepts. NY: Plenum Press

[12]Campbell, C. (1980) Some questions and problems related to homology *In* Armstrong, E. & Falk, D. *eds*. Primate Brain Evolution: Methods and Concepts. NY: Plenum Press

[13]Ploog, D. (1988) Neurobiology and pathology of subhuman vocal communication and human speech *In* Todt, D Goedeking, P. & Symmes, D. *eds*. Primate Vocal Communications NY: Springer-Verlag

particularly the fundamental characteristics of its syntax and semantics[14]. By following the prescription described by Kahn, any language can be decoded. The effort to decode the language of dolphin has proceeded well along the path outlined by Kahn. The syntax and semantics, while not yet understood, appear to comply with all of the mathematical criteria required of an efficient language. The demonstration by Markov & Ostrovskaya that the sounds of the bottlenose dolphin are well described by a Zipf function provides strong assurance that the dolphins are communicating using a formal language of considerable power and flexibility.

Gregg has taken the view that dolphins have not yet achieved the ability to communicate via language based on his ground rules. Herzing takes a different view based on her ground rules. The next sections will discuss some of these criteria and those author's assertions. Herzing has also presented a general discussion of her work with dolphins in the context of the Search for Extra-terrestrial intelligence (SETI)[15]. Turner has reported extensively on the behavioural aspects of dolphins socializations without taking a position with regard to language versus communications among dolphins.

Herzing has presented a paper[16] in 2006 describing her "Currencies of Cognition." "The identification and isolation of individual structural units of sound within the dolphin signal repertoire have not been adequately studied. Increased understanding of mechanistic and perceptual classification is needed to determine the natural boundaries of signal classification by delphinids, as it has been for other taxa." This work will attempt to build on that base and fill in some of the present voids. Herzing defines a set of 3 letter codes describing the physical positions of the dolphin associated with some of the observed vocalizations.

U.1.2 Recent popular books relating to dolphin speech and behavior

There have been a number of recent books in the popular press that provide useful background on dolphin behavior and communications skills, although these cannot normally be taken as authoritative. Several of the most recent are reviewed below.

Also reviewed is a 2009 paper by May-Collado & Wartzok because of how it extends the frequency band of the dolphin family and stresses the need to review only the latest papers (generally from 2000) when planning a comprehensive study of dolphin vocalizations. They document whistle harmonics extending up to 110 kHz in the families of smaller dolphins.

This work will continue to focus on the vocalization and linguistic capabilities of the bottlenose dolphin.

U.1.2.1 The extensive behavioral observations of Dudzinski et al.

Dudzinski & Frohoff published a book in 2008 focused on the communications skills of dolphins from a behavioral perspective[17]. Their Appendix 1, entitled "A behavioral guide to dolphin signals," is invaluable in relating the physical activities and body language of dolphins to their potential acoustic communications. Only a small portion of the book relates to acoustic communications and their discussion of the morphology of the dolphin merely repeats the common wisdom. They do note (page 96) that "The term *language* can be tricky to define: it is more than simply a complicated communication system." And, "Language is something else,—an ability that is clearly tied to the capacity for complex thought." The book is an invaluable addition to the literature leading to settling the question of whether dolphins employ a *language*.

Page 15 provides a block diagram of the individual families of Cetacea and page 169 expands the list to the various species names found in the text. Their description of the hearing apparatus on pages 42 & 43 is dated based on this work. They note on page 96, "The term *language* can be tricky to define: it is more than a complicated communications system." They cite Pryor & Herman in their subsequent discussion of their elements of language. On page 110, they assert, "The study of dolphin cognition suggests that the dolphin mind is capable of many complex activities: problem solving, planning, manipulating abstract symbols, and in a more general sense, *thinking*."

---

[14]Kahn, D. (1967) The Codebreakers; the Story of Secret Writing. NY: Macmillan

[15]Herzing, D. (2010) SETI meets a social intelligence: Dolphins as a model for real-time interaction and communication with a sentient species. Acta astronautica vol 67, pp 1451-1454

[16]Herzing, D. (2006) The currency of cognition: Assessing tools, techniques, and media for complex behavioral analysis Aquatic Mammals vol 32(4), pp 544-553

[17]Dudzinkski, K. & Frohoff, T. (2008) Dolphin Mysteries: Unlocking the Secrets of Communications. New Haven, CT: Yale Univ. Press

U.1.2.2 "Are Dolphins Really Smart" as examined by Gregg

Gregg, an active researcher in other aspects of dolphin physiology, has recently published a book for the popular audience in which he appears to take the negative view with regard to dolphin language[18]. The book resulted from a Doctoral Thesis submitted in 2008 focused on eavesdropping by one dolphin on the echolocation of a second dolphin. The dominant investigator cited in the index to the book was Lily in the 1970's rather than more recent investigators, even those working in his immediate geographic vicinity of Shannon estuary and Cardigan Bay such as Hickey of the University of Wales. Over 1000 references are cited (including a quotation attributed to the 2009 variant of this appendix in #132 for chapter 5). The book discusses primarily observations of dolphin sounds and dolphin behavior in the natural environment with little data supporting the text.. The Gregg text appears to rely upon the homeocentric philosophy that any language must be directly translatable into English on a near word for word basis (a feature not shared with most other human languages). The book is an interesting read but is clearly not as well balanced as the author sought to achieve. His focus on the dolphin as "A most gentle mammal" (Chapter 6) overlooks the fact the animal is a carnivore and must eat to survive. Its hunting skills in littoral waters appear to be anything but gentle.

Gregg defines ten features of what he defines as "natural language," the language of humankind, in order to separate it from machine and computer languages (C++, etc.). He then defines any language attributable to non-human animals as being within an animal communications systems (ACS). In defining a natural language, Gregg uses the section title "Ten essential ingredients of language" and states, "I would like to shoehorn my own list of ten attributes necessary for language into this discussion so we can compare humans and dolphins in a more formal way. This list is cobbled together from the ideas proposed by the big thinkers I've been referencing." To say the linguistic community of big thinkers is fractured is an understatement. The factions have a great difficulty agreeing on precise definitions. It is suggested that Gregg might want to perfect the wording of his ten *potential* ingredients of language.

> The ten features of Gregg are quite comprehensive and are clearly present in advanced human languages. However, he makes no attempt to demonstrate all ten features are present in all human languages (still more than 6000 of them according to the linguist, McWhorter[19]). It is unlikely the early human languages and current primitive languages get the maximum five points in each of the ten categories. McWhorter has identified a variety of situations where primitive languages cannot describe a variety of ideas or concepts clearly, or in some cases at all (clearly failing Gregg's first criteria of limitless expression). His seventh requirement, arbitrariness, contains several problem clauses. Most romance languages and English as a minimum incorporate onomatopoetic sounds (there is even this word for sounds that mimic a natural sound) and a variety of "natural languages" incorporate symbols representing the relevant image in their written form (such as early Sanskrit and both early and current Chinese "natural languages." Those languages originally used pictographs exclusively for their symbols in their written form. These languages do not deserve a five on Gregg's rating scale of one to five for arbitrariness. Gregg also recognizes (page 136) that, "Every one of these ten ingredients can be found in the communication systems of non-human animals to some degree. There simply is no fundamental break in continuity between the cognitive abilities of human and non-human animals—"

> Gregg did not diagram his communications system, whether human or otherwise (See the above figure supplemented by Figure S.1.1-1, Figure S.2.1-1 and Section S.2.1 in Appendix S of this work). Several of the features he associates with language (4, 5, 8, 9 & 10) appear to have little or nothing to do with the mechanism of language, i.e., vocalization and hearing, associated with stages 6/7 and 4 of the above figure, but involve cognitive functions of the prefrontal lobe of at least the higher primates, stage 5. Even the feature described as recursion (feature 3) is more relevant to the cognitive function than to the communications function (Wernicke's area of stage 4 and stage 5 of the brain). The stage 5 cognitive elements issue high level instructions to the vocalization elements of stages 6/7 as documented in Figure S.2.1-1. His feature 8, related loosely to the expression of emotion, is more closely associated to the facial, somatomotor and skeleto-motor modalities discussed in **Section U.1.1.2** than it is to the vocalization modality. He described the expression of emotion (usually involuntarily) in conjunction with speech as a paralinguistic element.

Gregg's report on the behavioral characteristics of the dolphin in its natural ecological environment, appears to

---

[18]Gregg, J. (2013) Are Dolphins Really Smart?: The mammal behind the myth NY: Oxford Univ Press

[19]McWhorter, J. (2004) The Story of Human Language. Chantilly, VA: The Teaching Company (36 lectures over about 18 hours)

restrictive. The contemporaneous reports of Hickey[20] on dolphins hunting in the Shannon Estuary or of Santos[21] trailing dolphins hunting in the Bay of Sardo or when two normally geographically separated dolphin families meet and appear to struggle over dialect differences (also in Hickey) describe a more capable species. He has asserted several decades of personal experience observing dolphin behavior. But the book came only five years after his PhD Dissertation. It would be useful to have a table (similar to that on his page 158) showing the range of capabilities used by humans in a selection of "natural languages" on the left and the range of capabilities thought to be observed in "dolphinese" by active field investigators on the right. He notes in his endnote 183 that this might result in moving the goal posts a bit in favor of the dolphins. It would more clearly define where the terms "language" and "talk" used to describe human activity blends into the terms "communication" and "communicate" used to describe similar non-human activity (page 130).

In reporting on the potential for dolphinese, Gregg does not address the limited recording bandwidth, typically limited to that of human voice recording equipment prior to 2000 (typical maximum of 20 kHz or less[22,23] while the "whistles" of dolphins are known to extend up to at least 40 kHz). The spectrographic recordings of Killebrew et al. clearly show the rich structure of the calls of even neonate dolphins. Reynolds et al. also describe the variety of calls, other than signature calls, employed by dolphins. Kassewitz among others is now recording dolphin sounds throughout the 20 Hz to200 kHz range. Many of the sonograms of dolphins exhibit complexities similar to or exceeding those found in the more complex tonal human languages.

> The term whistle as used in dolphin research is not analogous to its use in human vocalization. In humans, a whistle is formed by unvoiced and nominally laminar air rushing through the properly positioned and rigidized lips. In dolphins a whistle is formed within the nasal passages without the participation of its closure at the blowhole. No air is expended into the environment. The dolphin creates a "whistle" using unvoiced, nominally laminar, air passing through a resonant portion of one of its nasal passages formed by the muscular rigidizing of that portion.
>
> At an age of 7-10, my cohorts and I were able to generate a call much louder than our whistle by opening our mouth, rigidizing our throat muscles and expelling unvoiced air through our larynx with considerable force. The sound could be easily heard 200 yards distant. This capability was lost entirely with the approach of puberty.

Gregg only touches lightly on the common observation of mimicry among dolphins and its potential connection to communications or language learning among dolphins. Mimicry is a major component of language learning among human neonates. Interspecies mimicry between dolphins and humans is challenging; dolphins do not have a vocal cavity even remotely similar to humans and humans can only hear the lower 15% of the frequency range used by dolphins. A dolphin imitating a human by "speaking" through its nose (blowhole) is effective showmanship at various marine parks but of little scientific value.

Wells & Scott have summarized some interesting properties of the signature whistles of bottlenose dolphins[24]. "The signature whistles of many male calves are similar to the whistles of their mothers, while those of female calves are not. This study continues to search for the causes of this difference, but it may reflect the different social patterns displayed by males and females. Both young males and females disperse from the female bands in which they were born, but the females often return to their natal band to raise their own offspring." This suggests a need to maintain individual names within the band. "Intriguing observations by Tyack suggest that dolphins may mimic the signature whistle of other dolphins, apparently as a prelude to social interaction" (? Conversation). *Tursiops* has the ability to vocally label objects, and it may be that signature whistles function as labels or 'names'". Kassewitz has recently recorded the initial sounds of a newborn dolphin for a matter of hours using a full 192 kHz recording capability.

---

[20]Hickey, R. (2005) Comparison of Whistle Repertoire and Characteristics Between Cardigan Bay and Shannon Estuary Populations of Bottlenose Dolphins (Tursiops truncatus): With Implications for Passive and Active Survey techniques *In fulfillment of a Master's Degree*, University of Wales (Bangor), GB

[21]Santos, M. Caporin, G. et. al. (1989) Acoustic behaviour in a local population of bottlenose dolphins *In* Thomas, J. & Kastelein, R. *eds*. Sensory Abilities of Cetaceans. NY: Plenum Press pp 585-598

[22]Dudzinski, K. Clark, C. & Würsig, B. (1995) A mobile video/acoustic system for simultaneously recording dolphin behavior and vocalizations underwater. *Aquatic Mammals* vol 21(3), pp 187-193

[23]Killebrew, D Mercado III, E. Herman, L. & Pack, A. (2001) Sound production of a neonate bottlenose dolphin aquatic Mammals, vol 27.1, pp 34-44

[24]Wells, R. & Scott, M. (1981) xxx *In* Ridgway, S. & Harrison, R. *eds*. Handbook of Marine Mammals, vol 6. NY: Academic Press pg 161

Such a bandwidth is critically important if the true character of the animals utterances are to be evaluated. The changes in the character of these sounds over even this brief time is worth considerable study from a language perspective[25].

U.1.2.3 The papers of the Herzing team

Herzing has led a team for several decades studying the dolphins primarily resident in a bay of the Bahamas. A large library of behavioral features as well as vocal recordings have been acquired. Herzing has provided a diverse set of academic papers on dolphin behavior and communications, most readily available from a Google Scholar search[26].

U.1.2.3.1 The 2013 TED talk by Denise Herzing

She provided an excellent overview of her work related to their communications skills in her 2013 TED talk,

http://www.ted.com/talks/denise_herzing_could_we_speak_the_language_of_dolphins.html

She demonstrates the ability of the Atlantic spotted dolphin, *Stenella frontalis*, and researcher to communicate using only sounds generated on command by a computer in the hands of a scuba diver. They used artificial messages of similar form to normal dolphin tonal sounds using frequencies up to 110 kHz. They were working with Georgia Tech in the USA. So far, the extent of their messages are limited to only a few words (noun verb and object/target). While this activity does not meet Gregg's ten criteria, it is likely Gregg's criteria need modification to more properly define language as an adjunct to oral communication (similar to any other code) used within a wider spectrum of communication. This places his term language (in the singular) in the category of languages (in the plural) with each language having its own lexicon of sounds. In parallel, most languages have an associated written form that is a valuable adjunct to the language itself. In the American Heritage Dictionary, they describe a lexicon as "The morphemes of a language considered as a group."

Figure U.1.2-1 reproduces a frame at 6:38 minutes in Herzing's TED talk. She challenges the audience to identify which of the three spectrograms are from the dolphin and which are from humans.

---

[25]Kassewitz, J. Weber, M. Fulton, J. & Lingenfelser, R. (2007) A Preliminary Analysis of the First Hour of Sound after Birth Produced by a Dolphin Calf (Tursiops truncatus) Internal report available from www.speakdolphin.com

[26]Herzing, D. http://www.scholar.google.com

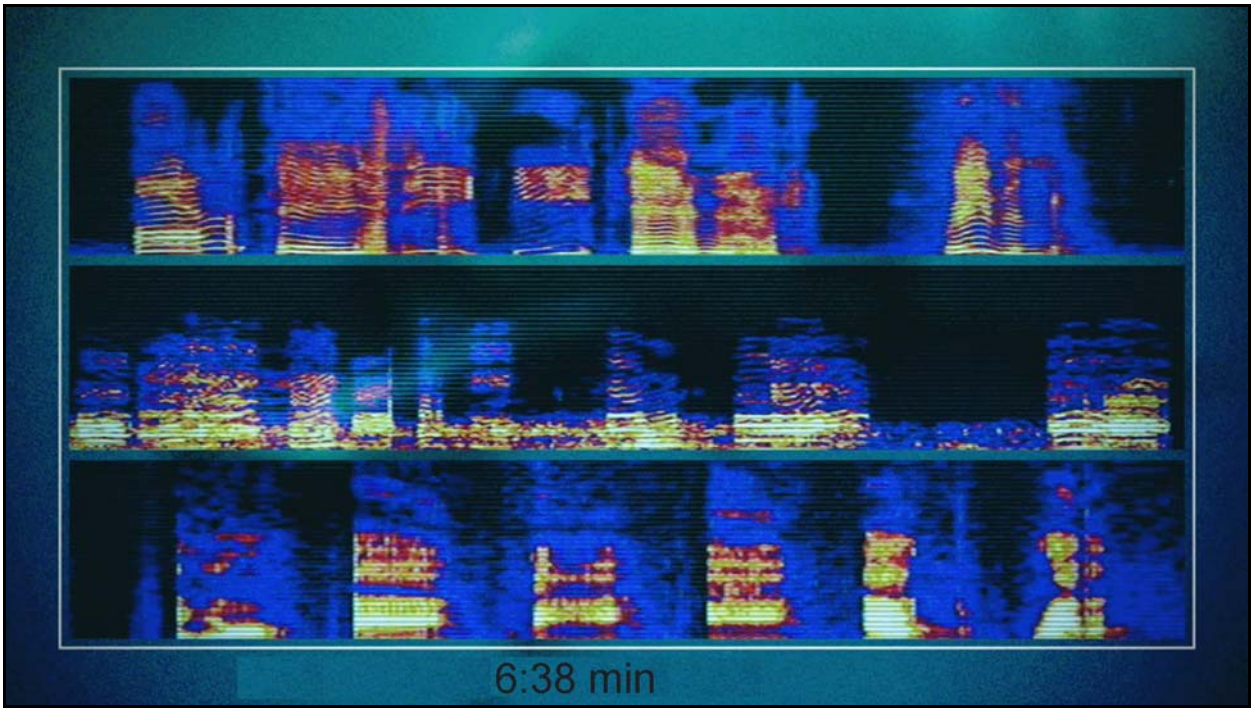

Figure U.1.2-1 Dolphin versus human spectrograms. From Herzing TED presentation, 2013.

She finally confirms the middle sample is from the dolphin. But, from the perspective of the visiting Martian introduced by Gregg, it would be extremely difficult to assert that one of these is merely a record of communications and the other two are truly more sophisticated records of speech (language).

Figure U.1.2-2 shows another fairly obvious feature of dolphin communications. The names dolphins assign themselves are clearly tonal in character (like in Chinese, not the Romance languages). The newly born dolphins traditionally take names similar to their mother's if female but differ more substantially if male. As noted below Kassewitz et al. actually observed a neonate dolphin initially mimic its mother's name but then adopt its own variant within a few hours of birth.

Within the context of Herzing's TED report, and under the assumption that dolphins exhibit a language capability, it appears Herzing introduced at least two new words and probably one new synonym into that language. The sounds used were tonal sweeps and generally did not contain any harmonic overtones. The dolphins would probably agree these were too guttural to be found in their natural lexicon. The words would describe a piece of rope and a piece of colorful woven fabric we know as a scarf. The synonym would be a new word for the seaweed that they already have a word for and humans identify as Sargassum. Human scientists have yet to identify the dolphin word for this material but they should be able to do this based on the code breaking procedures outlined earlier.

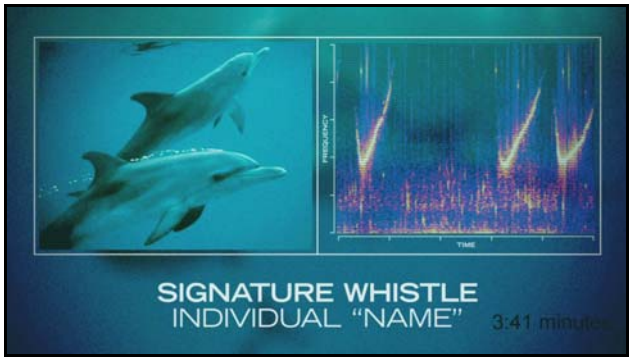

Figure U.1.2-2 The tonal quality of names in the dolphin world. Each dolphin of the bottle-nosed variety assumes a name based on its mother's name if a female but more distinctive if a male. From Herzing TED presentation, 2013

It is interesting that Herzing asserts the dolphin language exhibits a high entropy (suggesting it approaches human languages) without identifying a single word in that language. The calculation must have been made strictly on the variety of sounds in their lexicon (a standard human code breaking technique). Herzing subsequently confirmed her reference to entropy was based on the work of McCowan and colleagues explored below.

If a viable entropy was available and the necessarily accompanying set of frequency of occurrence tables for the

sounds, it should be possible to isolate the natural dolphin words for sargassum as a starting point. It is obvious from the imagery that the dolphins were more or less fighting to transport the pieces of rope, scarves and seaweed to the humans as requested acoustically. During these intervals, it is likely the appropriate words were used between the dolphins. When swimming in convoy, it is also likely that the dolphins share a restricted vocabulary. Herzing reported the synchrony of the dolphin sounds during these intervals. At least portions of this vocabulary should be decodable.

U.1.2.3.2 Other academic papers of the Herzing team

A 1996 Herzing paper classifies a large number of sounds according to the behavior of the dolphin at the time and using a recording capability described as flat to 22 kHz [27]. The paper includes a wide variety of spectrograms (of restricted bandwidth) and discussion of the need for a broader framework in which to analyze the results. An excellent bibliography is provided (where Santos is shown as "dos Santos.") along with a list of agencies interested in cetacean research.

A 2003 paper expands the recording bandwidth of the team to 100 kHz after filtering[28], . They note,

•"Efforts to study the social acoustic signaling behavior of delphinids have traditionally been restricted to audio-range (<20 kHz) analyses."
•"The burst pulses of both species were found to be predominantly ultrasonic, often with little or no energy below 20 kHz."
• "Many whistles had harmonics that extended past 50 kHz and some reached as high as 100 kHz."
• "The relative amplitude of harmonics and the high hearing sensitivity of dolphins to equivalent frequencies suggest that harmonics are biologically relevant spectral features."

As a result, dolphins hear little of what adult male humans say, and dolphins produce high frequency features in their speech that humans can not hear. Listening to a dolphin is similar to listening to the voice of a soprano over a common telephone. **Figure U.1.2-3** showing their figure 1 provides a good summary of the waveforms used by dolphins.

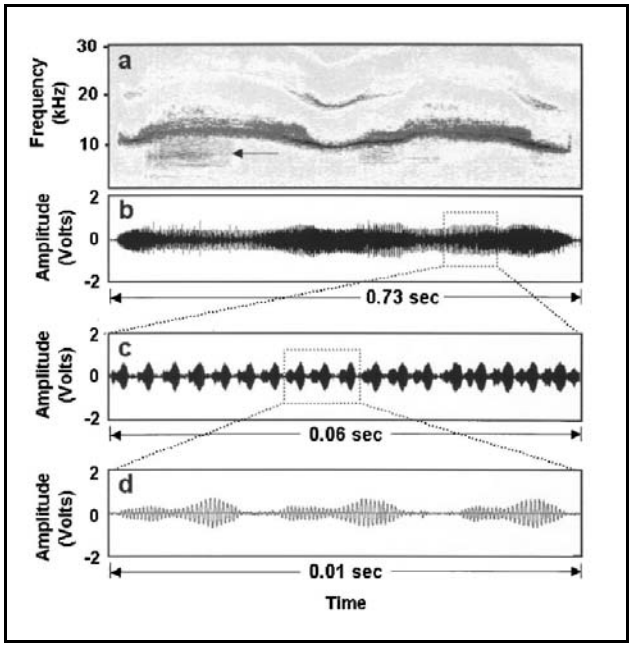

Figure U.1.2-3 Amplitude-modulated spotted dolphin whistle. A; sonogram. B; waveform versus time. and expanded versions of portions of the waveform. From Lammers et al., 2003.

[27]Herzing, D. (1996) vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus. Aquatic Mammals* vol 22.2, pp 61-79

[28]Lammers, M. Au, W. & Herzing, D. (2003) The broadband social acoustic signaling behavior of spinner and spotted dolphins *JASA* vol 114(3), pp 1629-1639

Figure U.1.2-4 shows the rich harmonic content of the typical dolphin emanation. The variation in this content would correspond to similar spectrograms for humans when their mouths were shaped differently in speech.

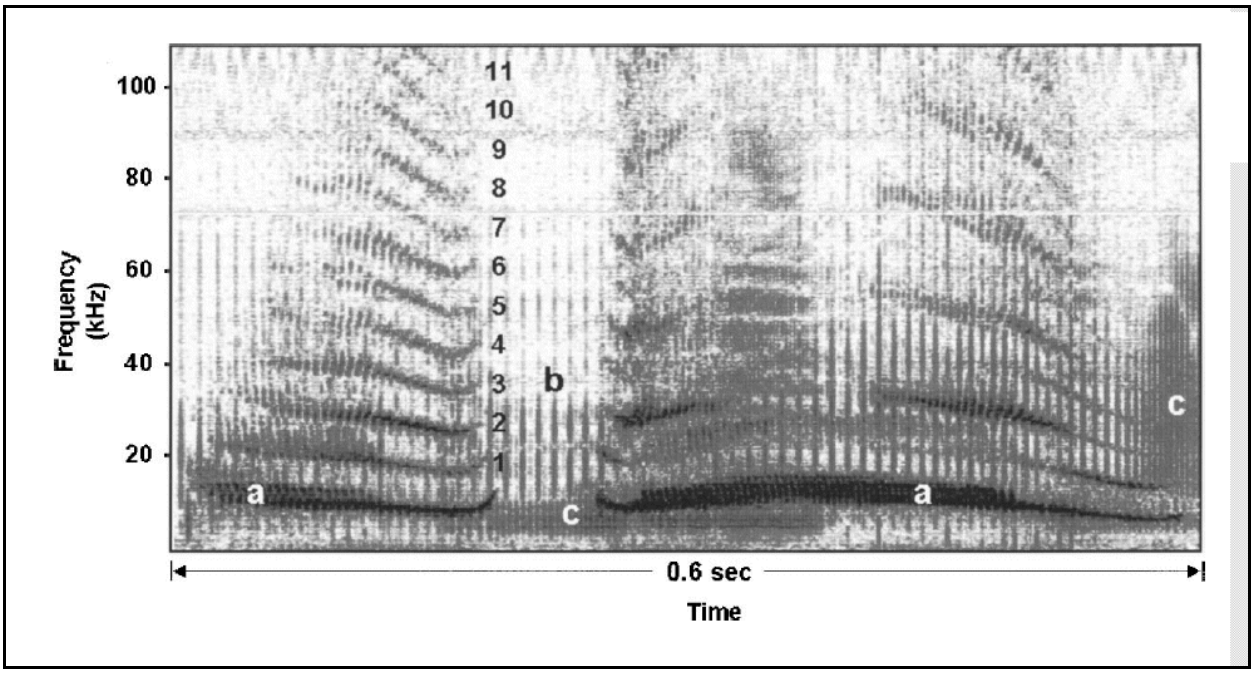

Figure U.1.2-4 A sequence of two spotted dolphin whistles (a) mixed with echolocation clicks.(b)! and burst pulses(c).
The first whistle contains 11 harmonics (numbered). Only the first harmonic is compatible with the hearing of (young) humans. From Lammers et al., 2003.

Note the regions marked burst clicks do not reproduce the actual waveforms in this image as the actual frequency of these clicks typically rises to around 150 kHz. The spectrograms typically indicate the envelope of these clicks.

Figure U.1.2-5 shows Figure 3 of Lammerts et al. It shows additional variations in harmonic content among dolphins suggesting verbalization. Such variations in human spectrograms are clear examples of different emphasis in the pronunciations of closely related words of human language. Frame (c) can be interpreted as representing two distinct words within the 0.65 sec interval (roughly twice the duration of the (a) and (b) spectrograms. There is much additional information in the Lammers et al. paper impacting on the question of whether these variations characterize the lexicon of language. Figure 6 compares the wide band spectrogram with earlier band-limited spectrograms and notes the virtual absence of evidence of a significant signal in the narrow band view.

Lammers et al. conclude with; "If we presume that dolphins pay attention to the harmonic composition of whistles and if we accept that burst pulses play an important social role, then the evidence presented here indicates that there is considerably more to the social acoustic signaling behavior of spinner and spotted dolphins than meets the human ear. In future efforts to better understand the function of whistles and burst pulses we will need to more fully explore and appreciate their design. This will require that we take into account their broadband patterns of production and adopt methodologies that reflect the dolphin's auditory acuity. Recording signals in a manner consistent with how they are produced and ultimately heard by their intended listeners will be an important key to future efforts to accurately match them with their social context."

A 2006 paper[29] explored the broader aspects of the social interaction of dolphins, noting; "These specific actions (e.g., gestures, vocalizations, gaze, body/head orientation, etc.) represent the potential media of information, or currency of cognition, available to dolphins. Such media are real-world, observable, and measurable signals through detailed behavioral analysis (i.e., Micro-ethology)." "Micro-ethology, or the detailed analysis of behavior in context, is the area of focus for this paper." This paper provides a wealth of information related to the communications skills of dolphins and divides the data into many clearly defined categories.

Indications are this additional work will lead to a further definition of these sounds and their entropy in the context of a real language.

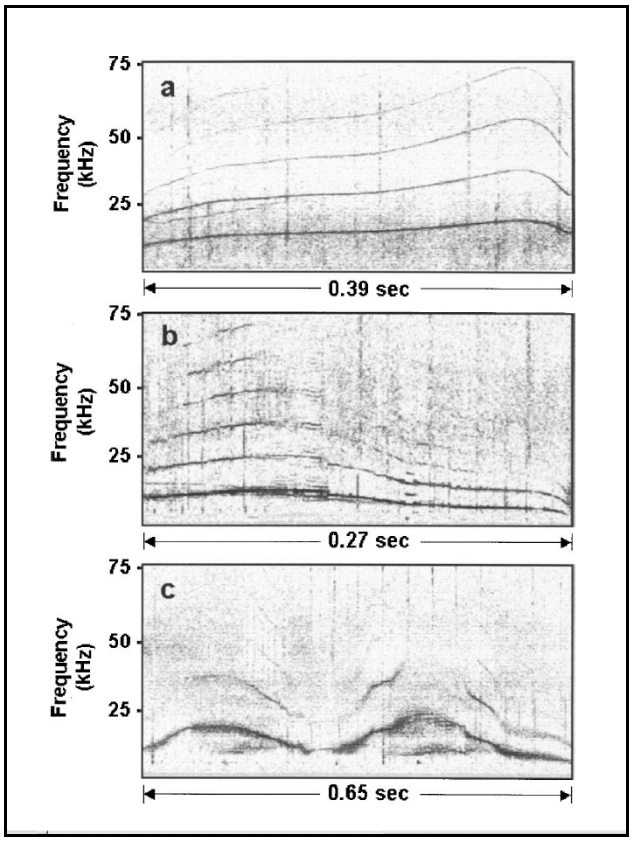

Figure U.1.2-5 Examples of the variation found in the harmonic composition of whistles; (a) spinner dolphin whistle with continuous harmonics throughout signal; (b) spotted dolphin whistle with harmonics emphasized on concave portion of the contour; and (c) spotted dolphin whistle with harmonics only on the slopes of the contour. From Lammers et al., 2003.

U.1.2.4 The behavioral observations of Turner at Shark Bay, Australia

Turner has published a popular book aimed at children on the lives of a large cohort of dolphins living in a free environment located in western Australia[30]. Their field location corresponded to Monkey Mia campground along the shore of Shark Bay, Western Australia.

The book presents many scenarios highlighting various features of dolphins routine experiences. An interesting observation is that some dolphins appear to change their personal names when they become associated with different groups. The new names may be a nick-name or merely a prefix or suffix added to the previous name. The name changing may be a relatively common occurrence in what Turner describes as the fission-fusion society employed by

[29]Herzing, D. (2006) The Currency of Cognition: Assessing Tools, Techniques, and Media for Complex Behavioral Analysis *Aquatic Mammals* vol 32(4), pp 544-553

[30]Turner, P. (2013) The Dolphins of Shark Bay. NY: Houghton-Mifflin

dolphins (continually forming and dissolving small groups within a larger cohort). There are many other anecdotes from a trained observer suggestive of a structured language (or minimally communications) between various groups of dolphins. Turner also describes the use of tools be dolphins in considerable detail (accompanied by many interesting pictures). She makes a number of suggestions concerning what dolphins may be saying to each other, depending on the situation they find themselves in.

For anyone attempting to discover the lexicon and syntax of dolphin communications, the Turner book can be useful. It is focused on the social aspects of communal living. It does not contain any signal waveforms.

U.1.2.5 The extended frequency range of the Guyana dolphin, (*Sotalia guianensis*)

With the improvements in hydrophone and recording capabilities since 2000, new information has become available about the vocalizations of other little known dolphins. May-Collado & Wartzok have reported on the Guyana dolphin and provided a bibliography of wider bandwidths recorded from a variety of dolphins[31]. They report the higher formants of whistles in this species extending up to at least 130 kHz based on a statistically adequate cohort size. **Figure U.1.2-6** reproduces a spectrogram from the Guyana dolphin. The caption is quoted directly. Its wording would be modified slightly based on the terminology developed later in this Appendix. Many of the whistles appear to consist of multiple phonemes not necessarily associated with a specific "signature whistle." Formants are clearly visible extending up to about 110 kHz at 0.27, 0.33 and 0.48 seconds There are also formants (or stray signals) shown at 4500 and 7500 Hz during much of the recording(s). Such high frequency components may be useful for short distance communications among river and littoral dolphins but are probably of little use to wider ranging marine dolphins (see the acoustic attenuation characteristic of sea water in Appendix L). It is clear from the figure that complete whistles can be missed if recording equipment limited to 24 kHz is used in investigations.

> Dawson asserted in 1991 that Hector's dolphin generated no whistles based on his measurements using test equipment probably limited to 20 kHz in that time period[32]. Hector's dolphin is one of the smallest of the dolphins (mature length below 1.5 meters) and could quite possibly generate whistles at frequencies above the range that Dawson was able to record.

The figure also illustrates a variety of unvoiced sounds during intervals between whistles that may be very significant in any speech capability of these dolphins. The authors describe these intervals as "break of contour" but they are not necessarily periods of silence. The "break" illustrated in one of their 2007 papers was about 30 ms. A "break" could be occupied by either a voiced or unvoiced consonant. An amplitude profile versus time accompanying the spectrogram would provide more information on this point.

The Guyana dolphin, *S. guianensis*, also known as the costero in Spanish, is a member of the oceanic dolphin family (*Delphinidae*). Physically, it resembles the bottlenose dolphin. However, it and its close freshwater relative the tucuxi, *S. fluviatilis*, differ sufficiently from the bottlenose dolphin to be placed in a separate genus, *Sotalia*. The mature Guyana dolphin is slightly longer than a grown man, 2.1 meters (6.9 feet).

---

[31]May-Colladoa, L. & Wartzok, D. (2009) A characterization of Guyana dolphin (*Sotalia guianensis*) whistles from Costa Rica: The importance of broadband recording systems *J Acoust Soc Am* vol125(2), pp 1202–1213

[32]Dawson, S. (1991) Clicks and communication: the behavioral and social contexts of Hector's dolphin vocalizations *Ethology* vol 88, pp 265-276

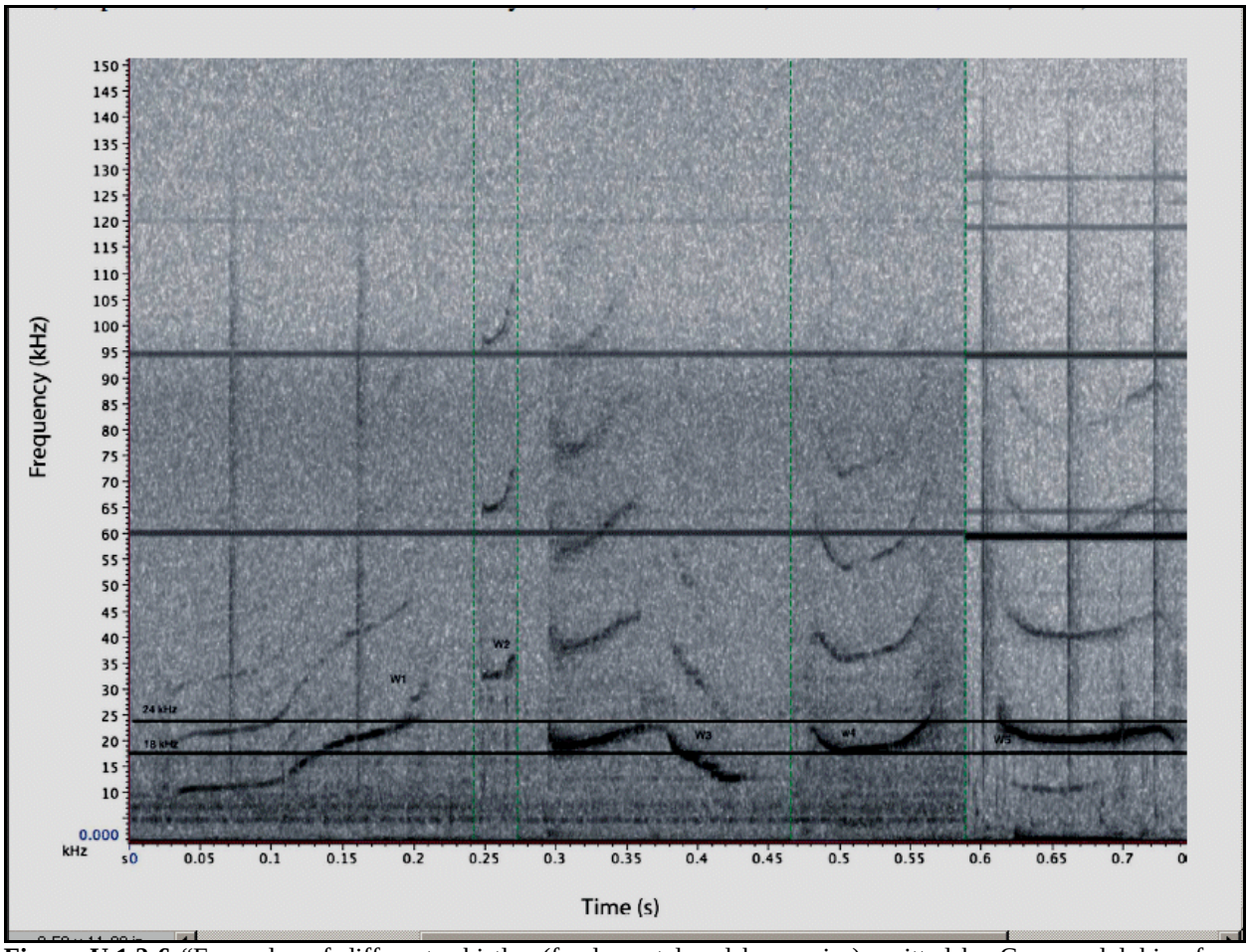

Figure U.1.2-6 "Examples of different whistles (fundamental and harmonics) emitted by Guyana dolphins from Gandoca-Manzanillo Wildlife Refuge, Costa Rica. The horizontal lines represent the mimicked bandwidth limits at 18 & 24 kHz." The lines at 60, 95 and higher kHz and the "W" labels were not addressed in their text. See text. From May-Collado & Wartzok, 2009.

They also provided a summary of the maximum frequencies noted by other researchers with extended test equipment capabilities.

> "Whistles are the most studied sound type, and several whistle acoustic variables have been recently described from Brazilian populations Monteiro-Filho and Monteiro, 2001; Azevedo and Simão, 2002; Erber and Simão, 2004; Azevedo and Van Sluys, 2005; Pivari and Rosso, 2005; Rossi-Santos and Podos, 2006 . Monteiro-Filho and Monteiro 2001 first described Guyana dolphin whistles as low in frequency up to 6 kHz but a more extensive study revealed a much wider whistle frequency range 1.34–23.89 kHz Azevedo and Van Sluys, 2005 . However, as noted by Azevedo and Van Sluys 2005 , some of the recorded whistles looked "cut off" by the upper frequency limit of their recording systems, suggesting Guyana dolphins can emit high frequency whistles exceeding the 24 kHz recording limit. Several toothed whale species have been shown to emit whistles with high fundamental maximum frequencies, up to 24 kHz in spinner dolphins and Atlantic spotted dolphins e.g., Lammers et al., 1997, 2003; Oswald et al., 2004 , 29 and 41 kHz in bottlenose dolphins Boisseau, 2005; May-Collado and Wartzok, 2008 , 35 kHz in white-beaked dolphins Rasmussen and Miller, 2002; Rasmussen et al., 2006 , 24 kHz striped and common dolphins Oswald et al., 2004 , and 48.10 kHz in botos May-Collado and Wartzok, 2007."

They close their introduction by noting, "The importance of selecting recording systems with bandwidth appropriate for the study species is fundamental in understanding dolphin whistle structure and its geographical variation Bazúa-Durán and Au, 2002; Au et al., 1999 , as well as for species classification Oswald et al., 2004 ."

> Their methodology was to record "Seven standard whistle variables were measured on the fundamental frequency of each: starting frequency (SF), ending frequency (EF), minimum frequency (MinF), maximum frequency (MaxF), delta frequency (DF=MaxF-MinF), duration (s) , and number of inflection points." "Peak

> frequency (PF) was also measured and is defined as the frequency at which maximum power occurs (azúa-Durán and Au, 2002, 2004; May-Collado and Wartzok, 2007) and number of harmonics (.g., Wang et al., 1995; Erber and Simão, 2004). Whistle contours were categorized as ascending, descending, ascending-descending, descending-ascending, and constant in frequency, sine, and others."

They carried out a significant study comparing wide band and narrow band recordings of the same species. They subsampled their wide bandwidth data to generate narrower band recordings that they use to "mimic" the narrowband recordings of other researchers. Their conclusion,

> "This study shows that limited bandwidth distorts the understanding of Guyana dolphin whistle frequency variables, particularly in whistle maximum, ending, and 3 4 frequencies. Whistles selected to mimic narrowband recordings systems with bandwidths of 18 and 24 kHz limited the characterization of the whistle frequency span of Guyana dolphins to a portion of the actual frequency range see Fig. 3 . For instance, about 73 whistles out of 422 had maximum frequencies that extended beyond the 24 kHz limit, and additional 14 whistles had minimum and starting frequencies above 24 kHz, *and would have been completely missed by narrowband recording systems* (italics added). In addition, most of the harmonics would have been missed with narrowband recording systems. In order to properly document whistle repertoire of Guyana dolphins including harmonic components a recording system with a bandwidth of at least 150 kHz is necessary."

Their conclusions ended with "this study provides evidence supporting the hypothesis that whistle frequency variables increase with latitude. Future studies on Guyana dolphin whistles should employ recording systems with bandwidth up to 50 kHz for whistle fundamental and up to 150 kHz when considering high order harmonics to ensure the inclusion of the entire whistle repertoire."

U.2 The broad view of inter and intra-species communications

It is widely recognized that many species from different phyla exhibit behaviors that express an internal or symbolic processing of external reality (Roitblat, Bever, & Terrace, 1984). There is also no social behavior without communication (Evans & Bastian, 1969); few species exist which do not possess a substantial repertoire of symbolic communicative devices such as vocalizations, displays, postures, and/or gestures. But these two symbolic systems may remain separate systems with little conceptual or feature overlap in most species.

> The interspecies communications literature between humans and animals is voluminous and is generally reported in volumes concerned with ethology[33], the behavior of the animal (generally in response to human voice commands).

Lilly was an early investigator of dolphin vocalizations and potential interspecies communications. His 1978 book provides an extensive bibliography of his earlier work[34]. The material is primarily observational and lacked our current knowledge of dolphin physiology. As an example, he speaks of the dolphin making sound through its pharynx. But the pharynx is not part of the "rebreathing" apparatus that allows the dolphin to make sounds under water. In fact, the dolphin does not make sound via its mouth even when on the surface. He notes humans communicate through facial expressions, gestures of the body, physical contacts and the production of sounds in the mouth, throat and larynx. A reading of his material did not locate a similar listing for the dolphin. A survey of the more recent literature confirms the dolphin employs essentially the same list of communications techniques (with portions of the nasal pathway replacing the mouth). Lily did note that dolphins release bubbles and turn their flippers perpendicular to the line of sight of adversaries when making a show of dominance. Such gestures are clearly visible to the high visual acuity of the dolphin.

Very recently, the NOVA series on PBS.org have explored the cognitive and language potentials of the bottlenose dolphin in comparison to other animals[35]. Many of the behaviorists introduced below participated in the preparation of this video material.

U.2.1 An evaluation of dolphin cognitive abilities

Before attempting to understand Dolphinia (or Truncatese), it is important to circumscribe their potential speech capability. One aspect of that capability is the ability of dolphins to cogitate at the level of both concrete and

---

[33]Bateson, P. Klopfer, P. & Thompson, N. *eds*. (1980-1993) Perspectives in Ethology. NY: Plenum Press *A series*.

[34]Lilly, J. (1978) Communication between Man and Dolphin. NY: Crown Publishers

[35]NOVA (2014) Inside Animal Minds: Who is the smartest? http://video.pbs.org/video/2365229680/

abstract ideas. Herman et al. have reported on an extensive study of two dolphins trained to interpret and respond to a series of hand signals or a series of underwater acoustic whistles[36]. Their results are discussed in **Section U.4.7.1** and are quite illuminating. Not only were the animals able to understand imperative sentences consisting of object + action + indirect object, they were able to respond properly to compound sentences, lexically novel sentences, structurally novel sentences and semantically reversible sentences. Interestingly, when given an impossible instruction the first time, they searched extensively for the missing object. After they were made aware that the instruction might be false (and with very little reinforcement), they performed a complete but more modest search before returning to their station and contacting the NO paddle. It was also clear that their interpretation of the object PERSON referred to the symbolic person, as it mattered little which attendant took the person-position next to the dolphin tank.

Their training extended up to five word instructions following a signature whistle describing which dolphin was to implement the instruction. This signature whistle was a computer-generated duplicate of the whistle the dolphin used to identify itself prior to the training program. One dolphin processed 168 sentences of various types with much higher performance than suggested by chance. The other dolphin processed 368 mixed sentences with similar performance. The specific words used are described in the Herman et al. paper. The acoustic words extended up to 35 kHz. The goal was to make the words members of an orthogonal set by varying their frequency, duration and modulation.

The lexical, syntactic and semantic breadth of language the dolphins have been demonstrated to understand, based on pragmatic studies, insures, in this analysts eyes, the inherent ability of the dolphins to communicate using language. The immediate tasks are two-fold. First is an attempt to isolate and define their phonetic symbol set. Second is to attempt to determine the range of objects and actions a dolphin is likely to use in everyday communications before attempting to decode the available recordings of these communications. These tasks are obviously closely intertwined. More details relating to their overall capability begin in **Section U.4.3.2**.

The very recent observation of dolphins forming toroid rings from air bubbles and playing with them extensively underwater has come to a shock to many experienced dolphin observers.

https://www.youtube.com/watch?feature=player_embedded&v=wuVgXJ55G6Y

This capability is impressive and illustrates both the physical dexterity and the cognitive ability of the bottlenose dolphin. The capability has been observed to be a learned skill (that is learned quite rapidly from other members of a cohort). These capabilities are far beyond those of a human child of less than eight or ten years of age.

U.2.1.1 Teaching dolphins a new structured acoustic language EMPTY

The Herman, Richards & Wolz paper is extensive. In essence, they taught a dolphin a totally acoustic second language using artificial symbols understandable to the dolphin. See **Section 4.7.1.1**.

U.2.1.2 Mirror self-recognition (or awareness) in dolphin

Previously, only humans and the great apes (specifically the chimpanze) have been shown to exhibit mirror self-recognition. Reiss & Marino, of the New York Aquarium, have performed a very statistically convincing set of experiments that show the bottlenose dolphin will use a mirror to examine areas of its body that it believes were marked, or sham-marked by a trainer[37]. Two animals were tested under carefully defined conditions involving carefully constructed protocols. The animals, as usual surprised the test team on a number of occasions. In several cases, the animals left their stations immediately after marking, without waiting for the release signal, in order to examine the mark as soon as possible. A feature of the experiments was the length of time the animals spent examining an actual mark compared to the much shorter time spent after recognizing that they had been sham marked and there was nothing to examine in detail. In some cases, they exhibited highly unusual swimming postures in front of the mirror in order to examine a mark on the right pectoral fin. In a one-time experiment, the animal was marked on the tongue. It immediately swam to the mirror and engaged in a mouth opening and closing sequence never before observed by him during the study

U.2.1.3 Initial definition of a language protocol for dolphins

---

[36]Herman, L. Richards, D. & Wolz, J. (1984) comprehension of sentences by bottlenosed dolphins *Cognition* vol 16, pp 129-219

[37]Reiss, D. & Marino, L. (2001) Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence *PNAS* vol 98(10), pp 5937-5942

Crystal has provided a comprehensive text on human language from the perspective of a linguist[38] and has provided a useful bibliography. He defined a "faculty of language" that can easily be confused with a "facility of language" (suggesting the more physical aspects of the communications capability). Here the term faculty of language will be replaced by a language protocol (more suggestive of the non-physical aspects of communications.

Crystal did not include the symbol set in his discussion of human language, although he did recognize there were different symbol sets among the written forms of natural languages. Here the focus is on the characteristics of a potential non-written language where a definition of the symbol set is a major challenge. Crystal also did not address the pragmatics associated with a language; The definition used here (2) is more explicit than that of Johnson[39] (1),

Pragmatics– 1. What gets accomplished as a result of communicative acts.
2. The study of language as it is used in a social context and affects the interlocutors and their behavior.

With these modifications, a language protocol includes the following elements,
- A symbol set
- A dictionary
- A set of syntactic rules
- A set of semantic rules
- A list of pragmatic results

This list has been incorporated into a figure modified from Crystal (**Section 4.9.1**)

U.2.1.4 The status of recursion in linguistic

The concept of recursion has been employed in mathematics for a very long time and follows a very specific set of rules. Chomsky appears to have introduced the term into linguistics in 1956. Considerable discussion has ensued ever since, particularly with the arrival of the computer age and major advances in linguistic theory that ensued. Hauser addressed the subject in 2001[40]. Hauser, Chomsky & Fitch (HCF hereafter) addressed the subject again in 2002[41]. Pinker & Jackendoff (PJ) provided a set of counter arguments in 2005[42]. Lobina attempted to provide an overview of the case for both sides[43]. He noted the background paper by Tomalin of 2007.

Lobina noted that multiple *rules* that can be implemented within the category of recursion. A clear definition of recursion within linguistics is difficult to locate because the word implies an open ended set of rules. Alex B. on the website, Linguistic Stack Exchange, has offered recursion "is a phenomenon where a linguistic rule can be applied to the result of the application of the same rule." Linguistics Stack Exchange is a question and answer site on the internet for professional linguists and others with an interest in linguistic research and theory. There are apparently both nested and tail-based (serial) rules as a minimum. Some rules apply to the cognitive generation of the sentences (or their higher level neural instructions to stage 6 command generation circuits) and others to the expression of the sentences by stage 7 circuits and the motor muscles of the articulation mechanisms.

The Lobina paper uses language totally unique to the linguistic specialty. It is focused on the architecture of

---

[38]Crystal, D. (2006) How Language Works. NY: Overlook Press

[39]Johnson, C. (1993) Animal communication via coordinated cognitive systems *In* Bateson, P. Klopfer, P. & Thompson, N. *eds*. (1980-1993) Perspectives in Ethology. NY: Plenum Press vol 10, chapter 7

[40]Hauser, M. (2001) What is so special about speech? Chapter 23 *in* Dupoux, E. *ed*. Language, Brain, and Cognitive Development. .
http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.207.8203

[41]Hauser, M. Chomsky, N. & Fitch, T. (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* vol 298, pp 1569-1579

[42]Pinker, S. & Jackendoff, R. (2005) The faculty of language: what's special about it? *Cognition* vol 95, pp 201-236

[43]Lobina, D. (2010) Recursion and linguistics: an addendum to Marcus Tomalin's reconsidering recursion in syntactic Theory. *Interlinguista* XX filcat.uab.cat/clt/XXIVAJL/Interlinguistica/.../LobinaREVF.pdf

language as interpreted by the individuals named in this paragraph. He notes, "Chomsky himself has been rather clear that it is the Merge operation that applies recursively," (Chomsky 1995:226). Lobina then proceeds to differentiate between *external* merge and other forms of Merge. In addressing the contributions of (PJ), he notes, "First, they define it as a "procedure that calls itself", a characteristic we have already noted. It then adds, however, that it also applies "to a constituent that contains a constituent of the same kind" (PJ: 203). The latter is surprising, as it constitutes a move from referring to procedures as recursive to describing data structures (constituents) as such. This is technically a mistake, and it is one that is rather widespread in linguistics. The employment of rewrite rules may account for this usage." Lobina goes on, "It is true that HCF's discussion of recursion is vague and unclear, somewhat justifying Tomalin's claim that this is the result of "the confusion that has enveloped the term 'recursion' since the late 1930s" (M. Tomalin, 2007: 1796).

Lobina does give four examples from Parker, "She starts by pointing out how vague linguists; definitions of the term 'recursion' have been . . ." Proceeding, "The first example is a sentence she might call iterative, while the last three (taken from her essay) are examples of so-called recursive sentences. (9) is a case of nested recursion, and the last two are cases of tail recursion.

(8) The nice vampire loves the girl and the girl loves the vampire and the friendly ghost loves the vampire and…
(9) The mouse the cat the dog chased bit ran.
(10)The man that wrote the book that Pat read in the café that Mary owns.
(11) John's brother's teacher's book is on the table."

Lobina closes his discussion with several very specific paragraphs addressed to the professional linguist.

> "In the 1950's, linguists correctly employed recursion in reference to specific rewrite rules. Ever since these rules were eliminated from linguistic theory, however, most linguists have used recursion, rather puzzlingly, to refer to those structures generated by recursive rewrite rules. This may well be the unfortunate legacy of employing rewrite rules. Other linguists have focused on the operation Merge, which is welcome, but a satisfactory treatment has yet to be provided. There seems to be a strong tendency to confuse hierarchically structured representations with recursion. Even though hierarchical data structures call for recursive mechanisms, the latter do not automatic arise because of the former."

The above examples demonstrate that no matter how generated or expressed, a level of recursion beyond three is largely unintelligible to humans without resorting to pen and paper to diagram the sentence.

Yip[44,45] has takes considerable exception to the premises of Hauser, Chomsky and Fitch in their 2002 paper arguing that their exclusion rules virtually eliminate the use of a language protocol by any non-human species.

U.2.2 Previous efforts to document the phonemes in dolphin language

Markov & Ostrovskaya attempted to establish the grammar of the bottlenose dolphin[46]. They provided a broad description of the individually identified sounds of the animal but without any frequency scale on the spectrograms or amplitude scale on their temporal plots. They worked diligently on over 300,000 signals to determine many structural elements of a potential language. They did note the ability of the dolphin to use multiple sound generators simultaneously to create very complex sound structures (page 610). It may be necessary to segregate some of these signals as related to echolocation rather than intra-species communications. This applies particularly to what they describe as synchronous operation of two pulse generators. They did offer a transcription method to describe specific messages.

Morgan attempted to identify the individual sounds of Beluga Whales and play them back to the animals in a variety of synthetic sequences[47]. The reaction did not lead to significant new understanding.

Li, et al. have recently provided a stimulating introduction that suggests different species of *Cetacea* may speak

---

[44]Yip, M. (2006a) The search for phonology in other species *Trends Cogn Sci* vol 10, pp 442-446

[45]Yip, M. (2006b) Is there such a thing as animal phonology? Linguistics Research Center, Univ. California, Santa Cruz pp 311-323

[46]Markov, V. & Ostrovskaya, V. (1989) Organization of communication system in *Tursiops truncatus montagu. In* Thomas, J. & Kastelein, R. *eds*. Sensory Abilities of Cetaceans. Op. Cit. Pp 599-611

[47]Morgan, D. (1979) The vocal and behavioral reactions of the Beluga, Delphinatperus leucas, to playback of its sounds *In* Winn, H. & Olla, B. *eds*. Op. Cit. pp 311-343

individual languages[48] (and as noted elsewhere in this discussion, different pods of bottlenose dolphins appear to exhibit dialects).

In any comprehensive program to study dolphin language, it is important to consider the critical nature of the filter bandwidth used to prepare spectrograms as part of the study. Section 8.1.4.1 of this authors work on hearing is relevant, www.hearingresearch.net/pdf/8Manipulation.pdf . It shows the same English sentence analyzed with two different bandwidths. The features highlighted are quite different.

U.2.2.1 Dictionary of relevant human ideas in words

A few definitions are in order. Using the terms from Pinker's Glossary of 1994;

morpheme– The smallest pieces into which (generally written) words can be cut: un-micro-wave-abil-ity.

phoneme– One of the units of sound that are strung together to form a morpheme, roughly corresponding to the letters of the alphabet: b-a-t, b-ea-t, s-tou-t.

Two additional definitions will also be important;

Speech– The actual sequences of sounds used between a speaker and listener of when communicating over a media, either air or water in the cases of interest.

Language– The protocol underlying speech and comprising the lexicon, grammar and more specialized semantics required to be shared between the speaker and listener to achieve adequate understanding. The protocol of human speech is normally defined as natural language (such as English or Chinese).

Dialect– Variation in the protocol/language used by speaker and listener of different cohorts of a species while maintaining communications (at a potentially reduced understandability).

Prosodics– the metrics of speech and singing, an element of the protocol/language.

Yip described the current stage of understanding of phonology (ca. 2006) as it applies to animal communications[49],[50]. While not defining phonology specifically, he does provide a range of sources for such a definition[51].

> **Phonology–** appears to focus on the mechanical ability of vocalization and its limited ability to articulate certain potential combinations of phonemes. It appears to be a detailed level of syntax primarily of interest to the dedicated linguist. Goldsmith[52] described the goal of phonology as, "the development of a formal model that allows for a simple and direct account of facts within a specific language. . ."
>
> **Back-off probability–** Goldsmith[53] also described the situation where "the number of words in a natural language is so large, that in practical terms we will continue to encounter as many as 25% of the time sequences that have never occurred in the data before." and "A probabilistic model must 'reserve' some of its probability for sequences not observed in the past." Thus, any non-infinite corpus must have a total cumulative probability of less than 100%.

---

[48]Li, S. Wang, K. & Wang, D. et al. (2008) Simultaneous production of low- and high-frequency sounds by neonatal finless porsoises (L) J Acoust Soc Am vol 124(2), pp 716-718

[49]Yip, M. (2006a) The search for phonology in other species *Trends Cogn Sci* vol 10, pp 442-446

[50]Yip, M. (2006b) Is there such a thing as animal phonology? Linguistics Research Center, Univ. California, Santa Cruz pp 311-323

[51]Goldsmith, J. (1996) The Handbook of Phonological Theory. NY: Blackwell

[52]Goldsmith, J. (1992) Local modelling in phonology *In* Davis, S. *ed*. Connectionism: Theory and Practice. NY: Oxford Univ Press

[53]Goldsmith, J. (2000) On information theory, entropy, and phonology in the 20th Century *Folia linguistica*

Since a putative dolphin language is not written, any investigation will necessarily be focused on the phonemes of their speech.

U.2.3 Brief overview of the language facilities of the dolphin

The major question addressed below is whether the dolphin, and particularly the bottlenose dolphin exhibits the capabilities necessary to support intraspecies communications using a recognizable language protocol. If they do, can humans learn to understand it?

> Let us separate inter-species from intra-species communications for now. While inter-species communications is clearly quite common (based on living with our pets if nothing else), some may not want to consider such communications as a complete language. Interspecies communications will appear occasionally below.

The following collage of materials, when placed in an appropriate framework, will demonstrate that dolphins exhibit all of the traits associated with intra-species communications. From a functional perspective, the basic communications ability relies upon the ability of an animal to recognize a pattern delivered to the cognitive elements of its brain and the ability of that animal to reproduce that same pattern in the medium of interest, whether acoustic, visual or tactile. If the animal can recognize a number of individual patterns and reproduce them at will, it clearly has the ability to communicate ideas to another member of that same species. This group of patterns is known as a language. The question becomes is the animal merely mimicing what it heard or is it adding meaning through inflections or other modifications to the original message.

> The word language is used as both a label and therefore a noun, and also as an adjective as in a language protocol. This work will focus on exploring the language protocol of dolphins while recognizing they may employ more than one language because of the dispersal of, and a degree of isolation over long periods of time of, large cohorts around the world.

If the animal can modify or rearrange these patterns, and generate them in the appropriate medium, we say it has expressed a new idea. The chimpanzees and other large primates are doing this regularly at the primate center in Atlanta. I expect the dolphins are doing this in the amusement parks around the country and world.
If two animals can exchange patterns back and forth that have been modified in the process, they can certainly communicate. The remaining question is whether the complexity of that communications warrant it being described as including a language protocol?

There is every indication that the dolphin can perform all of the necessary steps attributable to humans within their language protocol. Marino et al. have delineated these individual steps in the papers cited earlier. The problem is man has not learned how to achieve inter-species communications except using our language(s). In that context, the dolphin may be more capable than humans, dolphins readily learn to understand human spoken communications.

The second question, whether humans can learn to understand a putative dolphin language without at least an intermediate translation mechanism is more difficult to answer. Clearly humans cannot hear many of the vocalizations of the bottlenose dolphin. Ridgeway et al., Au et al. and Kassewitz et al. among other groups have worked with dolphins to the point they can ask them (by voice or whistle) to go to a certain point and carry out a task or respond to a situation (including return to station and press a button). These sequences constitute inter-species communications using patterns (symbols or signs if you prepare). The responses generally involve pragmatics rather than vocalizations.

Humans have become quite proficient at deciphering ever more complex cryptographic messages given a large enough corpus of coded material. This is particularly true when the messages are enciphered using a common symbol set. It becomes much more difficult if the symbol set is not known. Consider the problem of the Japanese trying during the 2nd World War to decipher the “plain language” text of the Navajo Indians serving in the US Marine Corp in the Pacific Theater. In the case of the dolphins, the first step in understanding their vocalizations and language protocol is to discover their symbol set. If that can be done by examining an appropriate corpus of vocalizations, many humans are prepared to attempt the decipherment of the dolphin language(s). However, many thousands to a few million man-hours have been spent listening to dolphin communications (mostly by investigators with limited linguistic and virtually no cryptanalysis training) with only limited results to report.

The task is no different than that of code breaking as practiced during the second world war. There are several well defined steps:
A. Perform traffic analysis to determine what symbols are issued by a given dolphin under a clearly defined (and delimited) set of conditions.
B. Perform a statistical analysis of the potential symbols used in the above traffic analysis. (E appears more often than m in English).
C. Perform an after action analysis to determine what symbols a second dolphin produced in response to specific symbols.

D. Refine the process.

To date, my traffic analysis shows that Cetacea make four distinctly different classes of acoustic signals,
(1) Signals at frequencies between 80-90 kHz and 150 kHz associated with the terminal predatory action (the capture sequence).
(2) Signals at frequencies between 20-30 kHz and 80-90 kHz used to search for food preliminary to predation. These would typically be frequency swept whistles, based on similar bat activities.
(2A) *The potential use of 20-90 kHz energy to stun fish within a few meters and generally directly in front of the dolphin (Ref. Castaway echo 3-23-2007).*
(3) Signals between 200 Hz and 40 kHz designed for intra-species communications within a pod.
(4) Signals at frequencies below 200 Hz designed to support long range communications among the whales.

It appears only the third category is used for intra-species communications.

A problem for humans is to determine what might be meaningful ideas to be shared between dolphins. They will clearly have a call for help, and more specifically a call for mother (ma ma is pretty universal among humans, although it begins to be tailored to fit a specific language after maybe 18 to 52 weeks). What is the equivalent among dolphins?

Breaking the code of dolphin language will not be an easy task. Very sophisticated analyses by very widely versed specialists will be required. Adding more graduate students (or novice Post-Doc's) to the task will not help.

Figure U.2.3-1 describes a nominal block diagram of any vocalization system[54]. A unique feature of marine mammals is that the air is not necessarily released following the creation of vocalizations. It is frequently recirculated within the vocalization system with the acoustic energy being directed into the fluid medium surround.

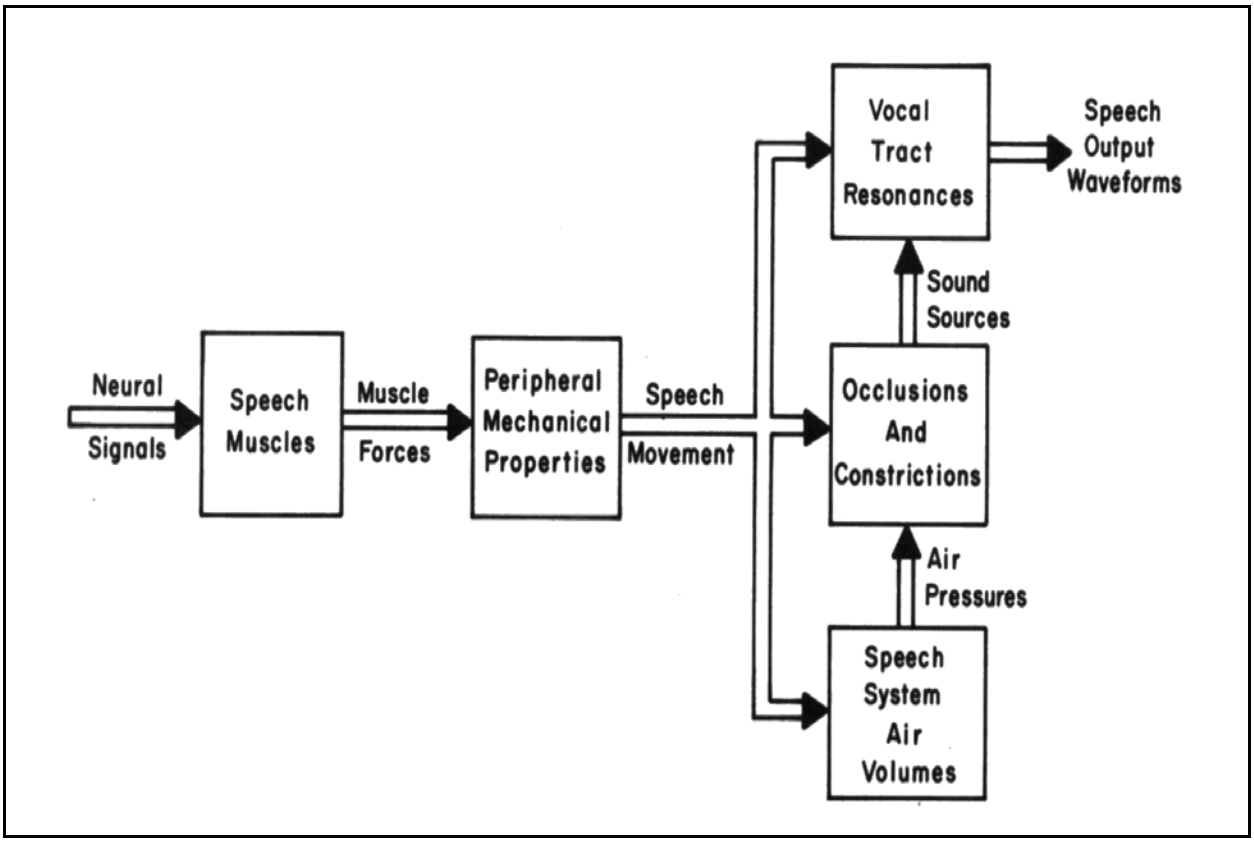

Figure U.2.3-1 Schematic representation of the speech production periphery. The elements on the right all contribute to the articulation of the voice originating in the element labeled peripheral mechanical properties (the larynx). The situation is more complex in dolphins where multiple sound generators are found. See text. From Lass, 1976.

[54]Lass, N. *ed*. (1976) Contemporary Issues in Experimental Phonetics. NY: Academic Press

U.2.3.1 Physiology of the dolphin nasal pathways

The vocalization mechanism of the dolphin is highly integrated with its echolocation apparatus, resulting in an extremely complex system that has not been adequately documented. **Section 1.1.2.3** of Appendix L. describes the overall apparatus using *Tursiops truncatus* as an example. To aid in isolating the active sound generation mechanisms from the hearing mechanisms, a complex combination of air spaces and soft tissue is employed. Simultaneously, the vocal apparatus exists with the purpose of radiating communications signals below 40 kHz nearly omnidirectionally.

- - - -

Figure U.2.3-2 describes the current understanding of the morphology pneumatic system of the dolphin. Contrary to the nasal passages of primates which are largely muscle free, the nares of the bottlenose dolphin are known to be formed of multiple sheets of muscle tissue. These sheets of muscle are a major part of the articulation capability of the dolphin, supporting the change in resonance among the multiple cavities of the nares and also supporting the pneumatic recycling of air between the various cavities when the dolphin is underwater.

It appears the nares can operate individually generating two separate and distinct vocalizations at one time. This is clearly the case for the high frequency signals designed for echolocation, and appears to be true for the lower frequencies used for interspecies communications.

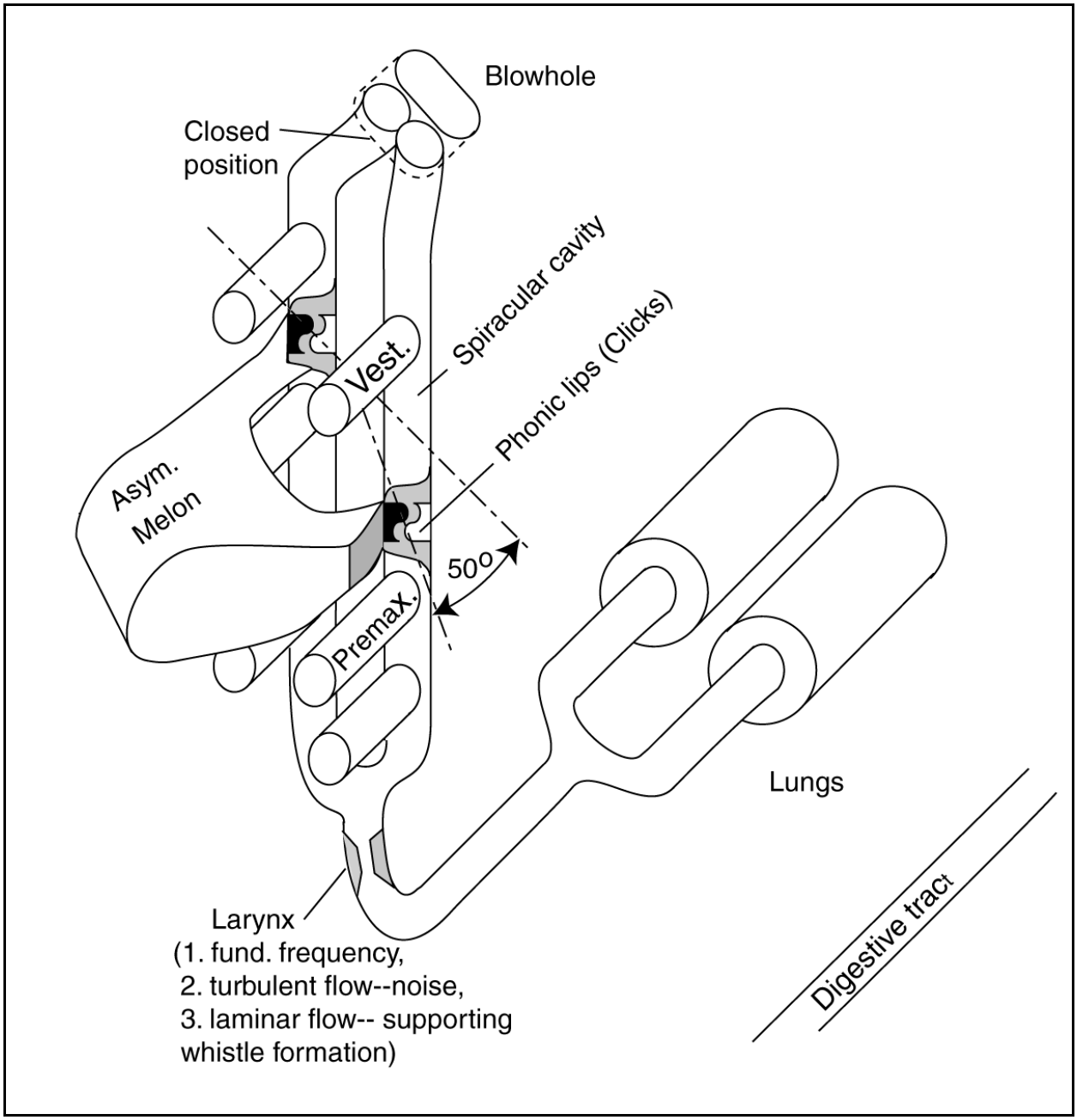

Figure U.2.3-2 The dual channel operation of the sound generation system in the bottlenose dolphin. This simplified schematic provides an explanation of several observed properties of the dolphins sound generation system. First, it isolates the sound generation system completely from the digestive tract. It provides two distinct air columns between the larynx and the blowhole. The individual sets of phonic lips can excite the melon separately and generate far field patterns that are significantly different (either simultaneously or sequentially. Note the angle between the plane containing the pairs of phonic lips and the horizontal plane. See text concerning the potential for the larynx to move up beyond the bifurcation of the pharynx and thereby be duplicated.

The areas labeled spiracular cavities are likely to be muscular walled cavities that are frequency selective and support the generation of a wide variety of higher formants in dolphin vocalization. Some of these formants may not be harmonically related to the other higher formants but be harmonics of the initial formant formed by the larynx.

Not shown are the additional multitude of air sacs surrounding portions of the melon and behind and below the spiracular cavities. These sacs provide effective acoustic mirrors that protect the brain case and the bulla containing the cochlea from the high energy acoustic energy delivered to the melon. They may also act to form the acoustic

output beam in conjunction with the melon. Rodionov & Markov have provided considerable information on the anatomy of the nasal passages but may not have documented all of the air sacs[55].

The dual channel character of the system is quite different from the preliminary description of the dolphin forehead presented in several forums by Aroyan[56]. His draft 2002 paper contains a mixture of morphology and soundings of the bottlenose dolphin, and caricatures of sound generation for a member of the super family *Delphinoidea* that is not homologous with the bottlenose. It also differs from the concept employed here in suggesting the ultra–high frequency (UHF) source employs;
1. A source mechanism of broadband pulses,
2. A damped resonator, and
3. A projector of the signal produced by the first two components.

In this work, it is suggested that the UHF system (up to 150 kHz) in T. truncatus consists of;
1. Two high frequency resonators acting as sources,
2. Two independently controlled envelope modulators of the Gabor type, and
3. A projector (the melon and its supporting structures) of the signals produced by the above first two components.

In the proposed configuration, the spectral width of the pulses would be less than one-third of the center frequency and would not normally be considered broadband by a communications engineer.

If the reports of Markov & Ostrovskaya are correct, the larynx should be moved up above the bifurcation of the pharynx and duplicated (**Section U.4.3.1**). As a result, the low frequency (LF) system (nominally 2 to 40 kHz) consists of;
1. Two low frequency sound generators consisting of a larynx and muscular tissue in the pharynx.
2. Neural circuitry to operate the larynx in three distinct modes, tonal, noise and laminar flow.
3. Neural circuitry to control the tissue tone of the nasal passage musculatura.
4. Morphology to efficiently transfer the sounds generated at the pneumatic/tissue interface into the surrounding environment.

The schematic is somewhat simplified in that Morris and others describe the movement of air from the premaxillary air sac upward past the phonic lips (more properly the valvular flaps of the nasal plugs according to Morris) into an accessory sac and finally into the vestibular sac. Whether the accessory sac is independent of the connecting spiracular cavity is difficult to discern from the published record.

U.2.3.1.1 Fundamental difference between high and low frequency sound generation

In 1986, Ridgeway and Carder reported on the pneumatic pressure in the nares of the white whale, *Delphinaterus leucas*[57]. For purposes of discussion, it will be assumed that similar results apply to the bottlenose dolphin. **Figure U.2.3-3** is selected from their set of spectrograms and shows a distinct difference between the character of the UHF pulses and the lower frequency pulses of *Odontocete*. It is clear that the high frequency pulses generated by the phonic lips do not involve changes in the pneumatic pressure within the nares. They are the result of physical percussion between hard objects. The changes in pressure are directly related to the lower frequency (voiced as shown here) sounds generated by air passing through the larynx. Note the higher amplitude of the percussion sounds irrespective of the pressure. "At times during whistles, intranasal pressure reached 800 mm of Hg or more than an atmosphere of positive pressure." Ridgeway and Carder provide a broad discussion concerning the physiology of *Odontocete* during deep diving. The generation of UHF pulses by percussion irrespective of the pressure environment simplifies this task.

---

[55]Rodionov, V. & Markov, V. (1992) Functional anatomy of the nasal system in the bottlenose dolphin *In* Thomas, J. Kastelein, R. & Supin, A. *eds*. Marine Mammal Sensory Systems. NY: Plenum pp 147-177

[56]Aroyan, J. (2002) Simplified models of the dolphin echolocation emission system http://members.cruzio.com/~jaroyan/publications.html & http://www2.cruzio.com/~jaroyan/Dolphin%20Model%204.htm

[57]Ridgeway, S. & Carder, D. (1986) Nasal pressure and sound production in an echolocating white whale, *Delphinaterus leucas In* Nachtigall, P. & Moore, P. *eds*. Animal Sonar: Processes and Performances. NY: Plenum Press pp 55-60

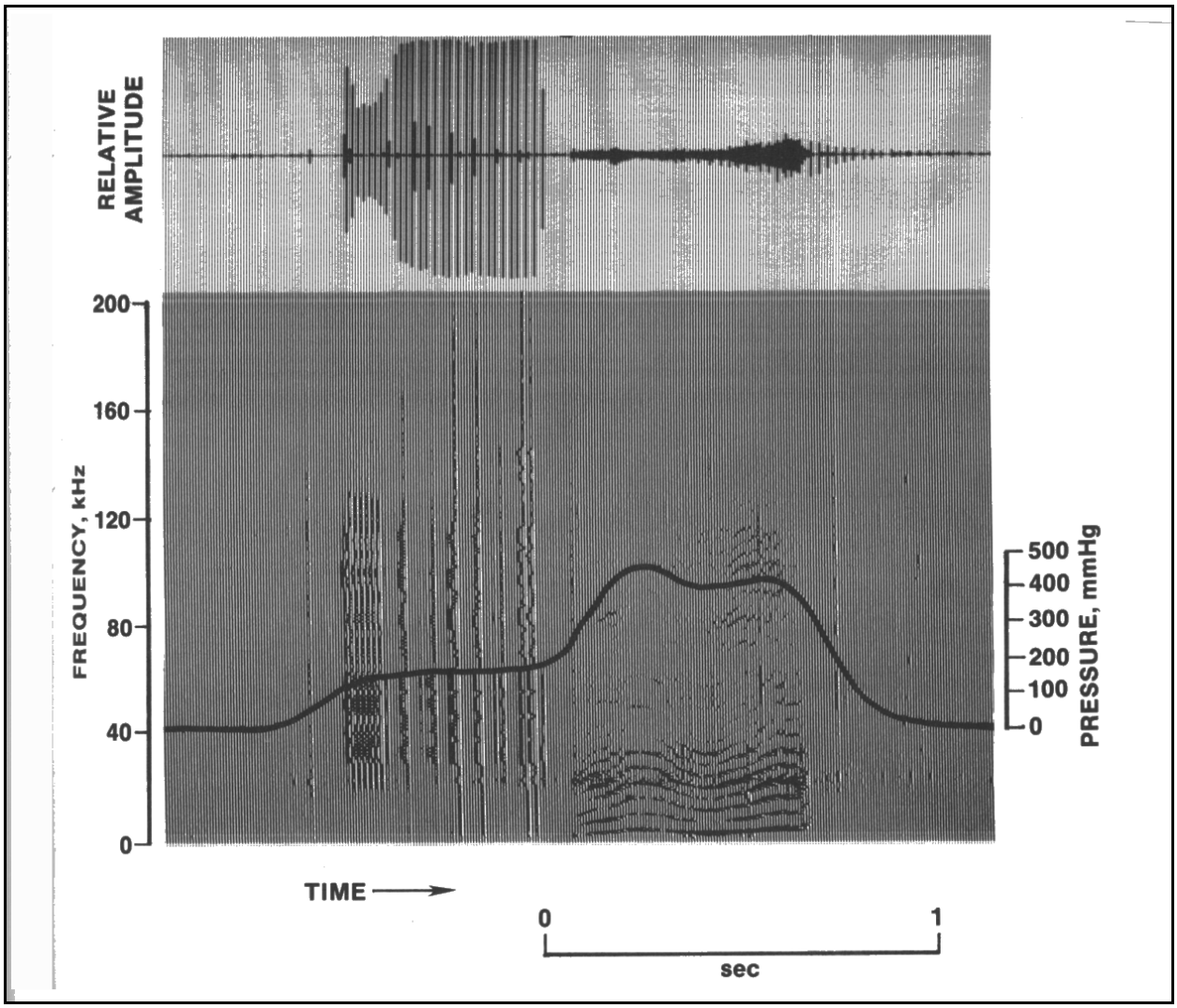

Figure U.2.3-3 Spectrogram and nasal pressure for a typical response by the white whale to a target present trial. Since the maximum recorded frequencies reached 200 kHz, the frequency bandwidth of the analyzer was set at 125 Hz. See text. From Ridgeway & Carder, 1986

U.2.3.1.2 Three distinct modes of sound generation involving the larynx

The fact that the dolphin recirculates the air within its pneumatic system while submerged offers a unique capability. The air may be passed back through the nares without first passing through the larynx. As a result, the musculatura of the nares can cause a resonance that results in a single tone whistle. Alternately, it is possible the larynx operates in three modes instead of the two commonly found in primates:

- first with the air passing from the lungs through the larynx while turbulent thereby generating noise.
- second with the air passing through the taught vocal chords of the larynx thereby generating a fundamental tone.
- third with the air passing from the lungs through the larynx while laminar thereby not generating noise or a fundamental tone but supporting the generation of a single tone whistle within the nares.

Connor & Smolker note (page 651) that what they define as pops are generally generated in the right hand nare of their limited data set (10 observations involving 3 dolphins).

U.2.3.2 Low frequency vocalization parameters

Lilly has provided early data comparing the sounds made by humans and dolphins in the course of vocalizations in

air[58]. The frequencies of the dolphin would not be expected to change when submerged since they still involve the movement of air through orifices, etc. **Figure U.2.3-4** begins with the fundamental (partial #1) and continues with the higher order harmonics versus the pulse rate of the enunciations. No discussions were provided concerning higher frequencies for the dolphin.

Caldwell & Caldwell published a variety of papers during the 1960 and 1970's related to dolphin communications[59]. While using primitive instrumentation including a mirror galvanometer, they provide a wide range of ethological observations of value. "On the basis of these observations we do not intend to state arbitrarily that a single animal has a vocabulary of a single somewhat variable whistle contour."

Winn & Olla have edited a volume on marine mammals including a major article by Caldwell & Caldwell[60]. The Caldwell & Caldwell paper focused on 126, mostly young, bottlenose dolphins. However, the data is seriously limited by the capabilities of the equipment of that era. Bandwidths were generally limited to 12 kHz. Most of the recordings only show a fundamental frequency in the 4 kHz region.

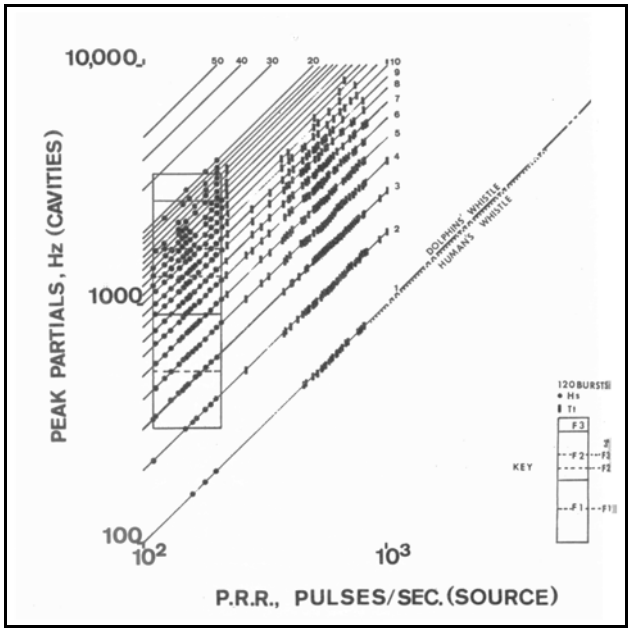

Figure U.2.3-4 Vocal exchanges between dolphin and human in the terrestrial environment. Partials refer to the fundamental plus higher harmonics. Box on left represents the human range as explained by legend on the right. The dolphins vocalizations are systematically higher in frequency and in pulse rate. F1, F2 & F3 are conventional human formants. From Lilly, 1978.

[58]Lilly, J. (1978) Communication Between Man and Dolphin. NY: Crown Publishers

[59]Caldwell, M. & Caldwell, D. (1965) Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*) *Nature*, vol 207(4995), pp 434-435

[60]Winn, H. & Olla, B. (1979) Behaviour of Marine Animals, Volume 3: Cetaceans. NY: Plenum Press

Connor & Smolker have provided more contemporaneous information on low frequency pulse sounds, labeled "pops," with considerable behavioral information about their use[61]. **Figure U.2.3-5** reproduces their small scale figures 2 (on the right) & 3 (on the left). The original caption for the left frame says, "The energy peaks marked by the cursors in the power spectrum are at 1280 Hz and 2080 Hz and the amplitude is ploted on a log (dB) scale." Three features are notable. First, the power spectrum is very broad relative to the noted frequencies, suggesting this is actually an unvoiced symbol likely to represent a consonant. Second, the given frequencies are not harmonically related, suggesting these peaks are likely to be coincidental and probably not at unique frequencies across a wide cohort of dolphins. Their text provides additional data suggesting for 45 sample pops, both the highest peak and the lowest peak have wide standard deviations. The data also suggests some of the spectra may represent voiced symbols based on the lowest prominent frequency. If unvoiced, the recorded sound is probably a form of "bray" as defined elsewhere in this appendix. If voiced, it would probably qualify as a "bark." Connor & Smolker suggest their pops may also be related to the thunks of McCowan & Reiss (1995, specifically page 659). The pulses in the lower frame are not equally spaced with the first and third intervals being significantly different from the others. The authors note the difficulty in characterizing the width of the individual "pops" (symbols) and provide the right frame. The time difference between the two cursors on the right frame suggest the pop durations range between 4 and 11 ms.

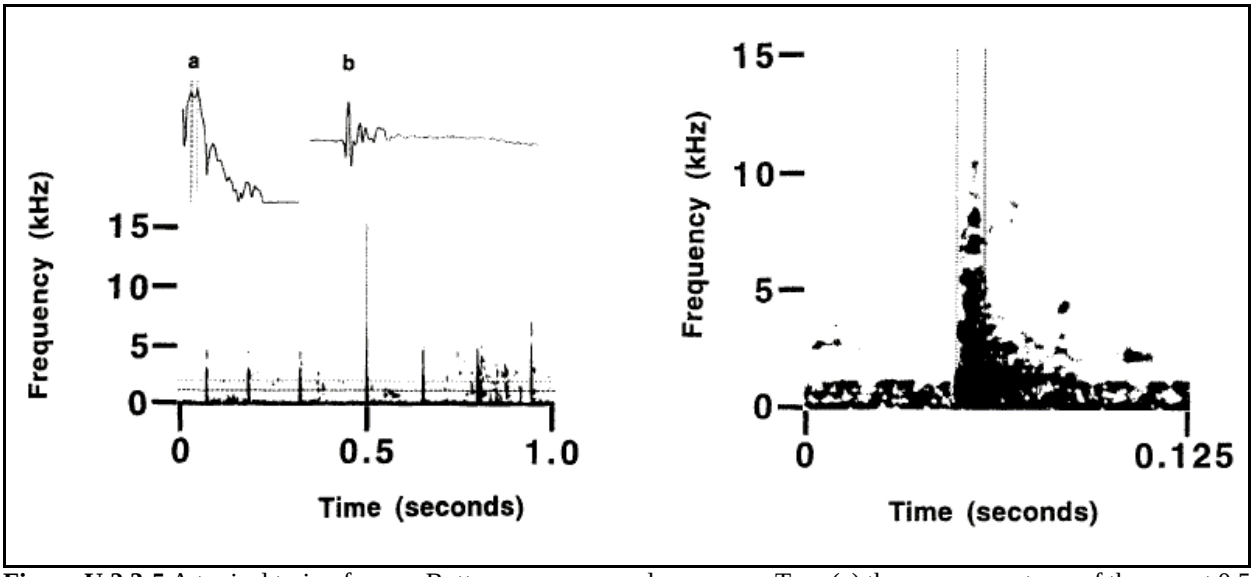

Figure U.2.3-5 A typical train of pops. Bottom; a one second sequence. Top; (a) the power spectrum of the pop at 0.5 seconds, cursor location. (b) an expanded amplitude versus time profile. See text. From Connor & Smolker, 1996.

Connor & Smolker provide additional figures showing other probably unvoiced sequences, including a sequence with two potential echolocation pulses interspersed among the much lower frequency pulses followed by a distinctly different pulse sequence involving much shorter pulse intervals and much higher maximum frequency (possibly a "buzz" as described elsewhere in this appendix). The limited quality of the figures and the failure of Connor & Smolker to indicate the bandwidth of their spectrum analyzer makes further interpretation of these data imprecise.

Connor & Smolker collected this data during their study of the herding practices of dolphins within Shark Bay. The significance of these signals as symbols in a broader symbol set for a dolphin language will be discussed beginning in **Section U.5.6**. They concluded, "Our findings confirm a communicative function for one kind of pulsed train in bottlenose dolphins."

U.2.3.3 Low frequency hearing in dolphins

Johnson reported in 1967 that the hearing of dolphins fell off rapidly at frequencies below 10 kHz. "An audiogram was obtained over the frequency range from 75 hertz to 150 kHz. The lowest thresholds (greatest hearing sensitivity) occurred in the frequencies near 50 kHz at a level of about -55 decibels (re 1 microbar). The effective upper limit of hearing for the experimental animal was determined to be 150 kHz."

---

[61]Connor, R. & Smolker, R. (1996) "Pop" goes the dolphin: A vocalization male bottlenose dolphins produce during consortships *Behavior* vol 133(9-10), pp 643-662

Jacobs & Hall have provided similar data for a fresh water dolphin[62]. "A fresh water Amazon River dolphin, *Inia geoffrensis*, was conditioned to respond to pure tones by pushing a response lever. By using this method an audiogram was obtained for frequencies between 1.0 and 105 kHz. The animal was most sensitive to tones between 75 and 90 kHz. The sensitivity threshold in this area was -50 dB re 1 dyn/cm2. The effective upper limit of hearing was 105 kHz."

Ridgeway & Carder have provided more complete data on several bottlenose dolphins and discovered one that was both deaf and mute[63]. They also cite several additional earlier sources of audiograms of dolphins.

Figure U.2.3-6 describes in conceptual form the hearing capability of the bottlenose dolphin based on the sparse empirical data and the model of hearing from **Sections 4.5** & **9.3.1** of this work. The partial human sensitivity is shown on a relative scale with peak performance normalized to that of the dolphin. The preferred dolphin area of speech is estimated from the similar area documented in human speech. It corresponds to the area of preferred whistles suggested by Richards, Wolz & Herman. It probably ends near 40 kHz based on the predicted highest frequency used for communications predicted here based on spectrographic data. The circle-dot data points of Ridgeway & Carder exhibit deviations from the projected performance based on the Marcatili Effect reminiscent of those of human performance (not shown) due to the parameters of the outer ear. The attenuation at frequencies above 150 kHz exceed 180 dB/decade. Note the negligible overlap in the optimal frequency sensitivity ranges of the human and dolphin. The nominal crossover point is 30 dB below maximum sensitivity in this normalized graph.

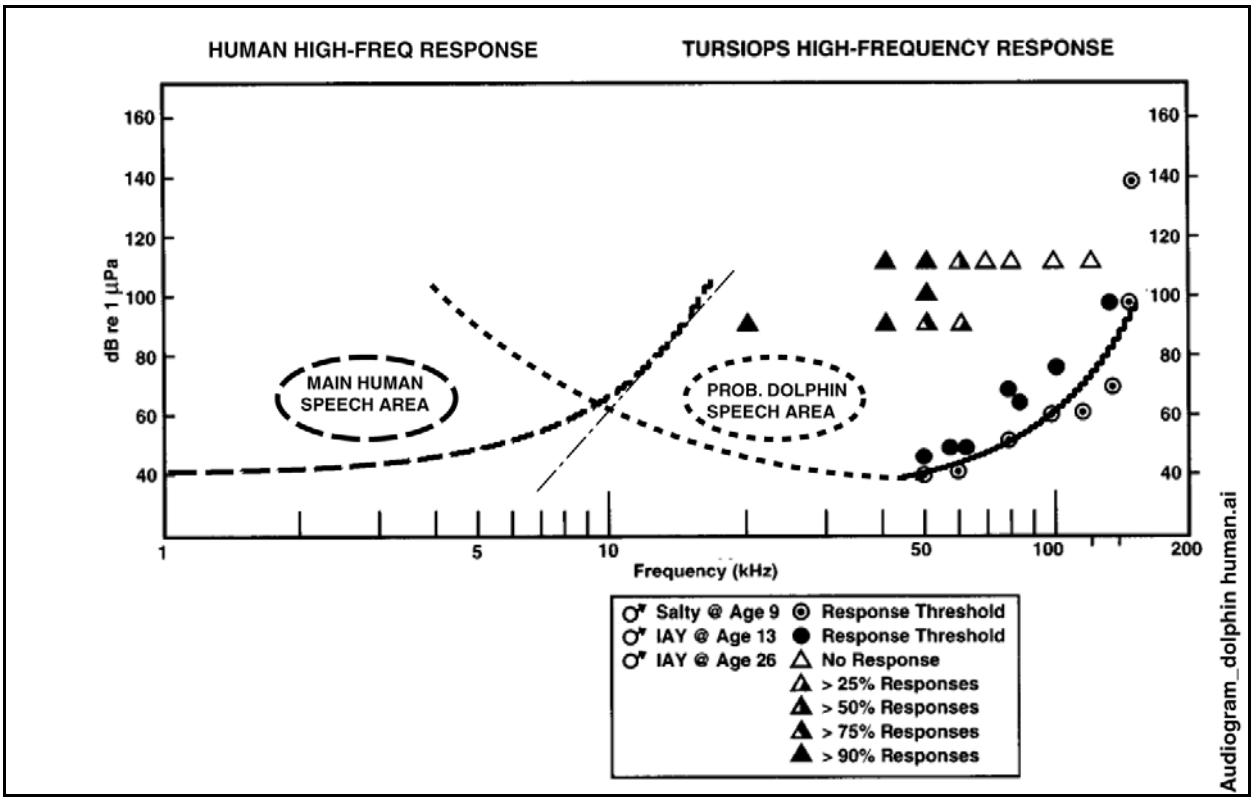

Figure U.2.3-6 Conceptual audiograms of human and dolphin. The dash-dot asymptote is falling at 180 dB/decade. The human hearing is falling off faster than 180 dB/decade at high frequencies due to the Marcatili Effect. It can be assumed the dolphin hearing falls off similarly. The dotted portion of the dolphin audiogram is estimated from the similar human audiogram. Measurements at greater than 100 db down from mid range sensitivity are difficult to make. See text. Data points for 100 ms tones from Ridgeway & Carder, 1997.

Ridgeway & Carder did not provide complete graphs of the dolphin's audio sensitivity although they indicated,

---

[62]Jacobs, D. & Hall, J. (1972) Auditory Thresholds of a Fresh Water Dolphin, *Inia geoffrensis Blainville. J Acoust Soc Am* vol 51, pg 530+

[63]Ridgeway, S. & Carder, D. (1997) Hearing deficits measured in some *Tursiops truncatus*, and discovery of a deaf/mute dolphin *J Acoust Soc Am* vol 101(1), pp 590-594

"Tone stimulus (St) duration was 100 ms with a 2-ms gradual rise in intensity at onset and decline on termination. The findings of Johnson ~1968! suggested to us that this duration was adequate. With three of the older males, some tests were done with 300- and 450-ms tones. Frequencies were 5, 10, 20, 40, 50, 60, 70, 80, 100, and 120 kHz." It is not clear how they obtained the data points above 150 kHz as plotted based on their indicated test frequencies.

The figure illustrates why attempting to communicate with dolphins using human vocalizations may not be effective, particularly when their hearing apparatus adjacent to their lower jaw is out of the water.

U.2.3.4 The division of the dolphin phonetic repertoire into three symbol sets

In the case of the dolphin, a complete symbol set, S(R) may consist of a whistle set W(O) plus a voices symbol set, V(P), and an unvoiced symbol set U(Q). If V(P) or U(Q) are missing from a particular investigators symbol set, or if any of the threes subordinate sets are open, the Symbol set, S(R) is necessarily an open symbol set.

These three sets can be described as; [xxx review for specificity and tie to above section]

• a voiced symbol set, V(P) – symbols of variable duration formed by the low frequency vibration of the vocal chords passed through the selective resonances formed by the muscular shaping of the aural cavity,

• an unvoiced symbol set, U(Q) – symbols of variable duration formed by air from the lungs converted into a turbulent stream as they pass through the larynx but not involving the vibration of the vocal chords resulting in a pseudo-noise sound that is then passed through the selective resonances formed by the muscular shaping of the aural cavity, and
• a whistle set, W(O) – symbols of variable duration formed by air from the lungs converted into a laminar stream as they pass through the larynx but not involving the vibration of the vocal chords resulting in little or no modulation (carrier frequency) until passed through a selective resonance formed by the muscular tissue of the nares. The carrier frequency created within the nares may be further modulated frequently at very low frequency (below 100 Hertz) by the musculatura

The distinctive members of these symbol sets can be illustrated easily.

U.2.3.4.1 Three distinct modes of sound generation in neonate dolphins

Figure U.2.3-7 reproduces a figure from Kassewitz showing several structures in the morpheme of a bottlenose dolphin shortly after its birth. A narrowband filter was used in the spectrographic analyzer to accentuate the harmonic structure of the formants. The signal consists of an unvoiced phoneme extending from 0.18 to 0.22 seconds and up to at least 18 kHz (with some content extending to 22 kHz). From 0.22 seconds to 1.04 seconds, the spectrogram may contain more than one phoneme. The region from 0.22 to ~0.45 seconds shows two formants generically labeled F2 and F3 because of their major frequency components. They do not contain significant harmonically related components. From ~0.45 to 0.73 seconds, both formants are accompanied by closely related harmonic elements suggestive of a larynx frequency, F1, near 700 Hertz. F1 does not appear prominently in this figure (possibly because of the bandpass character of the observing microphone.

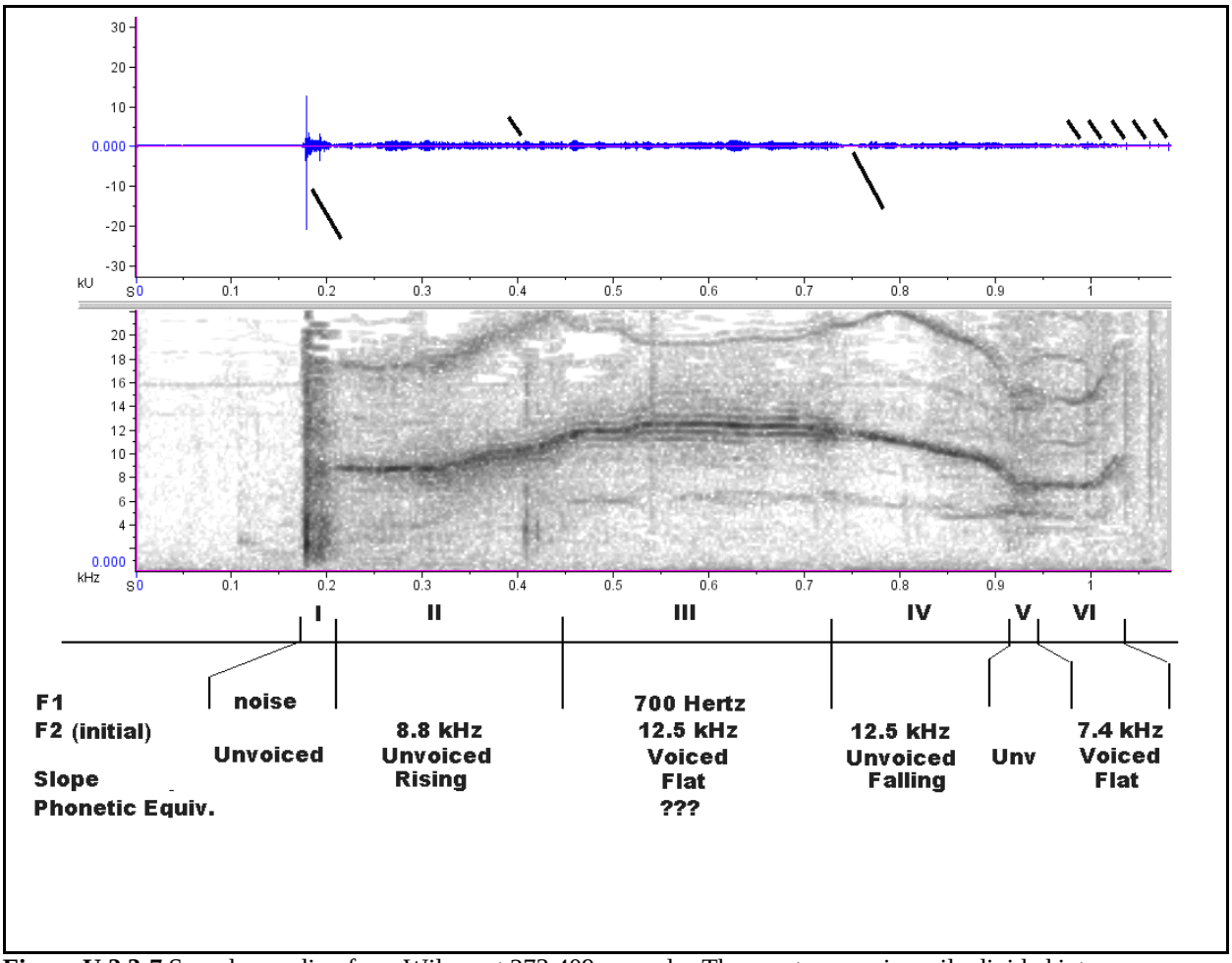

Figure U.2.3-7 Sound recording from Wilson at 273.409 seconds. The spectrogram is easily divided into one or more morphemes, including at least one unvoiced phoneme between 0.18 and 0.22 seconds. The waveforms between 0.22 and 1.04 seconds may consist of multiple phonemes. Note the major tones are not harmonically related, are independently dynamic and appear to originate from different resonators. Playback at 0.1 times the recording rate allows the segments to be identifiable to the human ear. See text. Data portion from Kassewitz, 2007.

The thin (low intensity) tones between F1 and F2 and between F2 and F3 are normally not labeled in human hearing research. Note this group of tones are not harmonically related and appear to vary independently suggesting they originate from different resonators. The fact that they are dynamic suggests the resonators were controlled by distinct articulators (muscular structures). These dynamic variations are normally associated with dialects and personal variations in human speech.

Note the minimum energy region near 0.75 seconds. From 0.73 (or 0.76) seconds to 0.92 seconds, another potential phoneme is seen that consists of two formants that are devoid of harmonic content. From 0.92 to 0.94 seconds, a largely unvoiced area is seen followed by an area structured quite differently and containing at least six identifiable formants. If correct, this interpretation would suggest a series of at least five phonemes during an interval of 0.82 seconds in a neonate bottlenose dolphin. whether the frequency contour near 6 kHz should be considered the actual second formant, F2, and causing the renumbering of the higher numbered formants, is open to discussion. In either case, this signal could easily be interpreted as as a single morpheme consisting of a consonant/vowel/vowel/vowel/consonant/vowel sequence. Alternately, it could be considered two-syllables consisting of a morpheme, consonant/vowel/vowel, a space, and a morpheme, vowel/consonant/vowel.

The tabulation provides a potential (but preliminary) description of the overall signal.

- Interval I is seen frequently in human vocalizations. It is produced by the larynx providing an air burst but not a fundamental tone. F1 is considered a noise source during this interval based on analogy with human vocalization research and is described as unvoiced. The high initial peak intensity (long pointer on the left) is usually associated with plosives.
- Interval II appears to exhibit a simple tone, F2, without sidebands. It may be considered a pure tone produced without the participation of the larynx and would therefore be considered an unvoiced whistle. It is rising.

• Interval III exhibits a tone, still labeled F2, with significant sidebands spaced at the presumed frequency of F1, the larynx frequency. This combination of F1 and F2 would normally be considered a vowel sound in human speech.
• Interval IV exhibits another simple tone, F2, without sidebands and can be considered an unvoiced whistle. It is falling.
• Interval V is complex and may be a transition region between interval IV and VI. Note the breakup of F3 during interval V.
• Interval VI is another unvoiced region of F2 that is flat in its slope. Note the introduction of additional formants above F3 during this interval.

The difference between the signal generated in interval II and III is significant and does not normally occur in most human speech. It suggests that dolphin phonemes may be built from three elements, consonants, two frequency vowels and single frequency whistles.

> Identification of three types of signals in dolphin signaling and the identification of two frequency vowels significantly undermines the CS technique of McCowan that suppresses any sideband structure of the frequency contours.

During intervals II and IV, the flow of air passing through the larynx could be considered laminar in character and not generating noise as in the case of interval I nor a tone as in interval III.

It is noteworthy that the transition segment between II and III and in segment IV contain non-harmonic contours.

Flanagan has addressed the “absolute identification of syllables” by humans in his section 7.32. He notes, “For [his] particular syllables, the one frequency variable (namely frequency of [his band limited] noise burst) appears adequate to distinguish the three consonants. High frequency bursts are heard as ‘t’ for all vowels. For /p/ and /k/ the identification depends not only upn frequency of burst but also on its relation to the vowel. Burst on a level with the second formant, or slightly above are hears as /k/; otherwise they are heard as /p/.”

Note how easily an investigator could delete the consonant from 0.18 to 0.22 seconds from the overall waveform and treat this signal as consisting of only a whistle of significant complexity. Without the amplitude presentation, it is easy to overlook the energy “void” in the spectrogram near 0.75 seconds.

Note also the apparent void near 0.75 seconds in the energy profile, a situation easily overlooked in the spectrogram. It occurs at the end of a voiced segment and at the beginning of a simple whistle.

On playback using Raven Light vers 1.0 (using a rate parameter of 0.1 to slow the playback and thereby lower the pitch of the sounds), these visually distinctive segments of the spectrogram are clearly evident to the human ear and are suggestive of separate phonetic symbols.

Interpretation of the other vertical elements of the spectrogram would require more information concerning the dolphin’s behavior during this period. These elements could be unvoiced phonemes or associated with Wilson’s earliest attempts to generate pulses for echolocation purposes. The energy profile and spacing of the right most five pulses (upper short pointers) would suggest they are echolocation signals.

U.2.3.4.2 Stress and the importance of the amplitude vs time profile

The above combined amplitude profile and spectrogram surface an important fact. Although the whistle is a prominent feature in the spectrogram, the amplitude profile associated with segment I shows that unvoiced signals can be of even greater intensity than the sum of the energy from all whistles contours and their harmonics at a given time. It is likely that these unvoiced signals, frequently described as barks or brays in the literature may be important phonetic elements in dolphin communications. If true, they are likely to be described as consonants in analogy to human communications. They may also provide information concerning the character of any dolphin language; whether it employs stress on various phonemes or syllables to modify the meaning of specific morphemes.

The analysis of the linguistic aspects of dolphin signals are greatly aided by the pairing of the amplitude profile and spectrogram of a given message. In the absence of the amplitude profile, any analysis is necessarily incomplete. The darkness of a whistle contour within a spectrogram is not indicative of its “loudness” in the overall message.

U.2.3.4.3 Example from Killebrew’s excited female dolphin

Killebrew et al. provided more specific data concerning their equipment than most investigators (“using a Hamming window, an analysis filter bandwidth of 175 Hz and an FFT length of 2048"). They do suggest that “burst-pulse sounds can be produced soon after birth and provide new clues about how sound production might develop in dolphins.” Their figures 2, 3 & 5 strongly suggest the bandwidth of their equipment was not adequate to record harmonics of their tonal sequences near 8 kHz. They conclude that more study is required using larger numbers of

subjects. **Figure U.2.3-8** reproduces their figure 5. The spectrogram contains communications elements not considered in the dolphin repertoire developed by McCowan and colleagues. These additional elements are almost certainly speech patterns with some meaning.

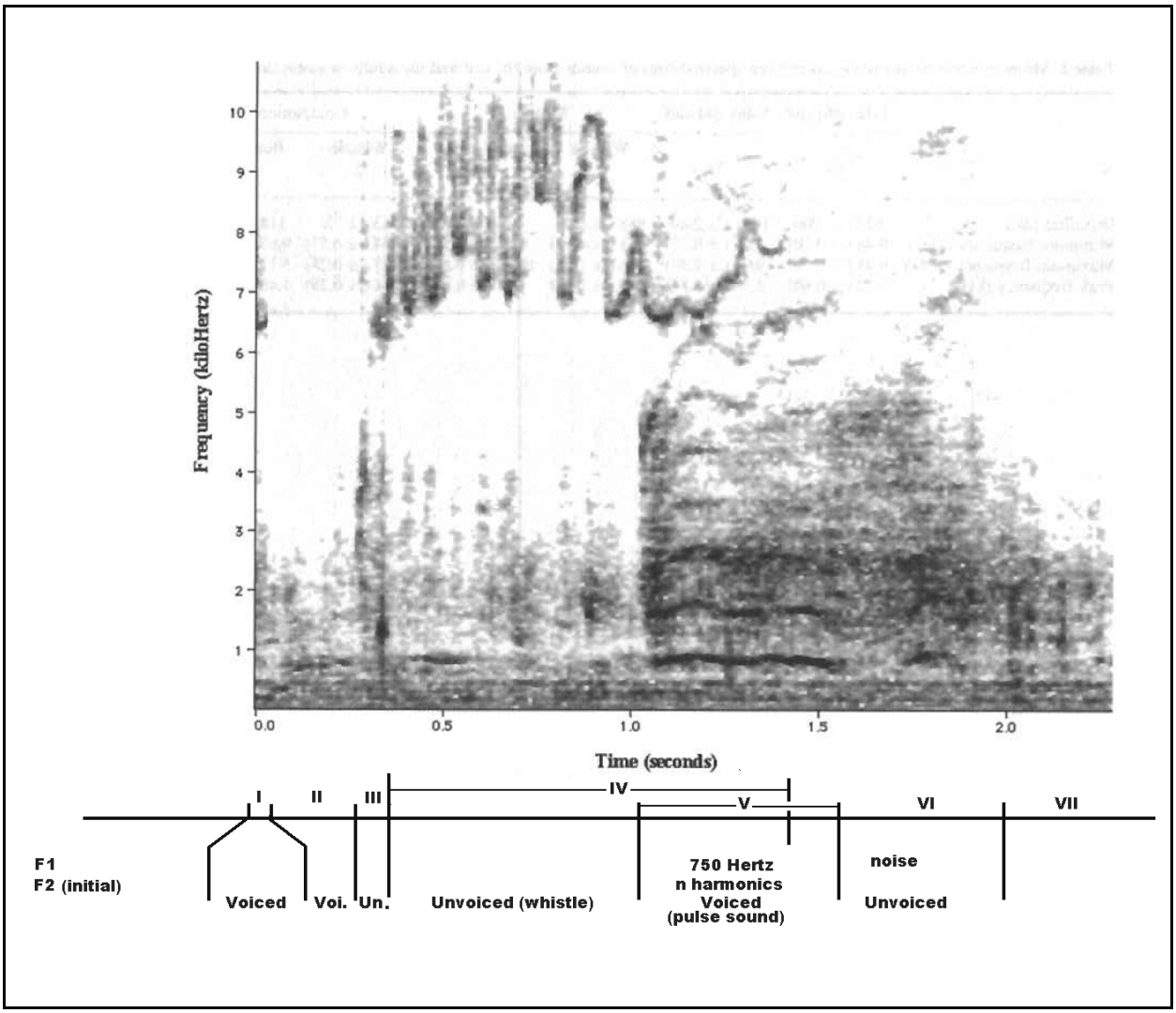

Figure U.2.3-8 Sounds produced by an adult female bottlenose dolphin in a state of excitement. The spectrogram exhibits a brief signal at 0.0-0.05 sec followed by a plosive interval of 0.25 sec, followed by a frequency modulated tone for one sec overlapping as burst-pulse pattern of 1.2 sec. "The frequency modulate element had peak energy at 6.93 kHz, whereas the burst-pulse elements had peak energy at 860 Hz." Frequency elements above 10 kHz, if any, were not recorded. See text. Spectrogram from Killebrew, 2001.

This spectrogram includes features not seen in the previous figure.

- Interval I contains what may be a brief voiced vowel with F1 at 200 Hertz and F2 at 6.4 kHz.
- Interval II is a voiced segment with F1 at 300 Hertz initially but moving to 700 Hertz (could be separate elements).
- Interval III is an unvoiced segment not currently defined in detail.
- Interval IV is a highly frequency modulated tone (trill) with a very low modulation frequency of ~12 cycles/sec.
- Interval V starts before IV ends and consists of a "pulse-sound" (See **Figure 8.1.4-1**) consisting of a strong F1 at 750 Hertz and a series of higher frequency formants that assume harmonic positions after about 0.2 seconds.
- Interval VI is an unvoiced segment with a noise-based F1.
- Interval VII is undefined at this time.

The trill in interval IV is similar to that found among birds of the Warbler family and occasionally heard from human female singers. In the warblers, the modulating frequency is generally between 3 and 4 cycles/sec and the modulation is similar to that observed here, ± 1.5 kHz. In the case illustrated here, the maximum signal intensity

occurs near the low frequency excursion of the waveform. There are some significant phase shifts associated with the ~ 12 Hertz modulating frequency. The pulse-sound of interval V is found quite frequently in human spectrograms. Here the formants above F1 begin by rising rapidly in frequency until they achieve a nominally harmonic relationship with respect to F1 (with some independent frequency variation seen in the higher formants during the middle of the interval. Pulse-sounds are usually interpreted as more complex vowels in human speech. The trill and the character of the higher formants in the pulse-sound are strong evidence of the muscular walls of the nares in dolphins.

> Identification of the pulse-sound type of signals in dolphin signaling along with the identification of two frequency vowels further undermines the CS technique of McCowan that suppresses any sideband structure of the frequency contours (including the harmonic structure of the pulse-sound element).

The overlapping of signaling elements as documented is not difficult to comprehend based on the complexity of the dual channel larynx system of dolphins documented in **Section 1.1.2.3** of Appendix L. Appendix L also contains additional spectrograms of dolphin signals.

U.2.3.5 Potential vocalizations and dialects among other members of *Cetacea*

Several efforts have been made to describe dialect differences among dolphins, particularly during fusion events between cohorts within the limits of Shark Bay and between visiting and resident cohorts of Shark Bay.

U.2.3.5.1 Studies of vocalization among killer whales in British Columbia

Thomsen et al. have attempted to categorize the vocalizations of the *Orcinus orca* of the Vancouver area[64]. They identified whistles and "pulsed calls" with some showing harmonics and some not. they subdivided the pulsed calls into "discrete calls, which are calls that are repetitive, remain stable over years, and are pod specific (Ford and Fisher, 1983; Ford, 1989, 1991); variable calls are calls that are nonrepetitive with a variety of forms such as squeaks, squawks, grunts, and growls. Variable calls were usually rich in sidebands and low in frequency (1–4 kHz)." These vocalizations may or may not exhibit sidebands. "Whistles are sounds based on a tonal format, generally with a continuous waveform which appears in spectrographic analysis as a narrow-band tone with or without harmonics." Their figures 1, 2 & 3 contain a wealth of information concerning the whistles but little concerning the pulsed calls. The spectrograms of the later figures show significant energy in frequencies below 2 kHz generally associated with a formant #1 in human vocalization. The lower frame of Figure 1 shows a spectrogram dominated by a first formant centered just above 7 kHz with significant second and third harmonics of lower intensity. The middle frame shows the second harmonic about 10 dB below the primary signal in intensity. The upper frame appears to show a sign wave at a nominal 8 kHz but it actually shows the summation of the primary signal, second harmonic and third harmonic. The middle frame of figure 2 shows a very clear noise background spectrum and a narrowband and unmodulated whistle centered on 9 kHz. The lower frame of figure 3 shows the appearance of a modulation during an otherwise narrowband signal of multiple harmonics. The modulation itself exhibits modulation as shown in the middle frame. There is every indication that the waveforms shown constitute a sequence of separate whistles and pulse signals. Thonsen et al. conclued, "Therefore most killer whale whistles appear to be structurally similar to those of other delphinids which in most cases show several harmonics" (citing a variety of old spectrograms in: Lilly and Miller, 1961; Busnel and Dziedzic, 1966, 1968; Caldwell and Caldwell, 1971; Caldwell et al., 1973, 1990; Sjare and Smith, 1986; dos Santos, 1990; Schultz and Corkeron, 1994). Their discussion is substantive but remains incomplete. they do draw a largely unsubstantiated conclusion concerning killer whale vocalizations, "Parameter measurements indicate that they are much more complex than whistles described for other delphinids."

U.2.3.5.2 Studies of dialect among killer whales in British Columbia

Deeke, Ford & Spong have reported on their studies of communications patterns among *Orcinus orca* over a 12-13 year period[65]. This period is probably too long to discuss dialect differences between nearby resident pods or cohorts. Their study was primarily statistical and involved characterizing primarily whistles by noting various visually apparent characteristics of the recorded signals from pods consisting of several generations along a matriarchal line (matriline).

---

[64]Thomsen, F. Franck, D. & Ford, J. (2001) Characteristics of whistles from the acoustic repertoire of resident killer whales (Orcinus orca) off Vancouver Island, British Columbia *JASA* vol 109(3) , pp 1240-1246

[65]Deecke, V. Ford, J. & Spong, P. (2000) Dialect change in resident killer whales: implications for vocal learning and cultural transmission *Anim Behav* vol 60, pp 629-638

They described a variety of signals recorded from these groups. "Resident killer whales emit a variety of vocalizations, including echolocation clicks, tonal whistles and pulsed calls. The pulsed calls consist of individual pulses of sound that resemble echolocation clicks in their physical structure. The pulse 'tone' is dictated by the initial rise time of the click, and pulses are repeated at a certain frequency termed the pulse repetition rate. Both pulse tone and pulse repetition rate are modified in killer whale calls (Schevill & Watkins 1966). In addition, some pulsed calls have an independent tonal component termed the high-frequency component with a fundamental frequency of 4–12 kHz (Hoelzel & Osborne 1986)."

"The most common pulsed vocalizations of resident killer whales are 'discrete calls', which are highly stereotyped and can be divided into distinct call types (Ford 1987, 1989)." They assigned an alphanumeric designation to each of their recognized calls.

They focus on multi harmonic whistles in the 1 to 5 kHz range. In their figure 1, they clearly include an interval of time (~ 50 ms) before the commencement of the whistle type call in their determination of call length. They do not account for any simultaneous voiced or unvoiced signaling during their "whistle" call length. They also ignore the presence of a second harmonically related signal beginning at about 350 ms and extending to at least 700 ms.

> The ability of the dolphin to create two simultaneous asynchronous tonal signals is not found in most human speech with the possible exception of the "Throat Singers of Tuva, Mongolia." The hyperlink provides examples of this unusual capability among humans. Recordings of these singers are available for more detailed analysis via the book and CD from Theodore Levin on the hyperlinked page.
>
> The throat singers of Tuva essentially use the musculatura of their throats in the same way dolphins use the musculatura of their nares to resonate at different harmonics of the signals produced in their larynx from those harmonics favored by the aural cavity in the case of the human.

They describe the statistical methods they used to determine divergence and conclude, "The patterns of vocal change in this study, as well as the vocal variation between different matrilines of resident killer whales described by Ford (1991) and Miller & Bain (2000) are therefore most parsimoniously explained by cultural drift." Their study of dialects among local *orca* families appears to be very preliminary in character.

U.2.4 Cryptanalysis tools available for deciphering dolphin communications EDIT

Discovering decisively whether dolphins employ a language protocol requires the use of sophisticated tools beyond the mere recording of their sounds.

There are a vast array of tools available to decipher any messages of unknown content. To date, the surface has only been scratched on bringing these tools to bear on dolphin communications. Decipherment begins with an exhaustive effort to determine the symbol set potentially used in such a language protocol. The projected symbol set can be attacked using a variety of cryptanalysis tools as described in Kahn's comprehensive book, "The Codebreakers."

The key task is to recognize that any dolphin language involves vocalization as opposed to a written record. Hence, the symbol set must be phonetically-based and not alphanumeric character-based.

The second key task is to recognize any dolphin language will likely employ its complete repertoire of sounds and not just whistles. Many of these sounds are sustained for periods of only a dozen milliseconds. Once these steps are taken, the task becomes virtually identical to any other cryptanalysis investigation. It relies upon Information Theory, based on Claude Shannon's primal work in the 1940's, augmented by a raft of tools. One of those tools is the n-gram discussed below.

U.2.4.1 The n-gram of cryptanalysis

A key step in deciphering an unknown phonetic language is to discover how the phonemes (the basic elements of the symbol set) are used in combination to form morphemes (words). **Figure U.2.4-1** from Wikipedia best illustrates how an n-gram is created, with examples from several disciplines as examples.
.

Field	Unit	Sample sequence	1-gram sequence	2-gram sequence	3-gram sequence
Vernacular name			unigram	bigram	trigram
Order of resulting Markov model			0	1	2
Protein sequencing	amino acid	… Cys-Gly-Leu-Ser-Trp …	…, Cys, Gly, Leu, Ser, Trp, …	…, Cys-Gly, Gly-Leu, Leu-Ser, Ser-Trp, …	…, Cys-Gly-Leu, Gly-Leu-Ser, Leu-Ser-Trp, …
DNA sequencing	base pair	…AGCTTCGA…	…, A, G, C, T, T, C, G, A, …	…, AG, GC, CT, TT, TC, CG, GA, …	…, AGC, GCT, CTT, TTC, TCG, CGA, …
Computational linguistics	character	… to_be_or_not_to_be…	…, t, o, _, b, e, _, o, r, _, n, o, t, _, t, o, _, b, e, …	…, to, o_, _b, be, e_, _o, or, r_, _n, no, ot, t_, _t, to, o_, _b, be, …	…, to_, o_b, _be, be_, e_o, _or, or_, r_n, _no, not, ot_, t_t, _to, to_, o_b, _be, …
Computational linguistics	word	… to be or not to be …	…, to, be, or, not, to, be, …	…, to be, be or, or not, not to, to be, …	…, to be or, be or not, or not to, not to be, …

Figure U.2.4-1 The creation of n-grams from a given sample sequence. Here, the sample sequence at the character and word level of computational linguistics are of primary interest. However, the sample sequencing for use in the amino acid and DNA fields may be more familiar. See text.

The DNA sequencing example provides the clearest example for the un-initiated investigator. The goal is to establish the frequency of occurrence (FOO) of a given combination of symbols in a long sample sequence (ultimately known as a corpus for the coded message, or language, under attack. The frequency of occurrence must be obtained for every significant symbol in the complete symbol set. The complete symbol set is known as a closed symbol set. An incomplete symbol set (such as just the whistles of a dolphin) are considered an open symbol set.

The 1-gram, or unigram, is can be used to create a frequency of occurrence versus the symbols in rank order to describe how the symbol set is employed in the target. The 1-gram has two major uses in cryptanalysis, to define the first-order entropy of Shannon, now described as H_1, where the FOO is critically important, as well as a very important parameter, the zero-order entropy or H_0. To compute H_0,the entire symbol set is used under the assumption that the FOO of all of the symbols are equal to the reciprocal of the number of symbols in the closed symbol set. All of these parameters will be discussed more fully in **Section U.4.7**.

As illustrated, the size of the various n-grams grow rapidly with n. This is the reason to have a large corpus available to analyze. Once H_0 and H_1 are established for the target, it becomes straight forward to calculate the FOO for higher argument n-grams. In abbreviated form, a 2-gram presents the FOO for all 2-symbol combinations of the symbol set, S(x,y), where x and y are allowed to be any symbol in the closed symbol set. The 3-gram presents the FOO for all 3-symbol combinations of the symbol set, S(x,y,z), where x, y and z are each allowed to be any symbol in the closed symbol set. These permutations of the variables obviously lead to very large frequency of occurrence tables and/or graphs. Wikipedia notes the 3-gram (without provision for a space) requires a 26^3 dimensional space. N-grams beyond n =3 generally require computer aided analysis. In earliere times, the word computer referred to mostly women employed to create n-grams. Such analyses are greatly aided by using a variant of the machine-readable phonetic alphabet, known as SAMPA, reoriented to fit the symbol set of the dolphin. SAMPA is an outgrowth of the more familiar International Phonetic Alphabet (IPA). The IPA code is unwieldy when used in computer-aided analysis.

Once frequently occurring groups of symbols are highlighted, it is possible to begin seeking the meaning of those combinations, frequently through behavioral analysis in the case of dolphins. This behavioral analysis is known as

pragmatics.

The field of computational linguistic research is a very active one at this time[66]. Google has made great use of word based n-grams to create their search aids.

U.3 (Reserved)

U.4 Research relative to defining and documenting the putative dolphin language(s)

Since the time of Lilly, and before, humans have attempted to communicate with dolphins. However, the methods have always been homocentric, teaching the dolphin a "foreign language" by training it to respond to specific sounds in a prescribed way. Even when the sounds chosen are recorded from the same or other dolphins, the result is the same. We attempt to teach the dolphin to respond to its own name (signature whistle) or some sound with an entirely different meaning in its native language.

> An easy example of this is as follows. An English speaker encounters several Russians in conversation. He knows no Russian. He hears a phonetic sequence that corresponds to his understanding of the English word "ocean." He then attempts to communicate with one of the Russians with the word ocean as he understands it, implying a large body of water. By sufficient arm waving and gesturing, he may be able to get the Russian to associate the phonetic sound for the English word for ocean with a large sea, even though in the Russian's language the word ocean is an adverb meaning "very" or "very much."

> Trying to teach any subject a foreign language, instead of learning his language, in order to communicate takes a lot of effort; and in an analogy to a cat, "bores the hell out of the cat."

It appears a more productive approach is to try and understand the language of the other species based on a sophisticated linguistic attack.

Before performing such a linquistic attack, it is important to know something about the social networking of the dolphins. Lusseau et al. have performed networking studies on at least one family of relatively isolated dolphins[67]. A literature search uncovers many more.

> The art of encryption/decryption of human written material is a very old one and in recent time has become a very sophisticated one. Encryption of written material has been undertaken at the idea level; and the sentence, word and letter levels. In spoken form, encryption has been used at the frequency band level (in elementary form during World War II), at the phoneme level and more commonly in recent times by digitizing the speech and then encrypting the digital values (generally on the individual bit stream level).

> In the case of putative dolphin speech, learning the language/protocol supporting the speech will probably require an attack starting at the phoneme level.

If a sophisticated linguistic attack is to be performed, this attack should be designed to discover the phonemes and other elements of their speech and how these are used in a syntax that the other species understands. This can lead to uncovering the lexicon of the speech and other features that together represent the language or protocol used by the target species. The basic premise of this author remains that no investigator with adequate multi-disciplinary background (such as Alan Turing, John von Neumann and William Friedman during World War II) has attempted to decode and define the dolphin morphemes, constructs (grammar?) and usage patterns as we currently know them.

Section S.4 of **Appendix S** has established a framework for a "language" that is hopefully acceptable to a linguist. It incorporates five corollaries to the null hypothesis that "Language is a protocol used in conjunction with a communications channel primarily among members of the same species (conspecifics) for the exchange of ideas."

Since the dolphins language is not a "written language," the procedures involved must be highly statistical. It must probably begin at the phoneme level. Since the language is not written, the statistical analyses are much more complicated than typical code breaking where a limited symbol set is used. **Appendix K** includes several attempts to document the phonemes/morphemes of dolphins (using a variety of recording equipment of generally inadequate bandwidth). Even these recording suggest that to learn Dolphinia or Truncatees, a very large symbol set must be

---

[66] Dunning, T. (1994) Statistical Identification of Language http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.1958

[67] Lusseau, D. Schneider, K. et al. (2003) The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations Behav Ecol Sociol vol 54(4), pp 396-405

attacked until some associations with physical behavior or intra-species communications can be ascertained.

Any successful attack at cryptanalysis of Dolphinia must recognize the potential for dolphins to employ multiple distinct languages based on the isolation of various cohort, or as a minimum dialects associated with these cohorts. May-Collado & Wartzok hint at the probability of dialects among the cohorts located along the east coast of South America.

An assumption can be made that the CNS of the dolphin employs a frame time similar to humans, nominally 30 milliseconds. As a starting point, it is reasonable to assume that a typical phoneme is three to five times longer than 30 ms based on experience with human languages. A complication arises because of the multiple sources of sound within the dolphin's nasal passages, recalling that the larynx of the dolphin is not in its throat (the passage along which food is ingested). These multiple acoustic sources suggest the dolphin can achieve an equivalent of the unique "throat singing" of the Inuit and Tuva people of Asia. Throat singing is also known as overtone singing. Wikipedia provides a good description of throat singing. In Dolphin, the equivalent is clearly nasal singing. Nasal singing in dolphin may be even more extensive.

As a starting point, the acoustic range of dolphin sounds (other than its forward focused echolocation sounds) potentially useful in communications (language) extends from about 2000 Hz to at least 80,000 Hz (with whistles extending up to about 40,000 Hz). In seeking to understand Dolphinia or Truncates, it makes no sense to restrict our analyses to frequencies that the human can hear. Nor is it rational to limit the dolphin's language to either a tonal or a stressed structure.

U.4.1 A conceptual sequence of dolphin acoustic activity

The following material will be restricted to, and focused on, a framework applicable to the study of Truncatese, the defined language of the bottlenose dolphin. This language probably involves significant dialects based on the isolation of pods over time. Whether the dialects are mutually understandable within the species is not clear. Whether it is understandable by other members of the genus and/or family is even less clear.

The sequence of sounds produced by bats during the target acquisition, tracking and terminal pursuit suggests that dolphins might also employ such a sequence of acoustic mechanisms. In the case of the dolphin, the sequence might be extended to include:

1. Intra-species communications to organize a search using low-band frequencies.
2. Passive location of potential schools of prey using low-band frequencies.
2. Location of potential prey using mid-band tonal echolocation to search.
3. Identification of potential prey using mid-band frequency sweeping in echolocation imaging (3-D).
4. Closing on and herding of prey using mid-band echolocation.
5. Potential stunning of prey using mid-band frequencies at maximum power level (0.12 ms pulses)
6. Closing on individual prey using high-band frequencies in echolocation and imaging mode (3-D).

This is only a skeleton of a strategy. Schusterman et al. provides a much broader discussion of dolphin foraging.

Kaiser has provided a broad discussion of intraspecies communications among the bottlenose dolphins[68].

U.4.2 A scenario for learning about dolphin intra-species communications

Researchers have been attempting to learn the mechanisms and content of dolphin speech for a long time. The problem is more complex than mere code breaking. Particularly in telegraphic code breaking, one knows the complete extent of the symbol set, has good traffic analysis data and has a reasonable idea of what might be in the coded messages. The problem is more complex in verbal code breaking, especially inter-species code breaking, where the symbol set is not known. In the case of dolphins, the symbol set may not be the same between unrelated and distant pods, just as human populations have developed widely varying languages, and dialects within these languages, based on their isolation from each other.

Because of the range limitations on underwater sound and the known background noise level in literal waters, it is likely the intra-species communications within a large bay would use low-band frequencies (maybe 200-5,000 Hz). This needs to be confirmed. At close range frequencies up to 30 kHz or higher could easily be used.

It is not clear from the literature whether any investigators have explored the potential symbol-set among a given pod of dolphins. Richards has provided a bibliography up to 1986. Analysis of recordings of mimicry among a pod of

[68]Kaiser, D. (1990) Linguistic Behavior in Nonhuman Species: A Paradigm for Testing Mental Continuity http://home.onemain.com/~dk1008206/html/cexam.htm

dolphins might be a productive approach to determining their symbol-set, particularly if sufficient raw recordings are available. If examples of mimicry between different pairs of individuals were analyzed, it may be possible to order the two symbol-sets. See **Section U.4.7.10** for recent literature.

U.4.3 Elements of intra-species communications

Humans communicate using a wide variety of techniques:
1. Facial expressions
2. Body language (including posturing)
3. Touching
4. Speech
5. Non-speech vocalizations

It must be assumed that dolphins use the same techniques, plus at least one other (blowing bubbles at appropriate times). Cahill has formalized some of these techniques in brief text[69].

Signal/Behavior	Potential Meaning
1) Contacts between two or more dolphins	
Rubbing bodies	Affection or affiliation, strengthening social bonds, reaffirming relationships, quieting an excited peer (usually a youngster)
Pectoral fin to pectoral fin rub	A greeting between two dolphins
Pectoral fin to side of peduncle	A solicitation for a favor or help sometime in the near future (He called this “contact position” in his research.)
Hits, rams, slams, bites	Usually irritation or aggression from older dolphins but when accompanied by soft angles of approach these are playful
Melon to genital contact	When a mom and her calf are swimming in echelon, the calf will often touch her genital area with its melon maybe indicating it wants to nurse. If the mom initiates contact then maybe it is telling the calf to nurse.
2) Dolphin vocalizations	

Dolphins are relatively homogeneous in their sound generating and sensing capabilities. However, Ketten and Wartzok have noted differences in the curvature of the cochlear partition in some river species[70]. The river species (which may qualify as a separate genus) use echolocation frequencies up to 200 kHz. Their type I cochlear partitions are nearly planar and deviate significantly from the Hankel function of the type II partitions of othe dolphins.

It is clear from the spectrograms of **Section U.1.2** that the dolphins exhibit essentially all of the sound sequences found in human language. The relevant sound shown in these spectrograms are described below. The ocean species exhibit the following sounds according to Cahill.

Whistles	Often dolphins produce a stereotypic whistle that is usually called a “signature” whistle. We now know this is used for maintaining a contact among individuals.
Chirps	Short in length sounds that resemble the sound of bird chirps, these sounds may signal a dolphin’s emotions. . . Sort of an “okay” message.
Click trains	Short pulsed vocalizations of high and wide-band frequency that are used to investigate objects or search for fish. Often these sound like creaky old doors opened slowly and have been called echolocation.

[69]Cahill, T. (2000) Dolphins. Washington, DC: National Geographic

[70]Ketten, D. & Wartzok, D. (1989) Three-dimensional reconstruction of the dolphin ear *In* Thomas, J. & Kastelein, R. *eds*. Op. Cit.

Squawks	Pulsed vocalizations that sound like "squawks" and are very high in repetition rate. They are used mostly in fights or during play by younger dolphins and may signal irritation or anger.
3) Bubbles	
Bubble stream or trail	A stream or trail of little bubbles that escape from a dolphin's blowhole. . . often seen from young dolphins during excited and playful swimming.
Bubble clouds	Large pockets of air from the dolphin's blowhole ("clouds") that are used to express anger or warning. They may also be used as a "shield" from another dolphin's harsh vocalizations (*or attempts to injure acoustically*).

The italics were added by this author. Reynolds et al. gave a slightly different list including frequency ranges[71]:

Sound Type	Frequency Range	Function
Clicks	0.2-150 kHz	Echolocation
Whistles (squeals)	0.2-24 kHz	Individual recognition–Group cohesion
Low freq., narrowband	0.3-0.9 kHz	Unknown
Rasps, grates, mews, barks,yelps	0.2-16 kHz	Communications?

The above descriptions are too brief for purposes of understanding the symbol set and potential messages used in acoustic communications. The rasps, grates, mews, etc. in particular need to be associated with individual spectrograms. The comparison of dolphin chirps to bird chirps might more closely relate to bird trills (See Nowicki & Podos in **Section U.4.73.4**)

Santos et al. have provided a somewhat more detailed breakdown of dolphin sounds[72]. They describe the "Types of calls" along with sample spectrograms of seventeen whistles (limited generally to 24 kHz) and their occurrence during a collection field trip. They also provide spectrograms of bangs, brays, etc.

Whistles– continuous, tonal sounds, occurring in a variety of frequency modulated contours, sometimes repeated in series. Very complex contours lasting up to 2.4 seconds are shown in the sketches in their Table 1.

Pulses sounds– the broadband clicks generally assumed to be associated with echolocation, and which come in trains of variable length and repetition rate. As the pulse rate increases, these trains sound like creaks (such as those of a rusty hinge), low creaks and moans.

Creaks– which usually are broadband, have a pulse rate extending over 40 pps.

Moans– similar to creaks but occurring at higher pulse rates (>40 pps) and frequently exhibiting higher harmonics simultaneously.

There appears to be a problem leaving the pulse frequency of moans unlimited. It is suggested that pulse rates above 100 pps be describes as a **Buzz** as perceived by human attendants. The moans would then cover 40 to 100 pps.

Bangs– a relatively loud broadband pulse (duration about 20 ms). These resemble a typical click waveform but on a longer time scale.

Brays– (resembling a donkey bray) a series of squeak-like sound followed by grunts. Typically 390 ms between the squeak sequence and the grunt sequence.

"Buzz Effect"– bursts of variable duration that seem to be modifications of other sounds, making them noisy, with the appearance of a buzzing wasp.

Blasts– noisy sounds appearing in sequences of 3 bursts each lasting about 150 ms. The whole set lasts

[71]Reynolds, J. Wells, R. & Eide, S. (2000) The Bottlenose Dolphin. Gainsville, FL: Univ. Press of Florida pg 76

[72]Santos, M. Caporin, G. et. al. (1989) Op. Cit.

about 700 ms.

They also noted long periods of silence, particularly when the group was traveling (Fast Directional Moving), when passing through sensitive areas like the mouth of the river or when they had to pass close to a stationary boat. They also noticed an absence of calls during presumed prey-searching, Erratic Group Movements or Slow Directional Moving.

Fast Directional Movement– During this activity, the surface of the external ears are faired into the body form to maximize swimming efficiency. As a result, the ears exhibit reduced forward gain and the dolphin loses forward hearing sensitivity at high frequencies.

Slow Directional Movement– The animals surface as a group and cover relatively short distances. There may be a zig-zag pattern or some interruptions in the linear displacement.

Erratic Group Movement– Similar to the slow directional movement, except that either there isn't a total displacement of the group or elsoe it occurs in a variable and unprdeicable fashion. They appear to be hunting for prey.

Spread Erratic Movements– the animals surface in a dispersed fashion, mostly alone but also in dyads or triads, spreading over a wide area. Seems to be a feeding or searching for prey activity.

Localized Surface Feeding– the animals surface close together, but in different directions, showing fast movements. All dives are very short. Many leaps of different types are visible. Appears to be a collective attack on a fish school.

Localized Surface Interaction– unrelated to feeding. Considered episodes of social interactions and play.

Some of these movement categories may be associated with passive listening for prey detection.

Lilly has also commented about the first sounds of baby and juvenile dolphins[73]. "Initial sounds: distress whistle for calling mother, "putts" for localization (release of air from blowhole). At 9 months, postpartum, control of clicking without air loss matures. Whistle control suddenly complex."

Kasselwitz (private communications, 15 June 2007) has recently recorded the first sounds of a newborn dolphin (within two seconds of birth). They included an initially constant then rising whistle with dominant frequency of 9 kHz rising to 18 kHz at 0.14 seconds, one distinct click of less than 0.01 seconds with most of its energy below 20 kHz, and a series of broadband sounds extending up to at least 90 kHz with durations of about 0.014 seconds repeated at a rate of 60 pps for 12 to 22 times The energy density of each pulse was highest at the beginning.

U.4.3.1 Elements of human speech as a potential model

Human speech offers a potential model of dolphin speech. However, it can not be relied upon in this respect. The physiology, and specifically the vocalization mechanisms are entirely different. There are indications the mechanisms of dolphins is considerably more capable, including the capability to employ multiple separate and distinct vocalization mechanisms independently and simultaneously[74]. It appears only the basic framework of the concept of a language protocol can be relied upon.

U.4.3.1.1 Background from Fletcher impacting more recent behavioral work

Much material from the study of English and other natural languages can be applied to the study of dolphin communications (See Appendix S and Chapter 8 of Hearing: A 21st Century Paradigm."). Figure 8.1.4-2 is reproduced here as **Figure U.4.3-1** to show how critical harmonics and time differences between consonants and vowels are in English. Fletcher presented two variants of the figure with the other (**Figure 8.1.4-1**) using a narrowband analyzing filter (~45 Hz) to emphasize the multitude of harmonics actually employed. The fundamental produced by the larynx, formant F1, varies between 100 and 500 Hz in humans.

---

[73]Lilly, J. (1978) Communication between Man and Dolphin. NY: Crown Publishers pp 47

[74]Markov, V. & Ostrovskaya, V. (1989) Organization of communication system in *Tursiops truncatus montagu. In* Thomas, J. & Kastelein, R. *eds*. Sensory Abilities of Cetaceans. Op. Cit. Pp 599-611

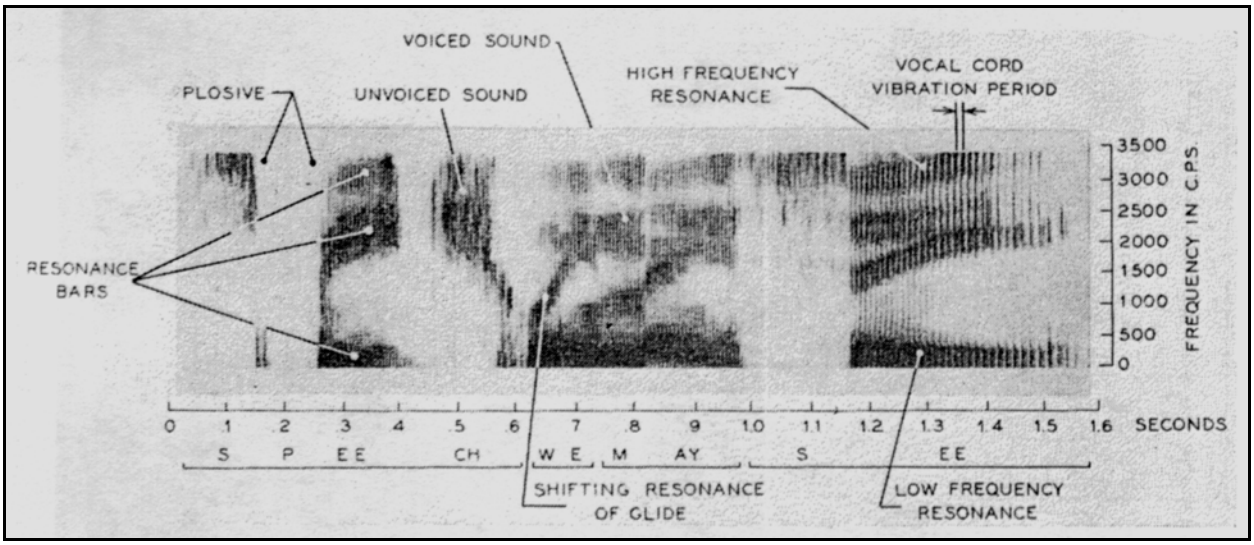

Figure U.4.3-1 Annotated spectrogram of human speech, "Speech we may see" using a wide band analyzing filter (300 cycles) to emphasize vocal resonances. Figure and captionfrom Fletcher, 1953.

Note the complexity of the sounds related to specific utterances. The term fricative was largely synonymous with plosive in Fletcher's time. The larynx typically vibrates at frequencies up to about 600 Hz. The vocal cavity is highly frequency selective with the resultant speech utilizing up through at least the fifth harmonic of the fundamental. The various phonemes associated with the sound, "e" involve multiple harmonics of the fundamental. This is characteristic of English, the vowels involve multiple harmonics of the underlying fundamental. Note the different patterns in the harmonics associated with "ee" in the expression "see." Note the "s" in "see" exhibits no fundamental content and no harmonic content below the fifth harmonic. The "s" in "speech" exhibits similar content. Note the nearly total lack of energy content in the "p" in "speech." This energy void extends for about 900 ms yet, it is critical to the meaning of the word "speech.". Removing the "p" would lead to an entirely different word, essentially the word "seech" spectrographically similar to the word "see."

Figure 8.1.4-3 is reproduced as **Figure U.4.3-2**, modified from Stevens[75]. It shows a spectrogram at an expanded scale. The spectrogram was sampled at approximately 85 samples per second (a 12 ms window). The content of the spectrogram for the t-sound in the word spectrogram suggests stimuli of this short duration are important in speech perception. Stevens noted the spectrogram showed the basic alternation between vowels and consonants (at least in English). "Vowels tend to be of greater intensity, and are produced when the vocal tract is open, whereas consonants tend to occur when the vocal tract is in a more constricted configuration." If the vowels are considered the nuclei of syllables, syllables occur at a rate of 3-5 per second and have an average duration of 200-350 ms. Within this interval, the vowel tends to occupy 100-150 ms on average. Stevens noted the necessity of interpreting individual sounds with durations of 40 ms or less while simultaneously perceiving patterns extending over 0.5 sec. or more. Even these patterns may be only a subset of a longer duration context. Stevens goes on to discuss the potential information extraction mechanisms. His last paragraph parallels the procedure believed to be at the heart of reading as well. "The process must involve temporary storage, hypothesis generation, internal synthesis, and matching of some representation of the signal."

A scale has been added to the original figure to highlight the harmonic character of the stimuli, based on an initial tone of 600 Hz. While there is a clear harmonic relationship between many of the stimuli segments, there is also variation with time in the harmonic relationships involved. These changes suggest the dynamic role of the aural cavity.

[75]Stevens, K. (1975) Speech perception *In* Towers, D. *ed*. The Nervous System. NY: Raven Press pp 163-171

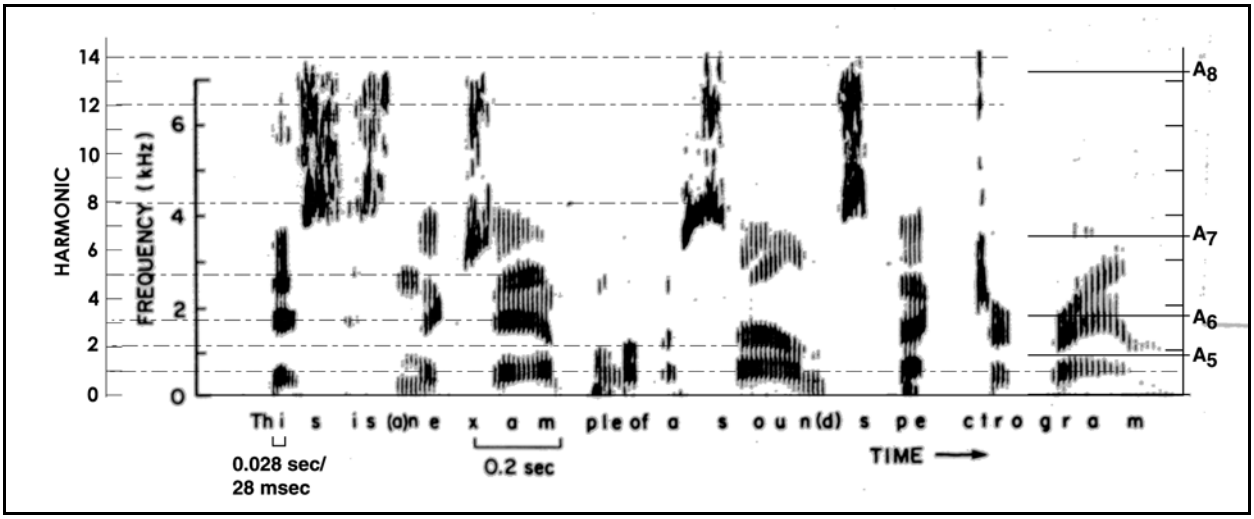

Figure U.4.3-2 Spectrogram of speech *generation* with harmonic overlay using the "summation" tonal channels EDIT re Zwicker. The scale at extreme right shows the harmonic on the assumption that the fundamental tone of the larynx was at 600 Hz. The scale on the right shows the nominal location of the "A' note in the octave defined by the subscript. Data from Stevens in Towers, 1975.

Stevens presented a summary of the features of speech important in its perception[76]. His emphasis on the importance of the context in which the speech was received is significant. He notes, "speech perception is not simply a passive extraction of acoustic properties from the speech wave. . . Rather, the process must involve temporary storage, hypothesis generation, internal synthesis, and matching of some representation of the signal." A very similar set of processes has been developed by this author in relation to reading[77].

It is clear that the English language depends on simultaneous, but minimally related, high order harmonics, of frequently long duration, and changes in tone frequency of individual harmonics during the production of words. It is important to maintain these options when implementing the analysis of any other potential language. Any Bayesian approach to speech analysis that does not maintain these options must be considered inadequate.

U.4.3.1.2 Simple two-section voiced stops of human speech

Liberman et al[78]. provided an important figure developed from spectrograms of human speech that is probably significant in discussing dolphin sounds, specifically whistles. Liberman et al. note, "In listening to speech, one typically reduces the number and variety of the many sounds with which he is bombarded by casting them into one or
another of the phoneme categories that his language allows." In developing his experimental protocol, he noted, "For this purpose a synthesizer was used to generate speech-like sounds and to vary them in small steps along an acoustic continuum known to contain important cues for the perception of the voiced stops, b, d, and g."

Figure U.4.3-3 shows a series of hand traced spectrograms with the first formant shown on the lower line and the second formant on the upper line.

"In the stimulus pattern at the extreme left of the top row of the figure, the second formant rises from a point 840 cps below its steady-state level, and in the pattern at the extreme right of the bottom row it falls from a point 720 cps above the steady state. Between these two extremes, the starting point of the transition varies in steps of 120 cps. For convenience these stimuli will be referred to by number, from 1 through 14, as indicated in the figure.

The rising transition of the first formant had been found previously to be a marker for the class of voiced stops (4) and, as can be seen in Fig. 1, this first-formant transition is constant in all the stimuli. In the steady-state part of the

---

[76]Stevens, K. (1975) Speech perception In Tower, D. *ed*. The Nervous System: Human Communications and its Disorders, Vol 3. NY: Raven Press pp 163-171

[77]Fulton, J. (2005) Processes in Biological Vision www.sightresearch.net/pdf/19Reading.pdf

[78]Liberman, A. Harris, K. Hoffman, H. & Griffith, B. (1957) The discrimination of speech sounds within and across phoneme boundaries *J Exp Psychol* vol 54(5), pp-358-368

pattern the first formant centers at 360 cps and the second at 2160 cps. Formants at these frequencies produce a synthetic approximation to the vowel *e* (as in *gate*)."

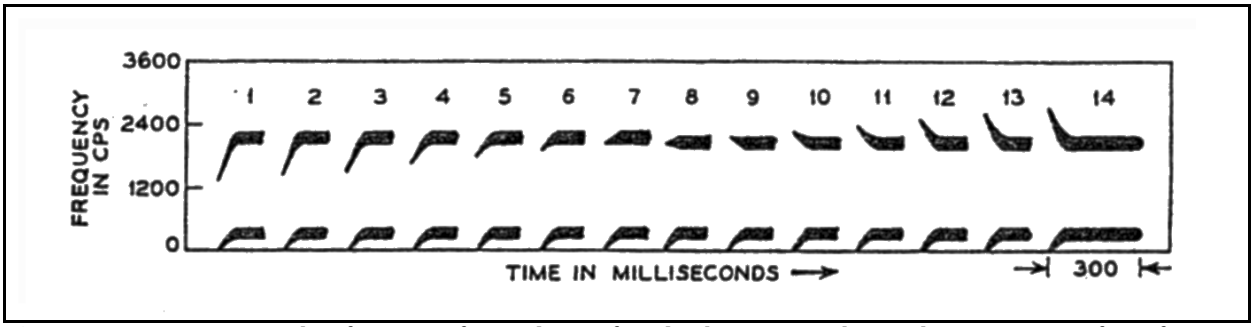

Figure U.4.3-3 Spectrographs of a series of voiced stops found in human speech. Each stop consists of two formants initially rising and then maintaining a constant value. See text. From Liberman et al., 1957.

The marginal variations in these voiced stops, b, d, and g indicate the variable frequency portion of the stop is the core consonant of the stop followed by the vowel component consisting of two constant frequency formants.

U.4.3.1.3 Recent contribution of Crystal

Crystal has provided a comprehensive text on human language from the perspective of a linguist[79], has provided a useful bibliography and has provided a copy of the current International Phonetic Alphabet. He notes there are 44 distinct phonemes in English out of a potentially much larger group. Some available combinations are not used in English but may be used in other human languages. He describes these 44 distinct phonemes (20 vowels and 24 consonants) in considerable detail. He also develops how they are typically used in groups of two or three and occasionally four phonemes to form syllables.

Fletcher, writing at a different time and using marginally different notation, provides spectrograms of 300 combinations of one consonant and one vowel[80]. The sequences look remarkably like dolphin sounds reported in the literature using low frequency recording techniques (less than 24 kHz). Even before Fletcher, engineers described the various spectral frequencies found in spectrograms of speech as formants. The lowest frequency formant was defined as the fundamental frequency generated by the larynx. The next higher formant, normally labeled F2 and occurring above 3 kHz, is that employed by McCowan in her contour similarity (CS) technique.

Paragraph S.4 in Appendix S of this work highlights the specific characteristics of a "language" used between a cohort of the same species (based on the studies of various linguists).

U.4.3.2 The complexity of the human language protocol

Words take on a broad range of meanings in any language over time. Some meanings are merely nuanced groups and others are entirely independent. The language protocol used within a natural (human) language can act as a framework for deciphering the communications of dolphins.

Figure U.4.3-4 shows a framework for the human language protocol expanded from Crystal[81]. Writing for an English speaking audience, he was unconcerned with the symbol set used or the pragmatics associated with the behavioral aspects (pragmatics) of significance when studying the communications of a species. The term faculty of language is easily confused with facility of language (reserved in this work for the physiological plant associated with vocalization). The term language protocol will be used in its place.

---

[79]Crystal, D. (2006) How Language Works. NY: Overlook Press

[80]Fletcher, H. (1953) Speech and Hearing in Communications. NY: Van Nostrand pg 61

[81]Crystal, D. (1987) The Cambridge Encyclopedia of Language. Cambridge: Cambridge Univ Press

In studying the vocalization of dolphin, ascertaining the symbol set employed, the dictionary of words (morphemes), the syntax, the semantics and the pragmatics related to their vocalizations are of major interest. While Crystal's term lexicon is appropriate in his discussion, it is the dictionary of unwritten vocalization that is of interest here. The actions (pragmatics) spawned by vocalizations can play an important role in defining the dictionary of, as well as the syntactic capability and semantics applicable to, dolphin communications.

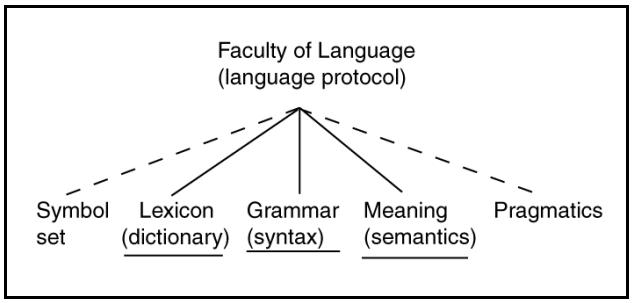

Figure U.4.3-4 The faculty of language or a language protocol. Writing for an English speaking audience, he was not concerned with the symbol set or pragmatics involved in the overall protocol. Expanded (dashed entries) from Crystal, 1987.

Phonology, a sub-field of grammar and syntax involves the rhythm, rhyme and intonation associated with words in vocalization. The phonology of a language can only be studied after the major fields are clearly understood. The elements of phonology will not be included in this appendix.

The lexical, syntactic and semantic breadth of language the dolphins have been demonstrated to understand, based on pragmatic studies, insures, in this analysts eyes, the inherent ability of the dolphins to communicate using language. The immediate tasks are two-fold. First is an attempt to isolate and define their phonetic symbol set. Second is to attempt to determine the range of objects and actions a dolphin is likely to use in everyday communications before attempting to decode the available recordings of these communications. These tasks are obviously closely intertwined.

A first step in the second task appears to have been accomplished. The dolphins respond to the names they have adopted for themselves previously. The challenge is to determine how they respond following a communication from another dolphin. If a sufficiently large set of detailed responses and detailed messages can be obtained, determining the meaning of the message content is reasonably straight forward using standard crytographic techniques. However, it is of primary importance that the symbol set used in communications be separated from other sounds and that this symbol set be determined as precisely as possible. The evidence strongly suggests that different pods use different symbol sets. Winn et al. have produced such a symbol set for the humpback whale[82]. The material is synopsized by Thompson, Winn & Perkins in Winn & Olla, 1979. More study of the literature will be necessary to determine if the appropriate analytical techniques have been applied to obtain more complete understanding. In 2002, Herman provided a good summary of the work he was aware of on The Dolphin Institute web page, www.dolphin-institute.org/resource_guide/animal_language.htm It expands on the recent work with the two dolphins described previously as Akeakamai and Phoenix.

The very rapid ability of at least one bottlenose dolphin to mimic a synthesized whistle is another example of how cognitively capable the animal is (**Section U.1.2.2**).

Pack & Herman have recently reported on the ability of the bottlenose dolphin to interpret posturing by humans, specifically the ability of the dolphin to understand human gazing and pointing[83]. In this case, gazing refers to the orientation of the head ***and*** eyes toward a point of interest. Phoenix and Akeakamai did exceedingly well on the defined test program without any prior training. It was obvious from the first test that the animals had an innate ability to interpret inter-species pointing and gazing that at least equaled that of the non-human great apes. It must be assumed that they are also aware of the posturing by their fellow dolphins. This posturing could be an important aspect of their intra-species communications.

The very recent observation of dolphins forming toroid rings from air bubbles and playing with them extensively underwater has come to a shock to many experienced dolphin observers.

https://www.youtube.com/watch?feature=player_embedded&v=wuVgXJ55G6Y

This capability is impressive and illustrates both the physical dexterity and the cognitive ability of the bottlenose dolphin. The capability has been observed to be a learned skill (that is learned quite rapidly from other members of a cohort). These capabilities are far beyond those of a human child of less than eight or ten years of age.

---

[82]Winn. H. Perkins, P. & Poulter, T. (1971) Sounds of the Humpback Whale *Proc 7th Ann Conf Biol Sonar and Diving Mammals* Stanford, Calif: Stanford Research Inst pp 39-52

[83]Pack, A. & Herman, L. (2007) the dolphin's (tursiops truncatus) understanding of human gazing and pointing: Knowing what and where *J Comp Psych* vol. 121(1), pp 34-45

U.4.3.2.1 Early work of the modern period– 1985 forward EMPTY

Pinkal & Thater, in a 2007 course in computational linguistics & phonetics, have reviewed the position of the dolphin with respect to a potential language using a framework from the human language protocol. The focii of their program includes "spoken language systems" and phonetics. The discussion is a sophisticated one involving such terms as ambiguity, in particular polysemy (the multiple uses of a single word), etc. These secondary traits highlight the complexity to be expected in dolphin communications and considerably complicate the deciphering of dolphin communications. However, these possibilities must be addressed or provided for within the overall framework.

They provided the following breakout of word-semantics in 2007[84].

Major word-semantic categories
• **Function words**:
– Connectives and quantifiers
– Auxiliary and modal verbs
– Temporal and modal adverbials
– Anaphoric pronouns, articles
– Degree modifiers, Copula, ...
• **Content words**
– Common nouns
– Full verbs
– Adjectives
• **Other**
– Named Entities (Persons, institutions, geographic entities, dates)
– Numbers
– Etc.

A major challenge is to determine the complexity of any dolphin language. Does it use both function and content words? Does it use named entities (tied to signature whistles)? Do the function and content words exhibit multiple meanings when used in different contexts?

> The following material will not use the International Phonetic Alphabet (IPA) symbols because of many readers may not have the correct Unicode installed on their computers. Most readers will not understand the symbols. Wells has provided access to the IPA phonetic symbols for a variety of computers. "If you cannot see a schwa here [ə, an upside down e] on your screen and you want to use these symbols, users of Windows 95/98/ME/NT/2000/XP and Macintosh OSX can download Lucida Sans Unicode free of charge and install it on your system, and/or install a newer browser."

The sophistication of the terminology used in the linguistic community is clearly seen in a document by Wells on syllabification and allophony (probably as viewed by a British academic)[85]. It includes a major discussion of whether dolphin is (should be) pronounced dol-fin or dolf-in. The paper provides precise definitions of many of the most specific terms used in linguistics, and probably used incorrectly frequently in this Appendix. It also includes many citation to high level academic sources in linguistics prior to 1984.

> The term RP is used throughout the Wells paper cited above. It stands for "Received Pronunciation." According to the BBC website, "Modern RP can be described as 'the speech of educated people living in London and the southeast of England and of other people elsewhere who speak in this way'. No two speakers of English speak the language identically. Every English-speaking country, and practically every region within those countries, has a distinctive accent. It is also true that meanings of words and grammatical forms can vary from one English-speaking country to another, but this entry is concerned solely with the standard pronunciation of British English. This is what linguists call 'Received Pronunciation' or 'RP'. RP refers exclusively to pronunciation, though it can be seen as analogous to Standard English ('SE')."

> As the BBC concludes, RP has practically disappeared except for its use by the Royal Family and some teachers of English as a second language.

---

[84]Pinkal, M. Thater, S. (2007) Semantic Theory: Lexical semantics I www.coli.uni-saarland.de Slide 9 *The overall program is dynamic and this lecture for the summer of 2007 may be hard to locate. Copy in Whdolphin/speech folder*

[85]Wells, J. (1990) Syllabification and allophony In Ramsaran, *ed.* Studies in the pronunciation of English. London: Routledge pp 76-86 http://www.phon.ucl.ac.uk/home/wells/syllabif.htm

U.4.3.2.2 SAMPA, a broadened (ASCII compliant) IPA symbol set

Various groups have proposed ASCII compliant phonetic symbol sets to replace the IPA standard that is now several decades old. While the recent expansion of the ASCII character sets and their inclusion in most PC operating systems, the problem has been ameliorated. However, character keys and keyboards are much less readily available.

A simple character replacement table, with examples, has been offered by Szynalski[86].

The Saarland University site(much of it in German) indicates a machine-readable phonetic alphabet is now available covering all of the major human languages. SAMPA (Speech Assessment Methods Phonetic Alphabet) is a machine-readable phonetic alphabet described in detail on the University College London website[87]. A translation table with annotations is also provided at that site.

The use of a version of the machine-readable SAMPA tailored for the sounds of the dolphin would greatly aide the deciphering process (a speech assessment method) anticipated in the plan of **Section U.5**.

U.4.3.3 A summary of relevant characteristics of Chinese - a tonal language

The characteristics of dolphin communications appear to exhibit the properties of a tonal language like Chinese more than an atonal language like English or Russian. Therefore, this section is provided as background to the preparation of a more detailed plan of attack on decrypting the language protocol associated with dolphin communications. First, an introduction to the phonetic character of Chinese.

Lisa Zhang has noted, “Among the reasons Moser thinks Chinese is hard are that the language is not phonetic, it has no alphabet, and it has a god-awful dictionary system[88]. These features or lack thereof appear mostly anecdotal. However, his supporting paper makes it abundantly clear that Chinese may be too difficult for a dolphin to emulate.

> Bopomofo[89] is the colloquial name of the zhuyin fuhao or zhuyin system of phonetic notation for the transcription of spoken Chinese, particularly the Mandarin dialect, and used alongside the Wade-Giles system, which used a modified Latin alphabet. The Wade system was replaced by Hanyu Pinyin in the Peoples Republic in 1958 and generally adopted elsewhere over time. Bopomofo remains widely used as an educational tool and electronic input method.
>
> The informal name "Bopomofo" is derived from the first four syllables in the conventional ordering of available syllables in Mandarin Chinese. The four Bopomofo characters that correspond to these syllables are usually placed first in a list of these characters.
>
> Zhuyin symbols include 21 consonants, 18 rhymes and medials and four tonal marks. Thus the total set is nearly twice the size of the English phonetic set but comparable to the set used in some European languages.
>
> Mandarin Phonetic Symbols II[90] is a romanization system formerly used in the Republic of China (Taiwan). It was designed to coexist with the above Zhuyin (non-romaized) system. The Wikipedia citation describes the classification of these phonetic symbols based on their nasal, plosive, affricate, fricative and lateral features as well as how they are formed by the vocal tract.

Zein has provided a tutorial on Mandarin Chinese Phonetics[91] designed for the novice. He notes, “In Chinese, each character corresponds to one syllable (which corresponds to a part of an English word, and entire word or more than one word). Chinese syllables consist of three elements: initial sound, final sound and tone. The initial sounds are consonants and the final sounds contain at least one vowel. Some syllables consist only of an initial sound or a final sound.”
“In Mandarin Chinese there are 21 initial sounds.

---

[86]Szynalski, T. (1990-99) Phonetic alphabets reference http://www.antimoon.com/ipa

[87]http://www.phon.ucl.ac.uk/home/sampa/index.html

[88]Moser, D. (2010) Why Chinese is so damn hard http://www.pinyin.info/readings/texts/moser.html

[89]http://en.wikipedia.org/wiki/Bopomofo

[90]http://en.wikipedia.org/wiki/Mandarin_Phonetic_Symbols_II

[91]Zein, P. (update 2008) Mandarin Chinese Phonetics, 3rd Ed. http://www.zein.se/patrick/chinen8p.html

In Mandarin Chinese there are 35 final sounds.
There are also seven special cases."
"The initial and final sounds make a total of 56 basic sounds. Combinations of initials and finals plus the special cases result in 413 possible combinations. Applying the four tones of Mandarin Chinese to this, we get a total of around 1,600 unique syllables." **Figure U.4.3-5** illustrates the four tones per Zein.

The use of only four tone contours in Chinese would suggest it is a much simpler language than the multiple tone patterns identified by the McCowan group (based only on the baseline frequency contour of whistles,1995b, figure 1). Alternately, consideration of the total whistle spectrogram as a single vowel may reduce the number of identifiable tone contours significantly.

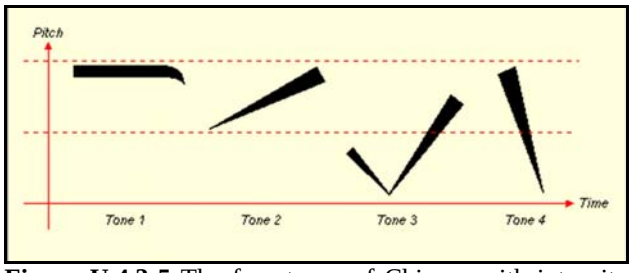

Figure U.4.3-5 The four tones of Chinese with intensity indicated by the width of the symbol. From Zein, 2008.

It is suggested by many linguists that the use of more than four tone patterns would make Chinese virtually incoherent to even native Chinese communicators. The same suggestion probably applies to any language used by dolphins.

Zein has recommended a frequency versus rank table of Mandarin (either modern or ancient) prepared by Jun Da[92] and containing the first 3500 characters (available as an Excel file. These files contain at least one link to a Zipf Diagram for Chinese. Zein notes a second statistical analysis is available on the website of Chih-Hao Tsai[93].

U.4.3.3.1 A Zipf diagram for Chinese

Figure U.4.3-6 provides a Zipf Diagram for Chinese provided by Ling[94] of Middle Tennessee State University. The straight line has been added. Even with 8,000 symbols in the data set, the curve is still not ideal. Multiple ranks show the same frequency of occurrence. However, the slope is very near -1.00 in this representation (the straight line).

[92]Jun Da (update 2010) Character frequency list http://lingua.mtsu.edu/chinese-computing/

[93]Chih-Hao Tsai (update 2005) Frequency and Stroke Counts of Chinese Characters http://technology.chtsai.org/charfreq/

[94]Ling, (update 2011) The Physics of Language http://www.mandarintools.com/cedict.html

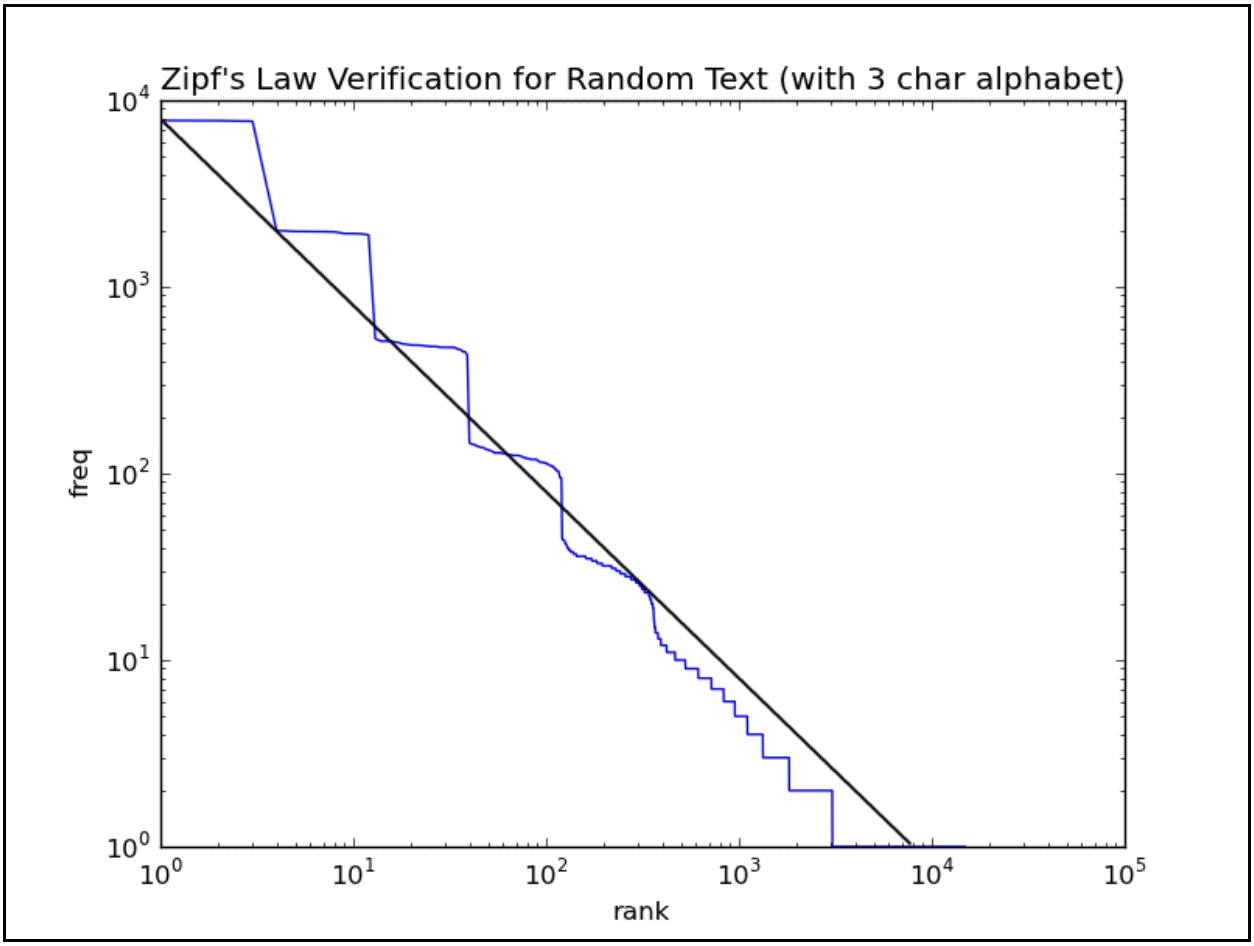

Figure U.4.3-6 Zipf Diagram for Chinese including 8,000 symbols. From Ling, 2011.

Ling developed this diagram in his search for meaning in speech communications. As noted in the introduction, he separates the use of the word language into a Sprachen and a faculty of communications. His current activity is well expressed:

> To further illuminate the distinction between Sprachen and Faculty, consider the means of transmission from parents to offspring. The Faculty is inherited by the offspring just as any other genetic traits are: through recombination of the parents genes. The Sprache, however, is not present in the genome, and it is not genetically transmitted from parents to offspring. Rather, the Sprache is trasmitted memetically to offspring through in-context audible utterances (speech). I am primarily concerned with the mechanism of evolution of Sprachen rather than that of Faculty
>
> Specifically, I believe that the Sprachen evolve over time by a process of analogous to neo-Darwinian (i.e., gene-based) natural selection. I will now give a sketch of this analogy:
>
> • A Sprache is like a species.
> • A speaker of a Sprache is like an individual.
> • Words are like genes.
> • The set of words known by a speaker is like the genome of that individual.
> • The average number of words known by a speaker is like the number of genes in the genome: Both numbers are approximately twenty thousand.
> • The lexicon of a Sprache is like the gene pool of that species.
> • The learning of new words is like the duplication of genes during reproduction.
> • Word changes in the new generation are like gene mutations.
> • The use of a word in an expression is like the expression of a gene in a biological function.
> • Words are communal: they never act alone, but rather in concert with other words to form meaningful expressions. This is like the simultaneous expression of multiple genes to perform a single function.
> • Words are subject to selection like the natural selection of genes. The mechanism works as follows: the probability of a word being learned varies with the usage frequency of the word. When a person makes a

speach act, he chooses to use words which best convey his meaning. In this way, we have a positive feedback loop which selects useful words and proliferates them.
• A speaker of a particular Sprache is resistant to the assimilation of words from other Sprachen. This is like how there is an incompatibility of genes, or at least of mating, between species.
• A word consists of form and function. The form of a word is the pronunciation. The function of a word is its symantics. This is like to the gene, whose form is its coding in DNA, and whose function is its biological role. Knowledge of the form of a word or gene in isolation is insufficient to explain the function of that word or gene: the form is only intepretable in context.

So, we see that there is a process of selection occuring at the word level which is highly analogous to the natural selection of neo-Darwinism.

I want to push this analogy to see how far it can go. I want to learn where it fits, where it doesn't, and why. I want to carry over the tools used in synthetic neo-Darwinism into linguistics. I want to look for empirical evidence for this process. I want to develop computer simulations and mathematical models where they are needed. I want to continue my litterature search for others' ideas on this topic. Finally, I want to tie this back into the biology, so as to gain a better understanding of the symbiotic relationship between Sprachen and Faculty."

While Ling's assertions may not all follow the findings and a ssertions of this work, they show he is seeking the same objective.

U.4.3.4 A summary of signals used by dog trainers

McConnell & Baylis have documented eleven "words" used by fourteen shepherds to work their dogs using whistles[95]. The spectrograms are simplified and without time or frequency scales. They also discussed a few instructions used by horse trainers as well. The individual whistles are strongly suggestive of a phoneme structure within the words, including pauses between such phonemes.

U.4.4 Potential sentence content in dolphin communications

Crystal has provided a brief discussion of how languages might have evolved. He describes five theories described by Jespersen in the early 20th Century (page 350). They do not include the navigation hypothesis described below. Crystal listed Jespersen's categories using fanciful names.

1. The *bow-wow theory*–Speech arose through people imitating the sounds of the environment, especially animal calls.
2. The *pooh-pooh theory*–Speech arose through people making instinctive sound, caused by pain, anger, or other emotions, such as ooh or tut-tut.
3. The *ding-ding theory*–Speech arose because people reacted to the stimuli in the world around them.
4. The *yo-he-ho theory*–Speech arose because, as people worked together, their physical efforts produced communal, rhythmical grunts, which in due course developed into chants.
5. The *la-la theory*– Speech arose through the romantic activities of humans, sound associated with love, play, poetic feeling, perhaps even song.

These theoretical evolutionary routes appear less than satisfying (convincing) even with the short justifications given by Crystal.

Kazakov & Bartlett have speculated on the potential for "Navigation" to form the instigating mode of intra-species communications[96]. They have noted the use of songlines by the Australian aborigines for this purpose during their development prior to the appearance of writing and maps. It appears likely that communicating routes to specific reefs, beaches and mouths of estuaries may be particularly important to dolphins.

Beyond the evolution of language, Crystal has also addressed the structure of languages. He describes the structure in terms of:

1. The vocabulary of the language–the ensemble of words used to express individual meanings.

---

[95]McConnell, P. & Baylis, J. (1985) Interspecific Communication in Cooperative Herding: Acoustic and Visual Signals from Human Shepherds and Herding Dogs *Ethiology* vol 67(1-4), pages 302–328

[96]Kazakov, D. & Bartlett, M. (2005) Could navigation be the key to language Proc 2nd Symp Emerg Lang @ York Univ.

2. The grammar of the language–the order in which words are presented.
3. The phonology of the language–the tone of voice in which a sentence is presented.

Phonology subdivides into pure phonology, how sounds are used in a language and phonetics, how speech sounds are made, transmitted and received. Pure phonology is often subdivided into segmental phonology (the study of vowels, consonants and syllables) and non-segmental phonology (the study of prosody–the style of word presentation as in verse).

Semantics is concerned with the choice of specific words of similar meaning within the vocabulary to more specifically shape the meaning of a sentence.

Grammar can be broken down into the study of word structure (morphology) and the study of word order (syntax).

The syntax of a sentence is straight-forward. Complex sentences contain clauses, that contain phrases, that contain words, that contain morphemes. The parts of a sentence can be diagramed using a tree diagram.

However, the grammar (i.e. the order of presentation of the elements in the tree diagram) is highly variable and depends on the language involved.

Two recent books provide significant background relative to human speech. McWhorter, has provided "The power of Babel : a natural history of language[97]" describing many of the common attributes of languages. Anthony has provided "The Horse, the Wheel and Language" as an exposition on how human languages have evolved beginning about 4500 years ago[98].

U.4.4.1 A symbolic notation accommodating four potential sound generators

Ostrovskaya & Markov have defined a structure of symbolic sounds for describing the short interval-type sounds of dolphin[99]. They have allowed for up to four simultaneous sources. This complicates their framework considerably. The paper suffers somewhat at the technical level in translation. It appears an initial exploration using their symbology but limiting the sources to one could be productive. However, it appears from recordings observed to date that a realistic model must include tonal sounds. The dolphin language appears to blend tonal– and stressed–language elements.

U.4.5 The ethology of dolphins responding to specific vocalizations EXPAND

Very significant data has been collected in the field regarding the response of dolphins to communications from other dolphins. Much of this data was collected in the third quartile of the 20th Century when the instrumentation required to accurate record dolphin vocalizations was limited.

The paper by Lilly & Miller of 1961 is very valuable from the ethology perspective but the use of a mirror galvanometer and an ink writer are indicative of their primitive technology[100]. Most of their data was at less than 20 kHz although one graph shows waveforms truncated at 32 kHz. They did assert the dolphins were capable of generating three classes of sounds based on at least two independent sound generators. Their classes are compatible with but form a different matrix than developed in this appendix. The paper explicitly exempted echolocation signals from its discussion.

The papers of Caldwell & Caldwell (including the 1968 paper) are quite valuable (in the waters near Los Angeles, California). Much of the work at Shark Bay in the fourth quartile is also very valuable in the ethological context.

A book by Leatherwood & Reeves in 1990 contains much good behavioral material[101], particularly the chapters by Caldwell & Caldwell and by Scott, Wells & Irvine. Harcourt & de Wal edited a book in 1992 on coalitions and

---

[97]McWhorter, J (2001) The power of Babel : a natural history of language. NY: W. H. Freeman

[98]Anthony, D. (2007) The Horse, the Wheel and Language. Princeton, NJ: Princeton Univ Press

[99]Ostrovskaya, V. & Markov, V. (1992) A language to describe the structure of pulsed sounds in bottlenose dolphins (*Tursiops truncatus montagu*) *In* Thomas, J. Kastelein, R. & Supin, A. *eds*. Marine Mammal Sensory Systems. NY: Plenum pp 393-414

[100]Lilly, J. & Miller, A. (1961) sounds emitted by the bottlenose dolphin *Science* vol 133, pp 16891693

[101]Leatherwood, S. & Reeves, R. *eds*. (1990) The Bottlenose dolphin. San Diego, Ca: Academic Press

alliances that includes good material by Conner and associates[102]. The 1996 paper by Connor et al. is also quite useful[103]. The material describes the courtship of female dolphins by males in considerable detail that provides a valuable background for combining the vocalizations of the males with the ethology of the females. The companion ethology work at Shark Bay by Connor & Smolker (1996) using more sophisticated instrumentation and more complex protocols is very useful. A related The book by Turner (2013) while written for a general audience includes a wealth of ethological data.

The 2000 text by Mann et al. is a treasure of ethological information, however, when it veers into interpretation related to the sensory and vocalization capabilities of the dolphins it reverts to considerable conjecture and the "common wisdom" of an earlier era as taught to even current students[104]. The description of the auditory modality of the dolphin is particularly archaic. Table 4.1 provides a comparitive table of dolphin cohort sizes employed in a wide variety of earlier behavioral studies. The chapter by Tyack is quite valuable but is not tightly edited. The text (page 289) and graphics (attributed to Overstrom, 1983) would suggest the dolphin was making sounds through its open mouth during "jaw claps." However, Tyack makes clear this is not the case on page 277; "This means that cetaceans do not need to open their mouths or blowholes when they vocalize underwater." The suggested framework for analyzing social structure attributed to Hinde (1976) on page 68 may remain viable to this day.

U.4.5.1 Major categories of dolphin activities

Slooten provided **Figure U.4.5-1** to provide a framework for describing the ethology of the small Hector's dolphin in its home waters off New Zealand[105]. At maturity, it averages only 1.5 meters in length. She begins with the assertion, "Behavior unfolds as a time sequence of different behavioral events and states, whether one is observing an individual or a group. In the case of a lone individual, behavior is likely to be influenced by internal motivational factors such as hunger and hormone levels as well as external factors such as weather conditions and food distribution. In pairs and groups, additional factors become important, as the behavior of each individual may be influenced also by the social context and by preceding behavior of the other(s)." She defined 25 specific behaviors in an appendix and then continued, "Sequences of behaviors can be analyzed to study the underlying principles, or rules, that structure behavior (Berridge, 1990; Meyer and Guillot, 1990; Waas, 1991; Wolffgramm, 1990) or, alternatively, to provide an empirically based classification of categories of behavior." And, "In this paper, I have attempted a repeatable classification of behavior of Hector's dolphin (*Cephalorhynchus hectori*) based on sequence analysis."

"In this paper, the term behavioral transition is used to mean a preceding-following relationship of two behaviors that occurred within the same group of dolphins. For example, the behavioral transition lobtail-vertical jump is a sequence where vertical jump follows lobtail within a specified time period (see Appendix I of the paper for behavior codes)."

---

[102]Connor, R. Smolker, R. & Richards, A. (1992) Dolphin alliances and coalitions *In* Harcourt, A & de Waal, F. *eds*. Coalitions and Alliances in Humans and other Animals. NY: Oxford Univ Press Chap. 15

[103]Coonnor, R. Richards, A. Smolker, R. & Mann, J. (1996) Patterns of female attractiveness in Indian Ocean bottlenose dolphins *Behavior* vol 133, pp 37-69

[104]Mann, J. Connor, R. Tyack, P. & Whitehead, H. (2000) Cetacean Societies: Field Studies of Dolphins and Whales. Chicago, Il: Univ. Chicago Press

[105]Slooten, E. (2014) Behavior of Hector's Dolphin: Classifying Behavior by Sequence Analysis *J Mammalogy* vol 75(4), pp 956-964

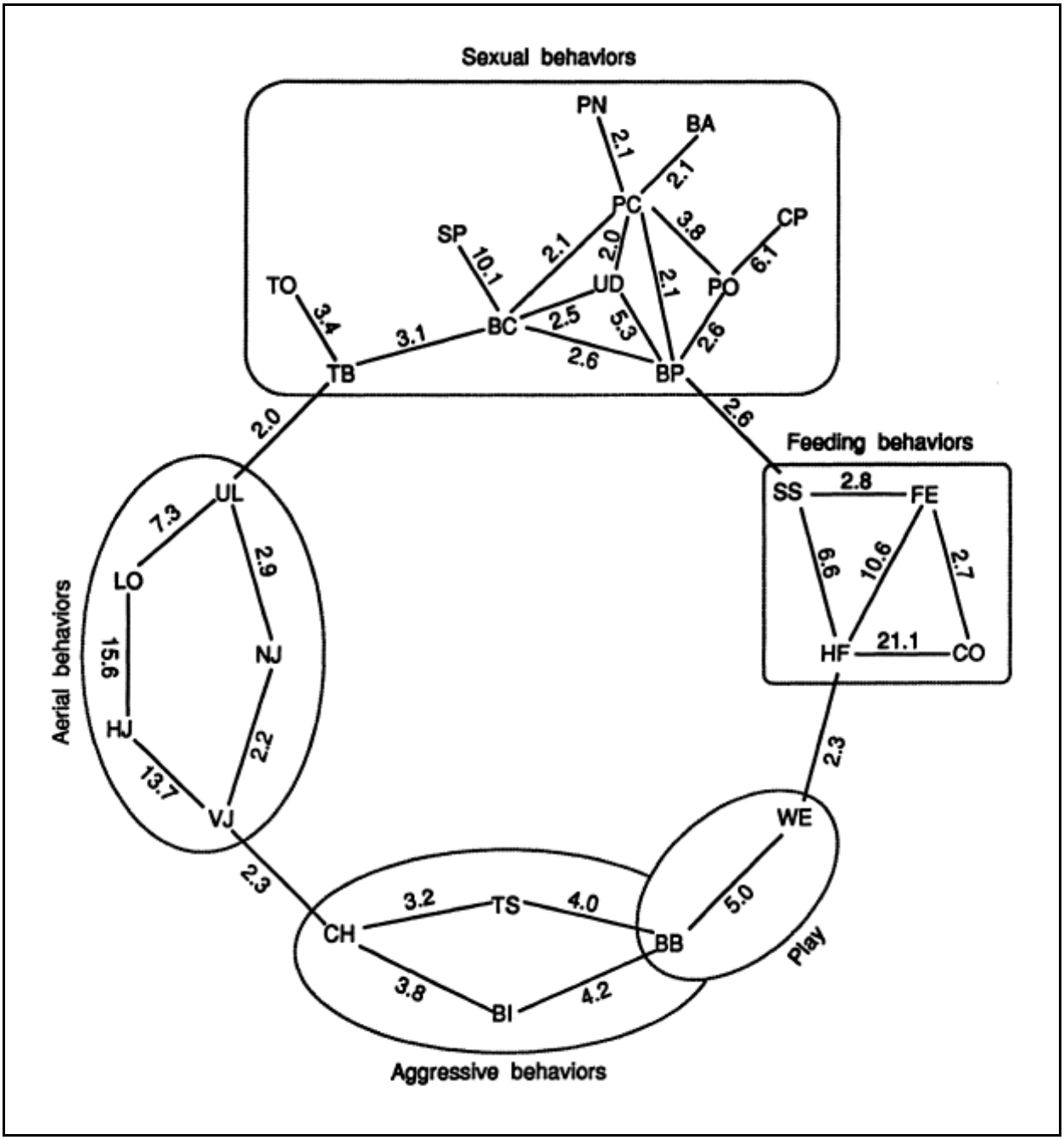

Figure U.4.5-1 Sequence diagram for 2-min interval between preceding and following behaviour, showing all links with a *z*-score ≥2. See text for behavior codes. From Slooten, 2014.

"The sequence diagrams (Figs. 1 and 2) revealed five behavioral categories with stronger associations within than between categories. The sexual category included copulation, possible copulation, and penis out. The aggressive category contained bite, chase, and tailsplash. The feeding behaviors feeding, swimming on side, chin-out, and horizontal flex were among the most closely associated behaviors and were negatively associated with most other behaviors. Three different types of leap and two forms of lobtailing were closely associated and were labeled aerial behaviors."

"The behavior bubbleblow was unusual in that it provided a link between some rather different behaviors. It was the only behavior that was difficult to place in a category and belonged to a different category in the two analyses. Bubbleblow was strongly associated with the aggressive behaviors and also with play with weed, the only clear-cut play behavior recorded."

"A striking feature of the classification of behaviors was that behavioral categories tended to be linked with two other categories forming a relatively simple circle shaped network."

"Most of the behaviors observed in this study have been observed in many other species of dolphins, porpoises, and whales."

U.4.5.2 Head snapping and jaw clapping EXTEND

Norris et al provided a paper on the collateral behaviors and rhythmic patterns of dolphins particularly associated with echolocation[106]. However these motions must be accounted for in any discussion of the pragmatics of dolphins related to language.

U.4.6 [Reserved]

U.4.7 Early attempts to analyze and describe dolphin speech or at least signature whistles

Since Lilly (1978) followed by Caldwell & Caldwell (1990), many people have attempted to describe the linguistic aspects of dolphin speech from a variety of perspectives. Most of these attempts have been heavily philosophical and generated by members of the general population. Most rely heavily upon anecdotal data and folklore. Prior to 1990, most of the data was recorded with quite primitive sound equipment. Herman & Tavolga provide a table of "known or suspected nonwhistling odontocete species." This table of records before1977 almost certainly contains small whistling dolphins with whistles not capturable with the available recording equipment. The footnotes to the table are informative.

Much of the vast literature using the terms dolphin and language, or dolphin and speech, have actually been concerned primarily with inter-species (human/dolphin) communications generally using hand or body signals on the part of the human. Occasionally, experiments using recorded dolphin sounds in an apparently completely irrelevant context have been performed (requiring the dolphin to treat these sounds as a foreign language that is only casually similar to the dolphin's native language).

U.4.7.1 Early work of the modern period– 1984 forward

A problem with most of the recorded dolphin sounds to date has been their limited bandwidth. Most recordings have not exceeded a bandwidth of 24 kHz whereas the communications (non echolocation) range of dolphins extends to at least 40 kHz. Another problem is the lack of a set of dolphin signatures that can be compared by other researchers. As will be seen below, the signature whistles described by the Richards team do not resemble the putative signature whistles of the McCowan, or other, teams.

U.4.7.1.1 The papers of Richards and associates, 1984

Richards and associates carried out a multi-year program at Kewalo Basin Marine Mammal Laboratory in Hawaii designed to explore two distinct situations. First, the ability of bottlenose dolphins to hear, interpret, understand and respond to artificial acoustic "words" in a "foreign" or synthetic dolphin language employing all of the elements of a human language protocol as shown in **Section U.4.3.2**. Second, to show the same communications could be carried out between one or more humans and the properly trained dolphins using a similar sign language protocol. The published reports are quite extensive and detailed. However, the test equipment was primitive by today's standards. The first was by Richards, Wolz & Herman describing the preparations and training program. It focused on characterizing the ability of the dolphins to mimic a set of presumed word like sounds and to associate those sounds with objects (thereby establishing a lexicon).

Richards, Wolz & Herman have quantified the repertoire of the bottlenose dolphin in preparation for their mimicry experiments[107]. They reported their dolphin was comfortable whistling in the 4 to 16 kHz range with modulation rates of 0 to 11 Hz. Abrupt square, sine and triangle waveforms were mimicked with relative ease. Akeakamai preferred to whistle in the 5 to 10 kHz range. For sounds outside of that range, she generally transposed her mimicry to a frequency within that range. **Figure U.4.7-1** shows a sampling of whistles obtained from a pair of dolphins. The time scale of this early work is coarse to near useless. Its only value may be in pedagogy.

---

[106]Norris, K. Howard, C. Marten, K. Wells, M. Wells, R. & Green, H. (1990) Collateral behaviors and rhythmic patterns In dolphin echolocation http://www.dtic.mil/dtic/tr/fulltext/u2/a246420.pdf

[107]Richards, D. Wolz, J. & Herman, L. (1984) Vocal mimicry of computer generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus. J Comp Psych* vol 98(1), pp 10-28

Because of the limitations of their equipment, Richards, Wolz & Herman could not know if the dolphin preferred a higher range or was in fact emitting a more complex signal that included a higher range. Most bottlenose dolphins are comfortable making sounds up to 40 kHz or higher.

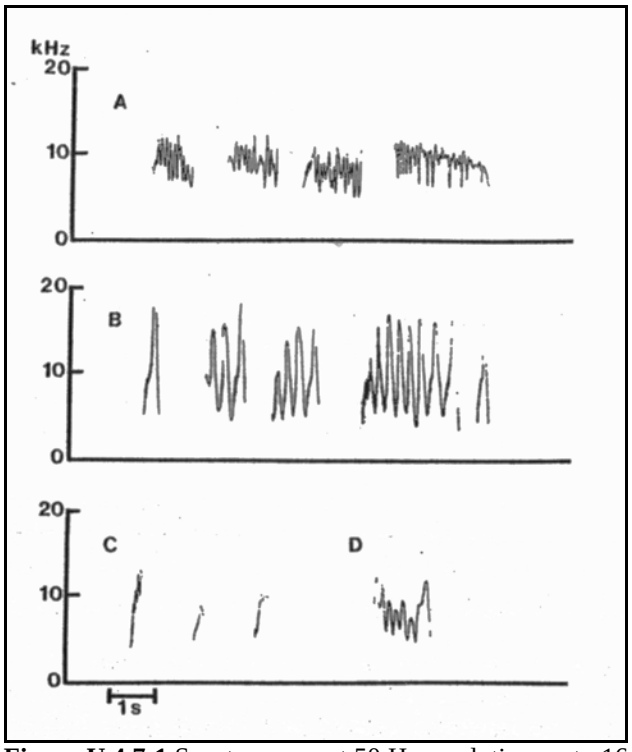

Figure U.4.7-1 Spectrograms at 50 Hz resolution up to 16 kHz from two dolphins. A; whistles characteristic of Akeakamai. B; whistles characteristic to Phoenix. C; short whistles of unknown origin. D; long whistles similar to those produced by both dolphins. From Richards et al., 1984.

Herman, Richards & Wolz provided the second paper describing the ability of their dolphins to understand and take action in response to acoustic instructions as well as human gestures[108]. The paper consisted of 90 typeset pages. The acoustic tests employed one dolphin, Akeakamai, Their multi-year program included exploration of all categories of the language protocol described by Crystal in **Section U.4.3.2**. It specifically included an exploration of the ability of Akeakamai to interpret and respond to recursive instructions.

Figure U.4.7-2 shows a brief selection of the acoustic signals they used as words in their "foreign" acoustic language experiments. Typically, they used objects from the middle row, actions from the lower row and action modifiers from the top row to construct sentences. Their table 1 provided a more complete set of objects, actions and action modifiers. The repertoire they worked with included about 35 individual words chosen to avoid conflict with common dolphin vocalizations. The specific words used are described in the Herman et al. paper. The acoustic words extended up to 35 kHz. The goal was to make the words members of an orthogonal set by varying their frequency, duration and modulation. It is interesting to note Herman et al. used a slowly modulated (2.5 Hz) tone at 30 kHz with a duration of 1.5 sec to mean "surfboard." This tone was outside the recording range of their acoustic test equipment.

The low frequency waveforms used as the words for pipe, water and frisbee appear challengingly low based on the audiograms for the dolphin.

- - - -

The sentences always began with a specific dolphin's name as defined in the computerized symbol generator as a salutation. A comprehensive set of sentences were created. Their table 2 describes the rules of syntax used to create sentences of up to five words.

The animals were able to understand imperative sentences consisting of object + action + indirect object and they were able to respond properly to compound sentences, lexically novel sentences, structurally novel sentences and semantically reversible sentences. Interestingly, when given an impossible instruction the first time, they searched extensively for the missing object. After they were made aware that the instruction might be false (and with very little reinforcement), they performed a complete but more modest search before returning to their station and contacting the NO paddle. It was also clear that their interpretation of the object PERSON referred to the symbolic person, as it mattered little which attendant took the person-position next to the dolphin tank.

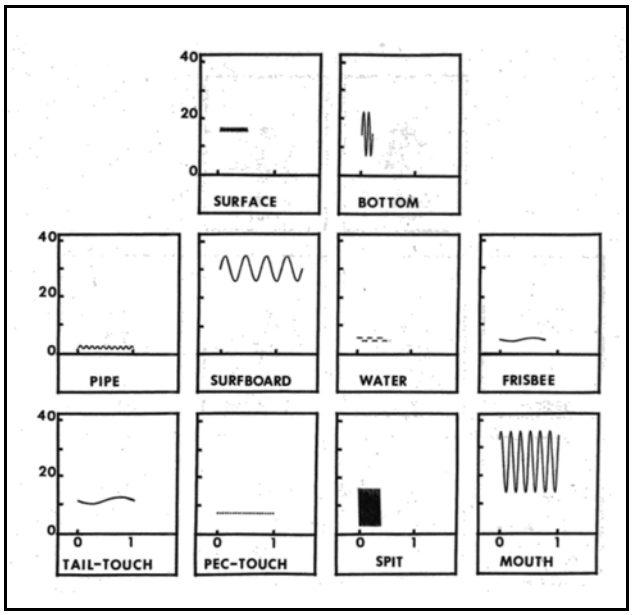

Figure U.4.7-2 The modulated waveforms of selected sounds of the acoustic "foreign" language used by Herman et al. Top row; action modifiers defined by Herman et al. Middle; objects defined by Herman et al. Bottom; actions defined by Herman et al. The word for surfboard was clearly outside the recording range of their test equipment and of human hearing. From Herman et al., 1984.

Their training extended up to five word instructions following a signature whistle describing which dolphin was to implement the instruction. This signature whistle was a computer-generated duplicate of the whistle the dolphin appeared to use to identify itself prior to the training program. As of April 1982, the repertoire of Akeakamai consisted of 464 unique sentences. Tables 7, 8 & 9 describe the statistical performance of the dolphin for each type of sentence. "Akeakamai responded wholly correctly to 83% of the 308 unique sentences given her. Both novel and familiar sentences were included in these totals."

The General Discussion surfaces a variety of results indicating some adjustment to the test protocol may be advisable, particularly relating to indirect objects. The dolphins managed to largely destroy one of the test objects made from plastic piping in the shape of an octagon.

They noted that the dolphins do not6 initially appreciate being given faulty or impossible instructions. A modification was introduced into their training protocol that allowed them to indicate a failed experiment not due to

[108]Herman, L. Richards, D. & Wolz, J. (1984) comprehension of sentences by bottlenosed dolphins *Cognition* vol 16, pp 129-219

their error (typically the absence of a specified object from their tank).

The dolphins were also able to understand recursive instructions similar to those understood by humans (page 184). They were observed to have trouble when their names were inserted into the recursive phrase. It is expected this disturbed the information extraction procedure just as it would in humans. Specifically, they were tested using two "recursive sentences" as follows;

- type a– Agent +Object + Action + Object + Action

- type b– Agent + Object + Action +*Agent* + Object + Action

The dolphin had a problem with the type b sequence because it had been taught that its name given after a sentence was to be considered a recall command after an incorrect response. Therefore, the second use of the dolphins name (agent in the sequence) in type b could be considered ambiguous. Herman et al. describe the resulting actions of the second dolphin, Phoenix, to this sentence type. It frequently interpreted its second occurrence of its name as a cancellation of the first instruction and it proceeded to perform the second action. The ability of Akeakamai to interpret and respond to anomalous sentences was explored in Table 11.

Begining on page 191, Herman et al. compared the performance of their dolphins to children. When passing from the one-word to the two-word level, children of 1:4 and 2:0 years of age performed similarly to the dolphins on 2-word tests. While the comparisons were brief, they noted, "What we are emphasizing here, however, is the use of the child data to illustrate that the type and levels of linguistic processing exhibited by the dolphins are substantial." Their discussion of each category of the language protocol is extensive.

The lexical and semantic breadth of language the dolphins were able to understand insures, in this analysts eyes, the inherent ability of the dolphins to communicate using language. This assertion includes the fact they were able to handle recursive sentences of moderate complexity. *The ability of at least one dolphin to process recursive sentences in a a totally acoustic "foreign" language appears to topple the last bastion of Chomsky's argument restricting language (as opposed to communications) to humans* (See citation 3).

U.4.7.1.2 The papers of Tyack and associates beginning in 1986

In 1986, Tyack presented a study on the capability of dolphins to mimic each other[109] as well as how they exchanged whistles that would normally be described as signature whistles using the definition of Caldwell & Caldwell. Sayigh, Tyack et al. followed the Tyack study with a more comprehensive study of mother-offspring pairs[110]. Their Abstract notes, "Mother-calf whistle exchanges were recorded from temporarily captured free-ranging bottlenose dolphins from 1975 to 1989. This is part of a long-term research project studying social structure and behavior of a community of approximately 100 dolphins in waters near Sarasota, Florida. Analysis of whistle exchanges from 12 mother-calf pairs shows that signature whistles can remain stable for periods up to at least 12 years. We looked for effects of vocal learning on the development of the signature whistle by comparing whistles of calves to those of their mothers. Eight female calves produced whistles distinct from those of their mothers, while four male calves produced whistles similar to those of their mothers. Male calves appeared to produce a greater proportion of whistles other than the signature whistle (termed "variants"). We hypothesize that these sex differences in whistle vocalizations may reflect differences in the roles males and females play in the social structure of the community."

> ***A caution here***. Many researchers describe the lowest frequency spectral contour above 5 kHz as the fundamental frequency of a dolphin whistle. This may be the fundamental frequency of unvoiced whistles but it is not the lowest frequency of voiced whistles and other voiced sounds. Such other sounds are typically harmonics of the signal produced by the larynx of the dolphin in the range of 200-600 Hz. A whistle in the 5 kHz region is more properly described as a baseline frequency when unvoiced. When voiced, it is typically the tenth harmonic of the fundamental frequency generated by the larynx. The laryngeal frequency is also known as the first formant, F1 in communications research. The baseline frequency is nominally the second formant, F2.

The many spectrogram sets of mother-offspring pairs provided by Sayith & Tyack are very helpful in establishing the fact that male offspring tend to mimic their mothers presumed signature whistle initially and then adopting a close simile as they mature. Female offspring appear to adopt a significantly different signature whistle at a very

---

[109]Tyack, P. (1986)Whistle repertoires of two bottlenosed dolphins, *Tursiops truncatus*: mimicry of signature whistles? *Behav Ecol Sociobiol* vol 18, pp 251-257

[110]Sayigh, L. Tyack, P. Wells, R. & Scott, M. (1990) Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: stability and mother-offspring comparisons *Behav Ecol Sociobiol* vol 26, pp 247-260

early age. They speculate on the sociological importance of this pattern. The sets also establish the long term stability of these waveforms over time. They note, “Favored whistles of free-ranging adult female Tursiops are markedly stable. Spectrograms of favored whistles from one matrilineal group, spanning three generations, further indicate consistency of individual whistles over time (Fig. 7). Mother no. 22 was recorded in 1976 with female calf no. 2, and again in 1987 with a different female calf, no. 148. The structure of the favored whistle of mother no. 22 remained stable over this 11-year period.”

Mother #8 exhibits a prolonged pulse-sound buried within her proposed whistle signature in 1987 (and apparently present in the 1976 and 1984 spectrograms) in figure 4.

One of Sayigh & Tyacks conclusions is important, “It is not entirely clear to what extent the similarities that we perceive in the general characteristics of whistle contours are meaningful to the animals. Richards et al. (1984) found that captive Tursiops were extremely adept at mimicking various computer-generated frequency contours. In some cases the animals would transpose the frequency while mimicking the contour. Tyack (1986) found that two captive bottlenose dolphins could mimic the frequency contour of one another's signature whistles very precisely, although duration varied by as much as 30%. These studies indicated that frequency contours recognizable by visual inspection were also recognized by the animals. It is very difficult, however, to quantify whistle similarity when certain aspects of the whistle such as duration, highest and lowest frequency, and number of loops are not characteristics that are consistently conserved by the animals.”

Buck & Tyack provided another valuable paper in 1983[111]. Their Abstract reads, “Bottlenose dolphins (*tursiops truncatus*) produce individually distinctive narrow-band "signature whistles." These whistles may be differentiated by the structure of their frequency contours. An algorithm is presented for extracting frequency contours from whistles and comparing two such contours. This algorithm performs nonuniform time dilation to align the contours and provides a quantitative distance measure between the contours. Two recognition experiments using the algorithm on three dolphin whistles f rom each of five individuals classified 15 out of 15 single-loop whistles correctly, and 14 out of 15 central loops for multiple-loop whistles correctly.” They employed a different protocol from most experimentors, “To evaluate the feasibility of the algorithm, we tested its performance on the task of "out-of-set" recognition of two sets of bottlenose dolphin whistles. Out-of-set recognition experiments use a library of "known" signals to identify a set of unknown signals. Each of the unknown signals is compared against all the known reference signals, and a preferential listing of the reference signals is generated in order of increasing distance. The unknown signal is assigned to the nearest reference signal. Each of the two experiments described here used a dictionary of five reference whistles chosen as typical of five different dolphins.” “It should be noted that all the recordings used for these experiments were made with contact hydrophones and had reasonably high signal to noise ratios.” They concluded, “The preliminary results presented here suggest that the fundamental frequency (baseline frequency, editor) is a primary key in individual identification for bottlenose dolphin signature whistlesBased on this criteria, the nonuniform time-dilation algorithm is well-suited for quantitative testing of the signature whistle hypothesis.”

- - - -

A problem with the spectrograms in the Tyack and colleague papers is their publication of small, frequently heavily cropped images. These frequently omit details related to the larynx frequency of the dolphin and any potentially important consonant(s) acting as a prefix (adjective) to the whistle signature. Some of the spectrograms in the Sayigh & tyack (1990) paper are more inclusive and even illustrate several burst-pulse-sound waveforms within a sequence of whistles. Figure 8, reproduced as **Figure U.4.7-3** also shows an unvoiced phoneme preceding a whistle sequence for “female #15” in 1987. The overall signature of female #15 is thereby significantly different from her mother’s overall signature. In the same figure, showing data recorded in 1984, female #15 also shows some unvoiced sound not appearing in her mother’s whistle sequence of the same time. The unvoiced content differs from that in the 1984 pattern. See **Section U.2.3.1** for a discussion of unvoiced sounds.

[111]Buck, J. & Tyack, P. (1983) A quantitative measure of similarity for *tursiops truncatus* signature whistles *J Acoust.Soc Am* vol 94(5), pp 2497-2506

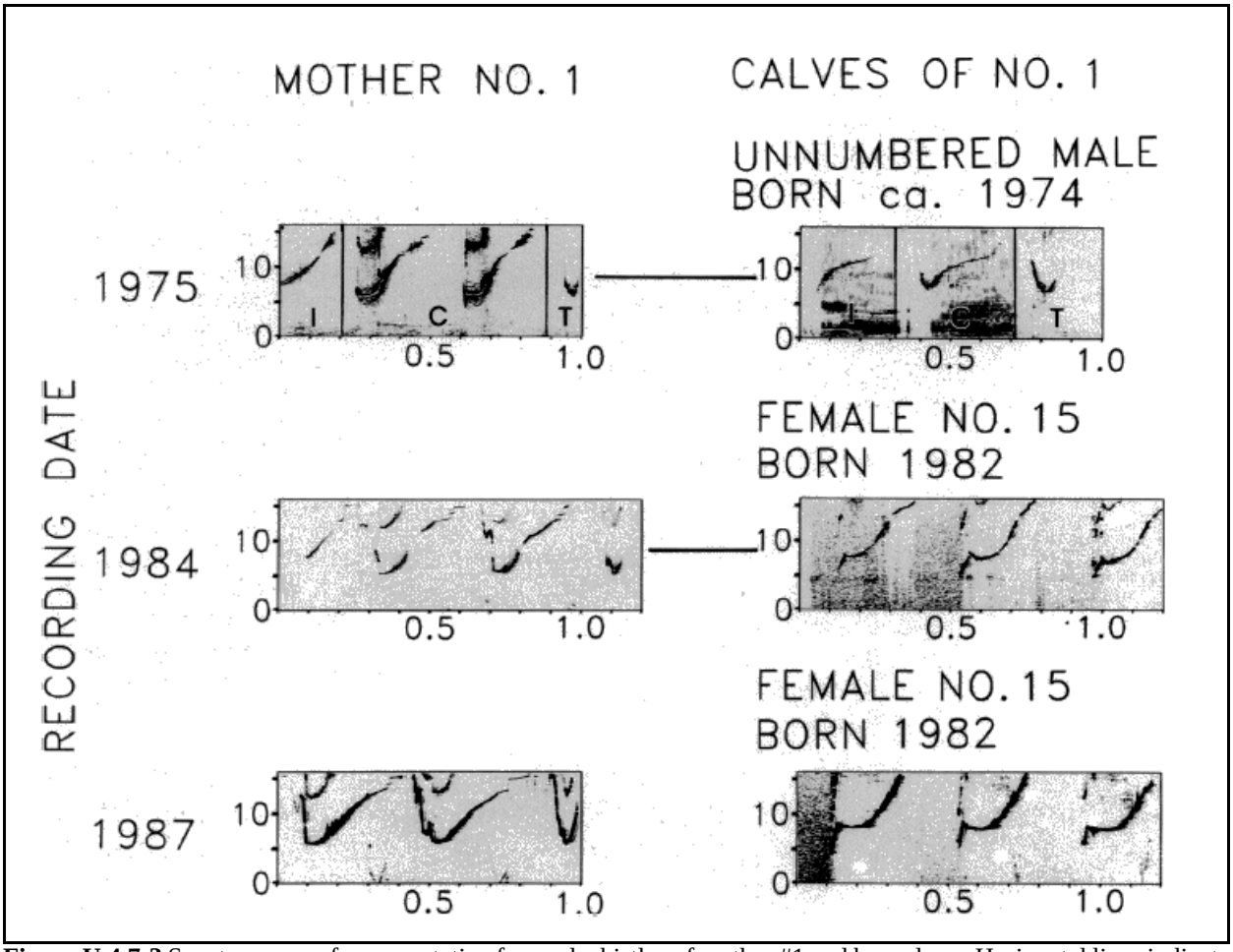

Figure U.4.7-3 Spectrograms of representative favored whistles of mother #1 and her calves. Horizontal lines indicate the waveforms were obtained at the same time. Their short term capture was simultaneous. I; introductory loop. C; central loop(s). T; terminal loop. See text. From Sayigh & Tyack, 1990.

- - - -

Preliminary analyses by this author up through 2014 suggest the sounds produced by dolphins are extremely complex and rival those of humans, except in a different frequency range. They are quite capable of generating two distinct whistles simultaneously that actually cross in frequency in the middle of the morpheme(s). Because of the multiple sources of sound in the dolphin system, their sound vocabulary may be even more versatile than that of humans. As a result, it appears a meaningful analysis leading to understand dolphin language will require establishing at least a framework representing the broad phonology of the dolphin and a skeleton of a vocabulary. Only with such a framework and skeleton in hand can any attempt be made to determine the grammar and syntax of the dolphins.

U.4.7.1.3 The 1990 paper of Caldwell, Caldwell & Tyack

Caldwell et al. provided a major paper in 1990 that included their definition of a signature whistle[112]. They make an important opening statement, "The sounds produced by dolphins fall into three categories. Two are broad-band: the click sounds used in echolocation and the burst pulsed sounds often describes as squawks, yelps, barks, etc. The third category, the whistle, is a frequency-modulated narrow-band sound *which appealed to us because of its relative ease of analysis and categorization*." The italics have been added. Based on their observations of a great number of whistles, they offered a "signature whistle hypothesis." "This hypothesis arose from our observation, made in 1965, that each dolphin in a captive group tended to produce whistles which were individually distinctive and stereotyped

[112]Caldwell, M. Caldwell, D. & Tyack, P. (1990) Review of the Signature-Whistle Hypothesis for the Atlantic bottlenose dolphin *In* Leatherwood, S. & Reeves, R. *eds*. The Bottlenose Dolphin NY: Academic Press.

in certain acoustic features. Since the individually distinctive features of these whistles were so striking to us, we called them signature whistles." They went on to recognize several important facts, "While some acoustic features of *an individual's whistles change as a function of behavioral context*, as discussed below, others do not. We have hypothesized that the individually distinctive attributes of signature whistles function to broadcast the identity of the whistler. *The more variable acoustic feature of the same whistles may communicate other information and serve other functions*." Again the italics have been added.

The italicized portions of their original hypotheses are critically important. They were observing dolphin vocalizations with quite primitive equipment that did not record the higher frequency portions of their whistles and they did recognize that the recorded whistles probably conveyed other information besides the identity of the vocalizer.

The whistles appealed to them because of the simplicity and ease of analysis and categorization of the whistles, not their demonstrated relevance in the broader context of language or their demonstrated relationship to a given animal to the exclusion of associated burst-pulsed sounds.

Their table 1 includes a wealth of information, but little with respect to the whistles of individual dolphins. Their subsequent discussion is voluminous and very valuable. However, they note, "We can sort spectrograms of signature whistles reliably, and dolphins are easily able to discriminate signature whistles of different individuals. The stereotypy of signature whistles has not been defined quantitatively, however, nor have indices for comparing the similarity of different whistles been developed. Development of quantitative indices of similarity will be important for studies of mimicry and the effect of early acoustic environment on the ontogeny of signature whistles. this will not be simple . . ."

They did not offer any graphical description of a "signature whistle." They focused on only simple whistles, recognizing they were frequently repeated in loops of 2 to 6 or more. The individual whistle was characterized by its start frequency, end frequency, duration and "total frequency modulation." They did note the frequent occurrence of intervals between the individual spectral forms. Some of these intervals included significant energy suggestive of an unvoiced element within their definition of a signature whistle. On page 205, they discussed the finer variations in their observed signature whistles that included, variations in the first and/or last forms within a series of loops, partial loops, chirps. They even went on to define partial whistles as intermediate between chirps and full whistles.

Caldwell et al. went on, "While there is general agreement concerning the acoustic structure of whistles, there is less agreement concerning the interpretation of variation in whistles." They conclude, they are "produced in contexts indicative of social communication. This interpretation of dolphin whistles is either clearly stated or implied by virtually every one who has worked with the sounds of this species." They also note, "The assumption behind the work on distress whistles was that dolphins share a large repertoire of stereotyped whistle contours, each of which is produced in a particular behavioral context." Dolphin cohorts exhibit a lexicon! They note the lack of enough understanding of these variations they relate to the term stereotypy. Their final sentence was, "Data on signature whistles and whistle mimicry point to a system of communication highly dependent on learning and to a social system based on individual-specific social relationships."

Later investigators, such as the McCowan team, suggested a considerably different structure for a signature whistle and suggested "personal signature" or "individual signature" as more appropriate labels. **Section U.5.6** suggests the label, "*signature whistle sequence*" or more concisely a "*signature sequence*" containing combinations of all of the available symbol types (consonants and vowels, etc.) with potentially multiple copies of each. This definition is compatible with their observation that a specific number of loops of simple whistles do not appear to represent a personal designator.

U.4.7.1.4 Local work of Santos et al. & Hickey

The best early studies concerned with dolphin communications is the work of Santos, Caporin et. al[113].

Their tables represent the start of a symbol analysis (step B) that can be related in a preliminary way with their limited traffic analysis (step A). While they speak of signals in the frequency range of 4 to 16 kHz as a mixture of echolocation and communications (personal signatures is their term), I do not believe these low frequencies have anything to do with echolocation (other than coarse obstacle avoidance like humans use to avoid running into a wall in a darkened room). To me, their symbols are all intra-species communications. The duration of their symbols is also much too long for echolocation unless they are using a Doppler system. The duration of their contours suggest each is a word of multiple symbol length (multiple syllables within one word) or even a sentence. (I only have preliminary information but it looks like essentially all mammalian brains operate on a 30 ms integration interval.)

---

[113]Santos, M. Caporin, G. et. al. (1989) Acoustic behaviour in a local population of bottlenose dolphins, *In* Thomas, J. & Kastelein, R. *eds*. Sensory Abilities of Cetaceans. NY: Plenum Press pp 585-598

One of these contours may say "I am proceeding to the left." The task is to figure out which one it is. This requires the very close delineation of what the dolphin did before or following this message (step C) followed by continuous repetition of the steps in order to build up a library.

The amplitude contours of Santos et al. without an accompanying spectrogram are of little value. Hickey has provided both amplitude profiles as well as spectrograms for the dolphins living around the island of Ireland (Eire)

U.4.7.1.5 Contributions of the Markov team before 1990

Markov and his team presented a series of substantive papers in the 1970's and 1980's. Beginning their discussion in a 1989 paper,, Markov & Ostrovskaya describe the many identified features of the dolphin vocalization system, including the simultaneous use of up to four independent sound generators. They note, “All this indicates that the communicative system of bottlenose dolphins should possess great capabilities, which suggest a great complexity of organization.” They then propose, “One can hardly imagine the existence of a developed communicative system in a species which does not need such a system or can not use it properly.”

Markov & Ostrovskaya introduce the Zipf diagram into their discussion but do not address entropy directly. They do note, “Zipf’s Law in linguistics is known as Pareto’s Law in economics, Lotka’s Law in sciencemetry, and Bradfort’s Law in informatics.”

Consistent with the above complexity, they note, “On the average, one can identify 5-7 blocks in the signal, though their number can reach 12.” Unfortunately they only use drawings to illustrate their blocks. They do indicate their individual blocks consist of whistles with typical durations of 17 to 80 ms. These whistles would correspond to phonemes in the context of the discussions below. The blocks correspond to either syllables or words. The paper contains immensely more pertinent information and proposals concerning dolphin vocalization. **Figure U.4.7-4** describes their basic “whistles” which they note, “Actually a whistle of any type amounts to an entire class of signals, since, while having the same shape of the contour, they can differ in duration, frequency range, register and the rate of frequency modulation.” Like most other dolphin researchers to date, they did not consider vowels consisting of two distinct simultaneous contours.

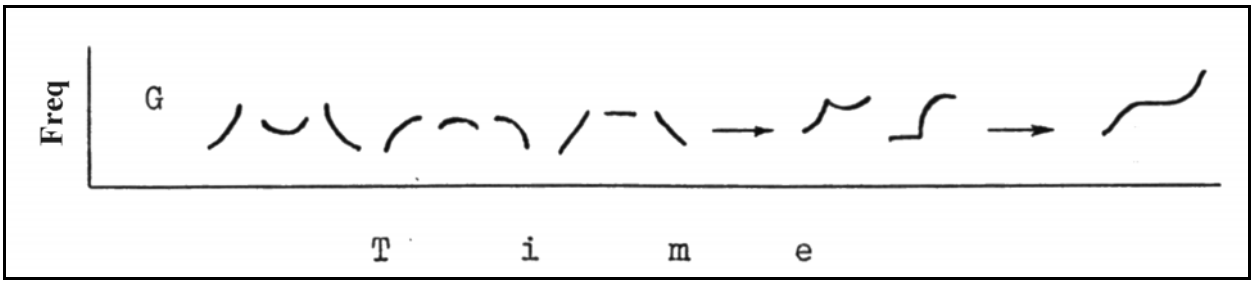

Figure U.4.7-4 A set of simplest structural elements and generation of two- and three-element signals out of them. From Markov & Ostrovskaya, 1989.

Figure U.4.7-5 shows the operation of various combinations of sound generators by a single dolphin. They note the individual tonal generators start at distinctly different times. This work differentiates between the two pulse type sound generators and the two tonal generators (**Section U.2.3.1** and in Appendix L.). Thus, it does not accept the figure as described by Markov & Ostrovskaya as viable. Markov & Ostrovskaya note, “Signals whose structure is formed with the participation of four sound generators are observed rarely in bottlenose dolphins. While the share of signals formed by one generator accounts for about 55%, by two generators - 40%, by three generators - somewhat less than 5%, the share of signals formed by all four generators constitute less than 0.1%.”

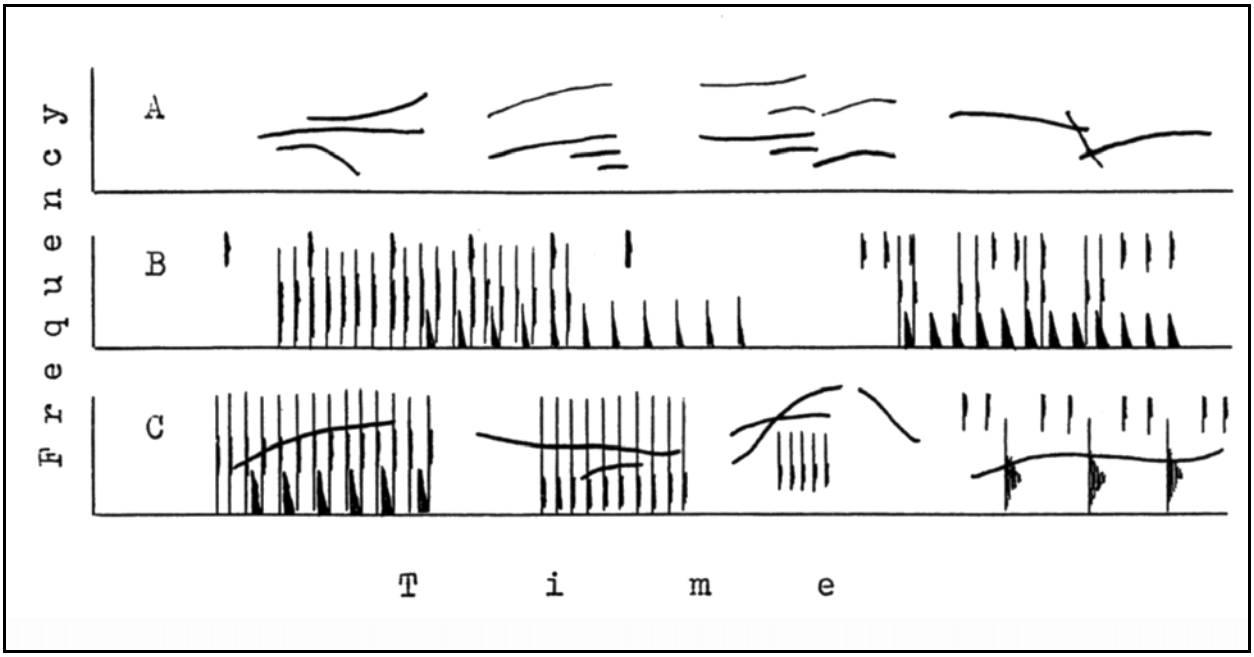

Figure U.4.7-5 Operation of three sound generators. Graphic presentation of the sonograms. A; all generators working in the tonal regime. B; all generaros working in the pulse regime. C; different versions of combined signals. From Markov & Ostrovshaya, 1989.

Markov & Ostrovskaya also show a set of four Zipf diagrams with ranks ranging from about 80 to nearly 1000 based on their interpretation of dolphin messages suggesting zero-order entropies of between H_0 = 7.5 and 10.0. These are the only Zipf diagrams found in the literature that apply specifically to dolphins. Accompanying these diagrams are frequency of occurrence versus duration of the signal lengths diagrams for a variety of situations (based on their criteria). They did not provide any list of phonemes although they indicate that such lists have been prepared by M.A. Ostrovsky in their laboratory. No papers by this author have been found using Google Scholar.

Their conclusions are thought provoking. "bottlenose dolphins have an open-type communicative system. Using a very flexible sound generation system, they can use multi-level combining for constructing a virtually unlimited vocabulary - a set of acoustic signals - and for forming a multitude of organized text messages out of its units. As to degree of complexity, the communicative system of bottlenose dolphins is unique, and nothin of the kind has so far been discovered in other animal species." These comments appear to include humans among the "other animal species." The multi-level combining capability may successfully compete with the recursive feature so loved by linguists as a means of separating human language from animal communications.

U.4.7.2 Initial introduction of information theory into dolphin communications

The 'first principle' developed by Shannon[114] in 1948 was supported by the parallel work of Zipf[115]. Shannon focused initially on the entropy of a single symbol used in a message to determine channel capacity. Later, he focused on the higher order entropies important in cryptanalysis.. Zipf focused on the entropic relationship between the first and second symbols, the first, second and third symbols, etc. They are generally described as the zero order entropy (single symbol space), first order entropy (pair of adjacent symbol spaces), etc. They summarized the value of these entropies as follows:

1. Zero-order entropy measures repertoire diversity.
2. First-order entropy begins to measure simple repertoire complexity (internal organizational structure).
3. Second-order entropy measures the repertoire complexity at the level of two-symbol sequences.
4. Higher-order entropies measure the communication system complexity by examining how signals interact within a repertoire at the three-symbol sequence level, and so forth.

---

[114]Shannon, C. (1948) A mathematical theory of communication *Bell Sys Tech J* vol 27, pp 379-423 & pp 623-656 *Reissued* as a standalone 80 page monogram # B-1598 in 1957 *See also Shannon & Weaver below*.

[115]Zipf, G. (1949) Human Behavior and the Principle of Least Effort. Cambridge: Addison-Wesley

The term entropy in communications is generally a contraction of the term *conditional entropy*. The entropy is conditional upon the character and size of the symbol set being addressed. Only a corpus of infinite extent can define the actual entropy of that symbol set as used with a specific language protocol.

While Shannon worked in a logarithm to the base 2 space, Zipf worked in a more common logarithm to the base 10 space (differing by a scale factor of 2.3). Zipf also developed an additional relationship. He plotted the distribution of the logarithm of the signal rank plotted against the logarithm of actual frequency of occurrence (i.e. percentage repetition of that signal), and asserted an ideal system exhibited a slope of minus one. This slope is independent of the base of the logarithms used. It is not the "slope" highlighted in the McCowan team papers discussed below.

To develop these entropies, they emphasized the need for very large data sets. Just to obtain statistically relevant first-order entropy for a 27 symbol language (English letters plus a space) requires at least 351 two-symbol groups. For a seventh-order entropy, including a majority of English words in common usage, 888,030 multi-symbol groups would be needed. The task appears to be greater in dolphins where McCowan, Hanser & Doyle identified 102 whistle types among a sampling of 600 classified whistles. Considering whistles as vowels, their study did not include any consonants. This situation says the dolphin system studied was an open system (there were probably more symbols than those identified by McCowan et al.). Including the additional symbols associated with the consonants would raise the number of symbols well above 102 and require exponentially more data samples to define the information capability of the language adequately.

U.4.7.2.1 The theory behind and character of Shannon's entropy formulas

Shannon first defined his famous equation for the zero-order entropy with considerable precision and accompanied by three conditions. He noted that his H(x) does not describe a quantity that is a function of x but this notation was meant as a label applicable to a specific formulation for the function H. The current terminology employs the notation H_x

where the zero-order entropy, $H_0 = -K$ times the summation of $p_i \log p_i$ over *all* n symbols in the symbol set.

He detailed the conditions for his definition of H_0 as,

1. $H_0 = 0$ if and only if all the p_i but one are zero, this one having the value of unity.

2. For a given n, H_0 is a maximum and equal to log n when all the p_i are equal (i.e., 1/n)

Since Shannon's time, K has been equal to 1.0 when the natural logarithmic base is used (symbol Ln). Shannon went on to develop higher orders of entropies based on specific symbol adjacencies within a longer sequence of symbols–or message,

3. For two events, x and y, with m possibilities for the first and n possibilities for the second, let $p(i,j)$ be the probability of the *joint* occurrence of *i* for the first and *j* for the second.

The third item defines an n-gram of size 2 and when expressed as a value as H_2. Note that changing the order from *i,j* to *j,i* represents a different joint event and may result in a different value of H_2. The order qu within a message is very common in English, the order uq is distinctly unusual. There is no requirement that the symbol *i*, with a probability p_i determined from a large corpus, occur more than once in a given short sequence of alphanumeric symbols (a message).

n-grams are defined in terms of adjacent sequences of specific symbols in a particular order. There is no requirement that any single symbol, or any adjacent sequence of specific symbols, occur more than once within a short sequence of symbols (a message) defining an n-gram or a specific order of entropy. Recurrence only contributes to the (statistical) quality of the n-gram or entropy value. The symbols used in linguistics for an n-gram can be defined as individual written characters, individual phonemes, syllables, words or in theory sentences or paragraphs. In protein sequencing, the symbols have been defined as base-pairs, amino acids, etc. Google has made extensive use of n-grams based on English words to construct their indices.

Determining the total number, n, of symbols in the vocal symbol set of a given animal communications system is a daunting task and a major requirement in decoding (decrypting) the meaning of the messages used in conspecific communications.

Attempting to decode a message without knowledge of the underlying symbol set, or other ancillary information (syntactic structure, semantics, pragmatics), is not likely to be fruitful. The breaking of the Japanese "Purple" code associated with the battle of Midway during World War II was based on pragmatics. The Japanese repeated a specific message in a coded message and took a specific action after being provided a specific uncoded message.

The label H_1 refers to the entropy associated with a one-gram or unigram. H_1 describes the real entropy of the complete set of symbols found in a very large corpus. Thus, it is based on the same equation as that for H_0 but where the actual values of p_i are used. H_1 represents the actual entropy of a symbol set and is necessarily less than the closely associated zero-order entropy, H_0, or H_1(max).

Other entropies can be defined describing unique symbol arrangements, such as the symbol y following symbol x by a given number of symbol spaces, z, in a symbol sequence (message). However, the result is a very specialized n-gram and very specialized order entropy.

U.4.7.2.2 The character of entropy presentations and the "information figure"

The labels, H_0, H_1, H_2,H_3, . . . have no intrinsic relationship. The subscript numbers are not ordinal, they are arbitrary labels. Shannon did not use the subscript notation in his original derivations. He did note that what is now called H_0 was a maximum theoretical value for H_1. They can be presented next to each other in any order and/or spacing to form a bar-chart. Excel uses the label bar-chart if the bars are horizontal and column-chart if the bars are vertical. Connecting the height of the individual columns by a line results in a line chart in Excel. However, the line so constructed has no underlying meaning. If the heights of the columns are arranged in descending order, the resulting line graph will be monotonic but remains without meaning. Cardinal numbers do not support fractions or decimals, as are only used for counting objects.. Lacking meaning, the slope of the line graph also lacks underlying meaning.

A line graph (chart) arranging the entropy values in monotonically decreasing values with equal spacing based on the subscript labels as cardinal numbers has been called an "information figure." As noted above, the presenter needs to justify the character of such a presentation.

In more recent times, Yaglom & Yaglom have provided extensive texts with the material at the level of Shannon and of Zipf. Their English edition of 1983 is both detailed and comprehensive[116]. Their discussions of entropy appear to be the source of the values used by McCowan and associates in several tables included here. When speaking of the entropy of a communication system, they stress the requirement to know how many discrete symbols are available in the symbol set (including the general need to include a blank space, giving the English alphabet 27 characters including the space). The modern Russian symbol set with a space consists of 33 symbols. Thus based simply on the number of symbols, Russian has a higher zero-order entropy than English. When discussing entropy, it is common to calculate the zero-order entropy as the logarithm of the size of the symbol set under the assumption that all symbols occur with equal probability (sometimes described as weighting). The English alphanumeric symbol set (including 0 to 9) expands to 37 symbols. They also stress the need for immense databases related to specific types of data (military communications versus specialized texts like Gadsby[117]) to obtain real values for the entropy values.

Yaglom & Yaglom provide distinct discussion based on alphanumeric symbol sets and phonetic (non-written) symbol sets. This distinction is critical when discussing dolphin communications. Chapter 4 of Yaglom & Yaglom is dedicated to the problem of the information transmission through communications channels.

U.4.7.2.3 The Shannon followup papers of the 1950's

In the 1948 Shannon paper, he outlined the fundamental concepts and equations that have led to the field of Information Theory. The paper had two distinct portions, the first looking at what came to be known as the entropy of a corpus of messages based on a closed symbol set. The second portion looked at the capacity of a communications channel and its ability to deliver a message of a given entropy and a closed symbol set. The focus at that time period was exclusively on sequential codes because of the state of the art in electronic signal processing. The alternative, block codes, are used widely today in human radio transmission systems based on the current state of the art.

The 1948 paper did not address any analog signal message. It was focused on decrypting teletype messages where the symbol set was dependent on the technical configuration of the specific teletype machine.

A book appeared in 1949 attributed to Shannon and Weaver[118]. It was in fact a reprint of the 1948 paper where the first word of the title was changed from "A" to "The" reflecting its great acceptance by multiple technical

---

[116]Yaglom, A. & Yaglom, I. (1983) Probability and Information. Boston, MA: D. Reidel

[117]Pierce, J. (1961) Symbols, Signals and Noise. NY: Harper Chapter 3

[118]Shannon, C. & Weaver, W. (1949) The Mathematical Theory of Information. Champaign, Il: Univ Illinois Press.

disciplines. The Weaver portion was essentially an extended introduction that synopsized and/or interpreted some of the Shannon material.

Shannon discussed entropy more fully in his first 1951 paper[119]. Shannon also developed his ideas relating to the redundancy of a language in 1951[120]. This second paper addressed many of the tools of cryptanalysis developed during World War II and extended during the post war period. Both of these papers can be very useful after a phonetic symbol set of dolphin sounds is obtained. Shannon addressed the redundancy in a language at both the symbol level and the word level, and noted the results varied significantly depending on the corpus used as the foundation for the study. In general, English text exhibited a redundancy of about 29% based on a digram analysis, and 36% based on trigram analysis at the symbol level. These values employed essentially all of the tables of recurrencies available to the Bell Laboratories team. He also noted the value was 54% based on a similar word analysis with some extrapolations because of incomplete tables at the word level. He was implying the use of block coding to remove these levels of redundancy and gave several examples of such future coding. The paper includes a roundtable discussion of the work by all of the luminaries of cybernetics of that day.

U.4.7.2.4 The character of the first order Zipf Diagram

Investigators have used the Zipf diagram as a rule of thumb in analyzing languages since Zipf introduced the diagram in 1949. It is typically drawn to analyze the statistical properties of a data set drawn on the prescribed axes of a figure. McCowan et al., 1999, provide citations to the progress in the understanding of the diagram. **Figure U.4.7-6** reproduces their figure 1. Frame 1a showing two cases, a plot of 10,000 randomly generated symbols from a set of 102 symbols corresponding to the number of potential symbols found by McCowan et al. to be potentially dolphin vocal communications symbols and an assimilation of data from a collection of human languages. The corpus of 10,000 symbols equaled 100 times the 102 symbol set. The fact that all of the symbols showed a near equal frequency of occurrence suggests each symbol carried the same amount of information but very low signal complexity, a zero-order entropy near zero.

McCowan et al. provide a table of calculated parameters for each frame of their figure 1. They speculate on the unusually large negative value of the slope for the 2-8 month old dolphins but do not consider the standard deviation associated with each data point presented. Elimination of only a few low rank data points would result in the slope calculation proceeding from a low value of -0.82 for a dolphin of less than 1 month of age to a value on the order of -0.95 for an adult dolphin. A value of 0.95 would be similar to the slope of various human languages. A slope with an absolute value exceeding 0.95 in a given language would suggest a very low level of redundancy for that language. Small slope changes near -1.00 may have large consequences. McCowan et al. do not explore the data set giving a slope of -1.07 in any detail. They do discuss the need for an ample sample size to obtain precise values for the point in a Zipf diagram. They note “extreme undersampling has the effect of increasing the slope (the higher-ranking signals are over-represented in a smaller sample), and thus the slope observed for the youngest infants might be expected to become even flatter if the infant data were sampled sufficiently.” Their sample sizes were marginal at best based on both the guide line of Zipf himself and their discussion, 4-5 times suggested by them and a sample size of 10 times the sample types by Hailman & Hailman (page 415) .

---

[119]Shannon, C. (1951a) Prediction and entropy of printed English. *BSTJ* vol 30, pg 50+

[120]Shannon, C. (1951b) The Redundancy of English Murray Hill, NJ: Bell Laboratories

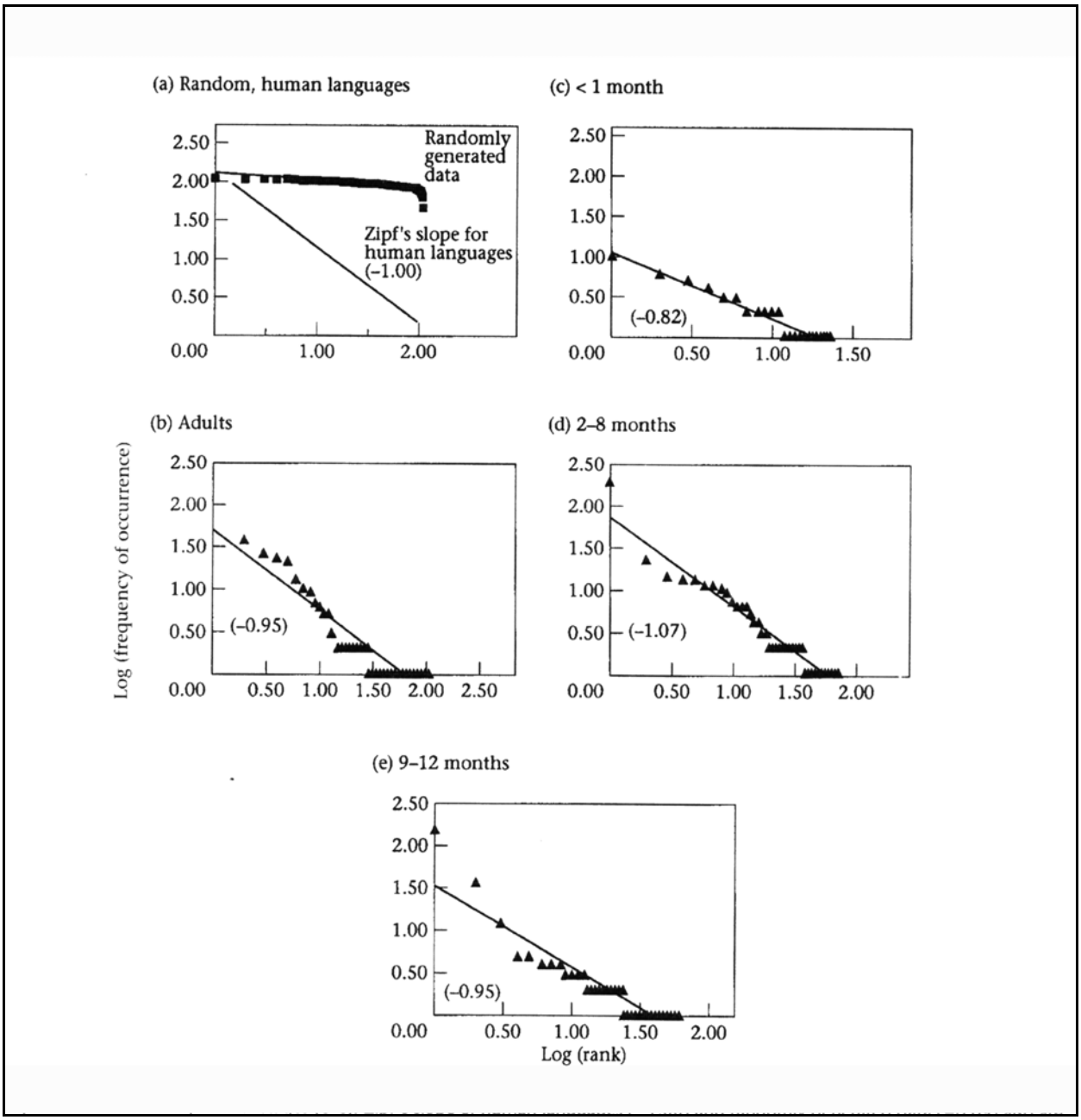

Figure U.4.7-6 Regression lines on a first order Zipf diagram. Frame (a) shows the typical situation obtained for a random number sequence and an assimilation of human languages. A horizontal line and a line with a slope of -1.00 can be considered asymptotes. The horizontal line shows the symbol set contains no information on a symbol to symbol basis. It is proposed the -1.00 slope line is an asymptote with respect to redundance and describing a highly efficient language protocol. A slope exceeding -1.00 is seldom reported and is probably due to limited precision in the measurements. If accurate, such a value would suggest a minimal slope asymptote with respect to needless repetition within the language protocol. Frame (c) shows a lower slope from data on <1 month old dolphins. The slope increases marginally with age until it is shown a -0.95 for adult dolphins. Frame (d) would have a slope of less than -1.00 if the data point at zero rank was eliminated from the regression calculation. Similarly for frame (e). See text. From McCowan et al., 1999.

U.4.7.2.5 The theory behind the first order Zipf Diagram

Nielsen has provided a popular description of Zipf's Law[121] including a plot of the same data on a diagram using linear-linear scales. That plot demonstrates the data can be represented on the vertical logarithmic scale by very negative numbers if the vertical axis is normalized.

The form of the Zipf plot appears to be that of a hyperbola, rank x frequency of occurrence = r x f = constant. The general hyperbola is symmetrical about the zero point (with the same form in quadrants 1 and 3). However, only positive values of "frequency of occurrence" have meaning, thereby limiting this discussion to quadrant 1.

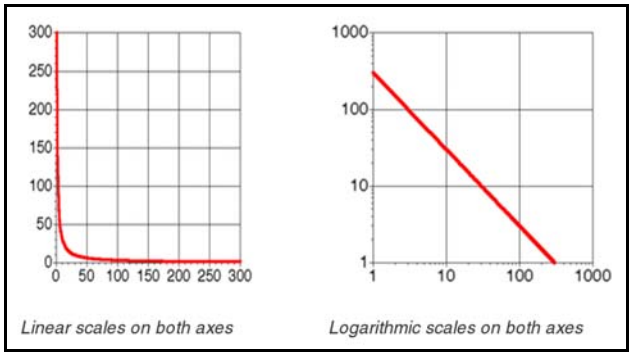

Figure U.4.7-7 The same data frequency of occurrence plotted on both linear and logarithmic coordinates. Horizontal axis, rank order of data set. Vertical axis, frequency of occurrence of the rank. From Nielsen, 1997.

Wyllys has provided a comprehensive discussion of the theoretical basis for Zipf's Law of linguistics and many empirical examples,[122] some beyond linguistics. He develops the more useful function for linguistics of $r^B f = c$ for B <1.00 which he defines as the "generalized Zipf's law."

Wyllys begins, "As one commentator, the statistician Gustav Herdan, has put it: 'Mathematicians believe in [Zipf's law] because they think that linguists have established it to be a linguistic law, and linguists believe in it because they, on their part, think that mathematicians have established it to be a mathematical law.'" He goes on, "Zipf's work shows that the approximation is much better for the middle ranks than for the very lowest and the very highest ranks, and his work with samples of various sizes suggests that the corpus should consist of at least 5000 words in order for the product rf to be reasonably constant, even in the middle ranks." The value of 5000 (200 times the 26 letters of the English alphabet) is much higher than the rules of thumb suggested by McCowan for exploratory investigations. The horizontal strings of data points in McCowan's data are a logical result of using an inadequate corpus of statistical trials. With sufficient trials, only the highest rank should exhibit a log frequency of occurrence = 0.0 (one occurrence). In some strings, the graphic size of the symbols used may obscure the small differences in actual ordinate value.

Wyllys' figure 1 shows results for a 21,354 word corpus. "The regression line, shown as a solid line, has a slope of -0.92; for comparison, figure 1 also shows a dashed line whose slope is -1."

"The study of Zipf's law can be broken into three areas: (1) the initial discovery that the simple equation does approximate the relationship between rank and frequency, (2) investigation of whether a better approximation exists, and (3)attempts to provide a satisfactory rationale for the close relationship of rank and frequency."

Wyllys notes, "Zipf's law only approximates the relationship between rank r and frequency f for any actual corpus. Zipf's work' shows that the approximation is much better for the middle ranks than for the very lowest and the very highest ranks, and his work with samples of various sizes suggests that the corpus should consist of at least 5000 words in the corpus being analyzed in order for the product rf to be reasonably constant, even in the middle ranks."

Based on his research, Wyllys notes, "In general, diagrams of the log(frequency)-versus-log(rank) relationship for natural-language data typically show a downward concavity for the low ranks. The full set of products rf typically shows a fairly consistent slow rise in the values of rf as r increases, rather than any readily identifiable constant value. Thus, the simple (hyperbolic) equation seems to represent actual data less accurately than does the generalized Zipf's law for B<1.00."

Wyllys also took pains to repeat Zipf's comment, "Before returning to linguistic considerations, let me say here for the sake of any mathematician who may plan to formulate the ensuing data more exactly, the ability of the highly intense positive to become the highly intense negative, in my opinion introduces the devil into the formula in the form of [the square root of -1]." The appearance of the square root of –1 introduces the entire mathematical field of imaginary numbers into Zipf type analyses.

Wyllys also notes that Mandelbrot has offered a more comprehensive form of Zipf's Law that appears to apply to linguistics. It includes additional terms to account for the frequent curvature at low ranks encountered in empirical

---

[121]Nielsen, J. (1997) Zipf curves and website popularity, http://www.nngroup.com/articles/zipf-curves-and-website-popularity/

[122]Wyllys, R. (1981)Empirical and theoretical bases of Zipf's Law, https://www.ideals.illinois.edu/bitstream/handle/2142/7182/librarytrendsv30i1g_opt.pdf?sequence=1

data. Such curvature would suggest a linear regression is not appropriate for the limited extent of McCowan's data. Edmundson has offered an even more sophisticated equation that includes both Zipf's and Mandelbrots equations as special cases.

"The essence of Mandelbrot's contribution was his considering communication costs of words in terms of the letters that spell the words and the spaces that separate them. This cost increases with the number of letters in a word and, by extension, in a message. Mandelbrot showed that Zipf's simple law follows as a first approximation from the minimization of communication costs in terms of letters and spaces. Linguistically, this amounts to minimizing costs in terms of phonemes, which is why the phenomenon holds for both written and spoken language."

"Many attempts have been made to provide other rationales for the Zipf phenomenon. Most of them are probabilistic in their approach, i.e., they consist of derivations, from various premises, of the probability that a word will occur with a certain frequency in an arbitrary corpus. . .only the nature of these attempts can be sketched; the principal goal is to emphasize their variety and, hence, the inconclusive current state of explanations of Zipf's law."

Wyllys concludes, "It seems intuitively plausible that some kind of general Poisson process should underlie the pervasiveness of Zipf's law and its siblings. . .On the other hand, it is clear that the process cannot be a pure Poisson process, since the choices of words are not independent, as the Poisson distribution requires. Already in 1955 Simon recognized this in employing a stochastic model 'in which the probability that a particular word will be the next one written depends on what words have been written previously.'"

U.4.7.3 Initial exploitation of information theory in dolphin research during the 1990's

McCowan and her team have approached understanding dolphin language from a different perspective, starting from the simple phonemes and working up toward semantics. McCowan et al., 1999, appearing at the end of a decade of significant groundwork is very well constructed but suffers at the theoretical level. At that level, it is more conceptual than rational. While it provides very provocative data, its comparison with the available theoretical framework is poorly structured[123]. The work relies upon a set of dolphin whistles collected previously[124,125]. They do confirm Zipf's assertion that very large data sets are needed to effectively employ Information Theory techniques to effectively characterize a language (and for a different species to "break the code" leading to an understanding of that language). Unfortunately the 1999 paper exhibits three specific areas of concern,

1. They make an assertion at the beginning of their Theory section; "Serious confusion in the nomenclature surrounding information theory has prompted us first to define our terms clearly." A clause appears to be missing. Any confusion is clearly limited to the application of information theory to the psychophysical world and specifically the signaling of dolphins and other *Cetacea*. Information theory, based on the work of Claude Shannon in the 1940's has been taken as gospel in the electrical engineering community for a very long time and has been used very effective in many non biological applications without requiring auxiliary terminology. It has long been considered flawless in its delineation within the field of Information Theory. On page 410, they credit Shannon as developing the concept of entropy in terms of channel capacity (1948) without recognizing his parallel writings with Weaver[126] on the entropy of symbols (1949) that proved so valuable in cryptanalysis during World War II. See **Section U.4.7.2**.

2.While citing Yaglom & Yaglom several times in the 1999 paper, it is not clear they understood the scope of the discussion in that text. They assembled several tables by accumulating numbers from different pages that are totally unrelated. Specifically, they failed to distinguish between the entropy values associated with written speech and oral speech. Yaglom & Yaglom discuss these modes separately and note written speech employs 26 letters and a space while oral speech employs between 41 and 45 phonemes (and a space). Any dolphin language should necessarily be compared to human oral speech. The zero-order entropy for human speech using equal probability phonemes is between 5.35 and 5.48 rather than the 4.75 value used by McCowan..

---

[123]McCowan, B. Hanser, S. & Doyle, L. (1999) Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires *Anim Behav* vol 57, pp 409-419

[124] McCowan, B. & Reiss, D. (1995a) Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins: a re-evaluation of the signature whistle hypothesis *Ethology* vol 100, pp 194-209

[125]McCowan, B. & Reiss, D. (1995b) Whistle contour development in captive-born infant bottlenose dolphins: role of learing *J Comp Psychol* vol 109, pp 242-260

[126]Shannon, C. & Weaver, W. (1949) The Mathematical Theory of Information. Champaign, Il: Univ Illinois Press.

3. While the effort of McCowan and associates during the 1990's is to be applauded, the results do not appear to warrant similar adulation. They took a very narrow view of dolphin communications and appear to have misapplied Shannon's formula for the zero-order entropy. They addressed only tonal signals band limited to 20kHz (whistles) and limited to intersequence intervals of less than 1600 ms (1999, page 411). They ignored a variety of pulse-sounds (also known as burst-pulse-sounds) such as bangs, barks, blasts, brays, creaks, moans and "buzz effects." Inclusion of these sounds would greatly expand the repertoire of dolphin sounds used in communications and contribute considerably to the structural features that should be included in the their calculations of entropy.

> *A caution is offered here*. The McCowan team discuss whether there was or was not a "signature whistle" identifying an individual during dolphin communications. The potential for an "individual signature" consisting of a whistle plus a consonant was not addressed and cannot be addressed in any study limited to whistles.

They adopt the Zipf-type diagrams mentioned above and introduce additional concepts that may be useful in applying information theory to the psychophysical realm of non-human communications. They characterize whistles more uniquely using a Contour Similarity Technique (CS). Their use of the slope (derivative) of the entropy parameter of Shannon is not clearly defined in the paper but is defined later in Doyle et al., 2008 as the slope of the regression line on the information graph of entropy value versus entropy order (presumably with the information order plotted linearly on the horizontal axis and the H-values plotted linearly on the vertical axis..

McCowan et al. (1999) describe their paper as "illustrating the nature of and predictive application of these higher-order entropies using a 'preliminary' sample of dolphin *whistle vocalizations* in both their Abstract and text. It is difficult to locate the source of their entropy numbers for the dolphin. They quote Shannon's formulas on page 412 but only the final values appear in Table 2 (see the discussion of subsets of entropy in **Section U.4.7.3.3**).

Based on the second concern and their preliminary nature of the 1999 paper, the McCowan et al. paper of 2001(The fallacy of 'signature whistles' in bottlenose dolphins[127]) must be considered preliminary also as it denigrates the concept of signature whistles in two respects; 1) without considering their concatenation (and/or overlay) with a variety of potential consonants to provide more definitive *personal signatures* and 2) their own assertion that the signature whistles they examined may include secondary features that could lead to *individual signatures* (2001, page 1152).

The 1999 paper clearly shows the inadequacy of recording dolphin whistles with equipment limited to 20 kHz. Many of the spectrograms of figure 2 are clearly going out of frame at 19.3 kHz. Figure 2 also displays a variety of brief chirps that may be associated with consonants. Some of these chirps are of significant amplitude suggesting they are not due to extraneous noise. They are also highly repetitive, suggesting their arising from a separate generating mechanism than that generating the tonal waveforms.

The 1999 paper does discuss one set of two-whistle sequences as shown later in this section. If statistically significant, such sequences strongly suggest a complex syntax in a dolphin language protocol.

- - - -

The 1999 paper is careful to define a very limited list of terms. The most important term is information, that they point out can be defined in at least three seemingly conflicting ways. Most commonly, information can be defined

1. In the "contextual" sense, that is, what information is conveyed during a communicative exchange (i.e. meaning).
2. Information can also be defined in the "communicative" sense, that is, how much information can be transferred during a communicative exchange.
3. Finally, information can be defined in the "statistical" sense (as Shannon defined it), as the degree to which data are non compressible, that is how much redundancy remains in the language. The latter term leads to the concept of entropy within a signaling stream. There is a minor semantic problem concerning the term random. McCowan and associates used the expression random to describe the symbology of a signal of maximum content. For real communications signals, the signal is typically pseudo-random. It appears random using conventional statistical tests but hides a structure known only to the communicating parties (the language protocol, or the transmission medium protocol if a translated from a natural language).

Pseudo-random signaling is now widely used in spread-spectrum radio transmission. Without the key, the signal appears completely random. However, with the key, the message becomes perfectly clear text (that may or may not include spaces, pauses, etc). McCowan, Hanser & Doyle defined *information* in the statistical context of Shannon, as the entropy of a message, in the 1999 paper. They defined entropy exactly as Shannon did, "a measure of the information degree of organization," a construct distantly related to the entropy of thermodynamics. Information in

---

[127]McCowan, B. & Reiss, D. (2001) The fallacy of 'signature whistles' in bottlenose dolphins: a comparative perspective of 'signature information' in animal vocalizations Anim Behav vol 62, pp 1151–1162

the communicative sense was defined as the communication capacity of the system, i.e. the complexity of the signaling system. They avoided any discussion of contextual information in their paper.

They only briefly note that information can be present as alphanumeric information employing a conventional alphabet representing written messages, as vocal information that can be expressed using a phonetic alphabet, and as contextual information expressing the meaning of signals in some other form. No third form has been readily defined for any language but defining such meaning based languages are a current and serious area of research in natural language translation. The McCowan team have only addressed information in the statistical sense using a set of waveforms limited to whistles that have been identified by their contour similarity technique in place of a phonetic alphabet. *No consonants have been included in their analysis.*

The McCowan team have been attempting to describe the entropy and other properties of dolphin communications without the availability of a conventional written alphabet or a phonetic alphabet by using their contour similarity technique to isolate approximately 62 whistle waveforms and assign them individual identifiers. As a result, their comparison of these identifiers with the entropy values associated with written languages is awkward. A comparison between the parameters of their identifiers with a set of phonetic identifiers used in natural languages would be more satisfactory. However, such a comparison remains limited by the elimination of consonants in the dolphin material.

They summarize their CS technique on page 410 of their1999 paper, without providing any individual identifiers, and cite McCowan 1995 and McCowan & Reiss 1995a and b.

McCowan 1995 develops her concept of a contour similarity parameter and its scalability. Although she does not use the term, it appears she is treating the first formant of a vowel (at a frequency above 5 kHz) in her contour similarity technique. The fundamental frequency of the dolphin larynx is in the 200 to 600 Hertz range.

McCowan & Reiss, 1995a examined 185 whistles and identified 27 waveforms from a small cohort of dolphins based primarily on a K-means cluster analysis and a briefly defined correlation of frequency measurement.

McCowan & Reiss, 1995b examined 1,281 whistles and identified 128 waveforms using their CS technique which included a calculation of the coefficient of frequency modulation (COFM) for each waveform. They then explored when these waveforms occurred in infant and adult populations.

Figure U.4.7-8 reproduces figure 3 from the McCowan 1995 paper. It shows a sample data set normalized with respect to frequency and duration, and includes the standard deviations about the mean of the resulting data set.

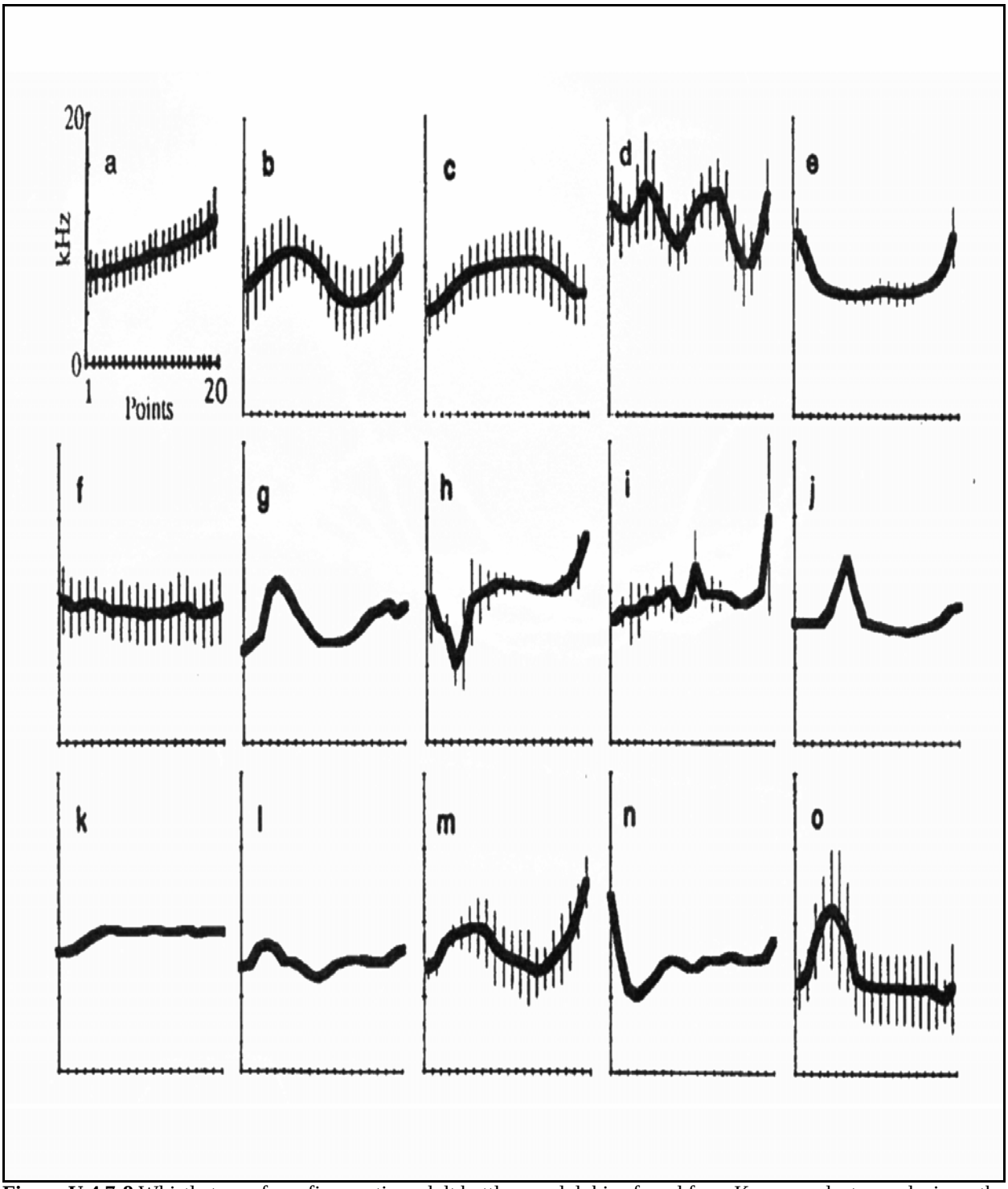

Figure U.4.7-8 Whistle types from five captive adult bottlenose dolphins found from K-means cluster analysis on the factor scores generated from the corrleation coefficients of frequency measurements. Whistle types are represented by the contour of the mean frequencies (line) and SD (bars) of the 20 frequency measurements (points). Whistle types: a, a whistle type, that was shared by all individuals; b-f, whistle types that were shared by two or more individuals; g-o, whistle types that were unique to individuals. From McCowan, 1995.

A caution is offered here. McCowan speaks of dividing the duration of a whistle into equal intervals (most simply referred to as points but more appropriately labeled sampling times). She then refers to the most important frequency component in the spectrogram of that whistle measured at each sampling point as a frequency point. It is clearly the frequency at the sampling time (point) x where x varied between start of the

waveform (labeled time 1) and the finish of the waveform (labeled time 20). In her figure 4a the horizontal scale is labeled "Frequency Measurements." The scale is more appropriately labeled "Relative Time"where the value of 20 corresponds to the end of each waveform regardless of its actual time duration (shown in figure 4b).

McCowan applied a variety of mathematical analysis techniques to her sets of data points, $f(t_n)$. She developed a rectangular correlation matrix that was used in a principal component analysis to reduce the number of collinear variables. The key technique was to employ a K-means cluster analysis to further evaluate the data set. She provided references describing these techniques in detail. She also used a variety of commercially available software routines in these analyses. McCowan only briefly addressed these routines. It is not obvious that all of the appropriate parameters, other than the data sets, were input into these routines.

Jain has long focused on cluster analysis techniques and offered a significant overview of this field in 2009[128]. He notes that thousands of data reduction algorithms have been spawned from the K-means algorithm of 1955. The K-means cluster analysis attempts to distinguish between members of a data set using a mean squared distance criteria. He provides an operational definition of clustering, or cluster analysis as: "Given a representation of n objects, find K groups based on a measure of similarity such that the similarities between objects in the same group are high while the similarities between objects in different groups are low."

Quoting Jain, "The K-means algorithm requires three user-specified parameters; number of clusters K, cluster initialization, and distance metric. The most critical choice is K." "Different initializations can lead to different final clustering because K-means only converges to local minima. One way to overcome the local minima is to run the K-means algorithm, for a given K, with multiple different initial partitions and choose the partition with the smallest squared error."

Jain addresses the "User's dilemma" in some detail. Among other concerns, he notes, "Clustering algorithms tend to find clusters in the data irrespective of whether or not any clusters are present." He illustrates this concern using the K-means logarithm and a data set with no "natural" clustering. The result is a figure with three clearly defined clusters based on a K-means partition with K = 3.

- - - -

McCowan et al. (1999) recognized the limitations of their small data sets to develop "Shannon-type" entropy values and explored a few Markovian sequences to gain further understanding. They continued to define a whistle sequence by an intersequence interval of 1600 ms and a typical interwhistle interval within a sequence as 300 ms. **Figure U.4.7-9** shows a second-order probability tree for a two-whistle sequence. Illustrating a probability tree for a three-whistle sequence is awkward.

Their paper contained much more information that clearly demonstrated (within the confines of the available data set) that the dolphin whistle repertoire was as sophisticated statistically as human language.

Markov & Ostrovskaya have demonstrated the entropy of the wide band sounds (0.5 to 120 kHz) of the bottlenose dolphin of the Black Sea can be described by a Zipf function. They provide a variety of functions obtained under two-way and one-way communications between calm dolphins and Sidorova et al. extended

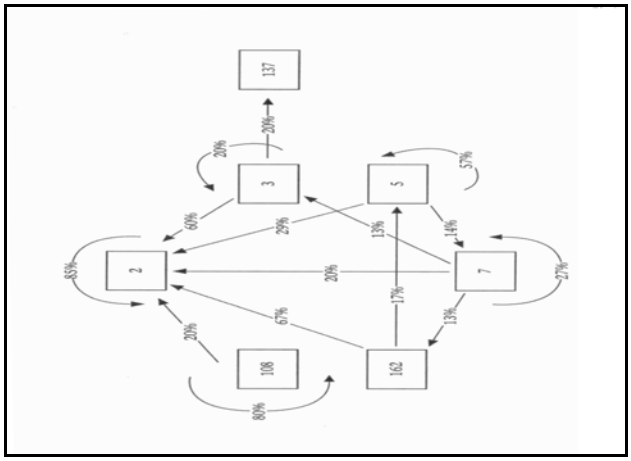

Figure U.4.7-9 One set of two-whistle sequences shown as a probability tree based on a Markovian first-order (i.e. Shannon second-order entropy) analysis. Percentages and direction of arrows shown represent the probability of one whistle type immediately following a second whistle type. A curved arrow indicates the probability that a whistle of one type immediately follows itself. From McCowan, Hanser & Doyle, 1999.

[128]Jain, A. (2010) Data clustering" 50 years beyond K-means *Pattern Recog Letters* vol 31(8), pp 651–666

the studies to highly stressed animals[129]. They all but assert their findings confirm the use of language by dolphins. The suggested language consists of at least 100 structural types or blocks with the potential of well over 10^5 identifiable signal groups. They did not succeed in defining the semantic meaning of any individual words in their messages.

Boisseau has recently reported on bottlenose dolphins living in several fiords along the New Zealand coastline in considerable detail based on both broad and narrow band recordings[130]. Using a variety of correlation techniques, including principal component analysis and spectrographic cross-correlation, he has defined 12 principal sound components that may constitute principal speech components. His figure 1 providing a comparison of simultaneous broad band and narrow band recordings and assignment of potential semantic or phoneme labels is very useful.

Fletcher provided figures (1953, figure 53) not unlike those of McCowan showing individual spectrograms of all of the vowel and consonant sounds in English. His spectrograms typically show third and fourth harmonics in the formants for most vowels.

It is important to note the human auditory modality utilizes separate and distinct neural channels (in **Figure2.4.2-1**of Chapter 2) to analyze the fricative and tonal aspects of speech before combining the analyses and extracting the overall meaning of the material. The tonal channel is described morphologically as the MOC path because of its source at the medial superior olive of stage 2 signal processing. The atonal channel (time and intensity information) is called the LOC path since its source is the lateral superior olive of stage 2.

U.4.7.3.1 Alternate findings of Janik group in the 1990's

Also working during the last decade of the 20th Century, Janik and associates took a noticeably different approach to the study of dolphin communications/language. The 1999 paper compared and then described the pitfalls in categorizing dolphin whistles[131].

> "I compared the classification of bottlenose dolphin, *Tursiops truncatus*, whistles by human observers with the performance of three computer methods: (1) a method developed by McCowan (1995, Ethology, 100, 177–193); (2) a comparison of cross-correlation coefficients using hierarchical cluster analysis; and (3) a comparison of average difference in frequency along two whistle contours also using hierarchical cluster analysis." "The discrepancies between methods show how crucial it is to obtain an external validation of the behaviour classes used in studies of animal behaviour."

His whistle sample consisted of 104 waveforms. "The sample contained five individual-specific signature whistles and several non-signature whistles. Five human observers, without knowledge of the recording context, were more likely than the computer methods to identify signature whistles that were used only while an animal was isolated from the rest of the group."

Opening his introduction, Janik notes, "A crucial step in any study of animal behaviour is division of the observed behaviour into separate categories. If those chosen have any relevance to the animal itself, a selective usage of these patterns according to some external variable should be observable. Examples of such a variable are a particular context or individual. Thus, if a category is used only in one particular context or by only one individual, it confirms the biological significance of this category. This is one of the most basic principles in animal behaviour research."

Janik noted it was "Caldwell et al. (1990) who termed the most *common whistle type* that is produced by an isolated individual its *signature whistle*."

Figure U.4.7-10 reproduces figure 1 in the 1999 paper, originally in Janik & Slater[132]. It combined all of the signature whistles used in their studies. A separate figure 2 identified all of the non-signature whistles encountered in the study. Janik noted, "Bottlenose dolphins often produce multiloop whistles in which separate whistles follow

---

[129]Sidorova, I. Markov, V. & Ostrovskaya, V. (1989) Signalization of the bottlenose dolphin during the adaptation to different stressors *In* Thomas, J. & Kastelein, R. *eds*. Sensory Abilities of Cetaceans. NY: Plenum Press pp 623+

[130]Boisseau, O. (2005) Quantifying the acoustic repertoire of a population: the vocalization of free-ranging bottlenose dolphins in Fiordland, New Zealand *JASA* vol 117(4), pt 1, pp 2318-2329

[131]Janik, V. (1999) Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods *Animal Behav* vol 57, pp 133-143

[132]Janik, V. & Slater, P. (1998) Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls *Animal Behav* vol 56, pp 829–838

each other closely and occur together most of the time. For the analysis here each separate whistle from such multiloop whistles was considered on its own. Each whistle was given an identification number. These numbers are used purely to refer to a particular whistle in the sample and they have no further meaning." *While the recurrent pattern of whistles may be semantically crucial, it was ignored in the Janik studies!* Also of major significance was that, "All 104 line spectrograms were printed on separate sheets and five observers were asked to classify calls independently by their shape." *The whistles were taken out of context before the human observers visually interpreted them.*

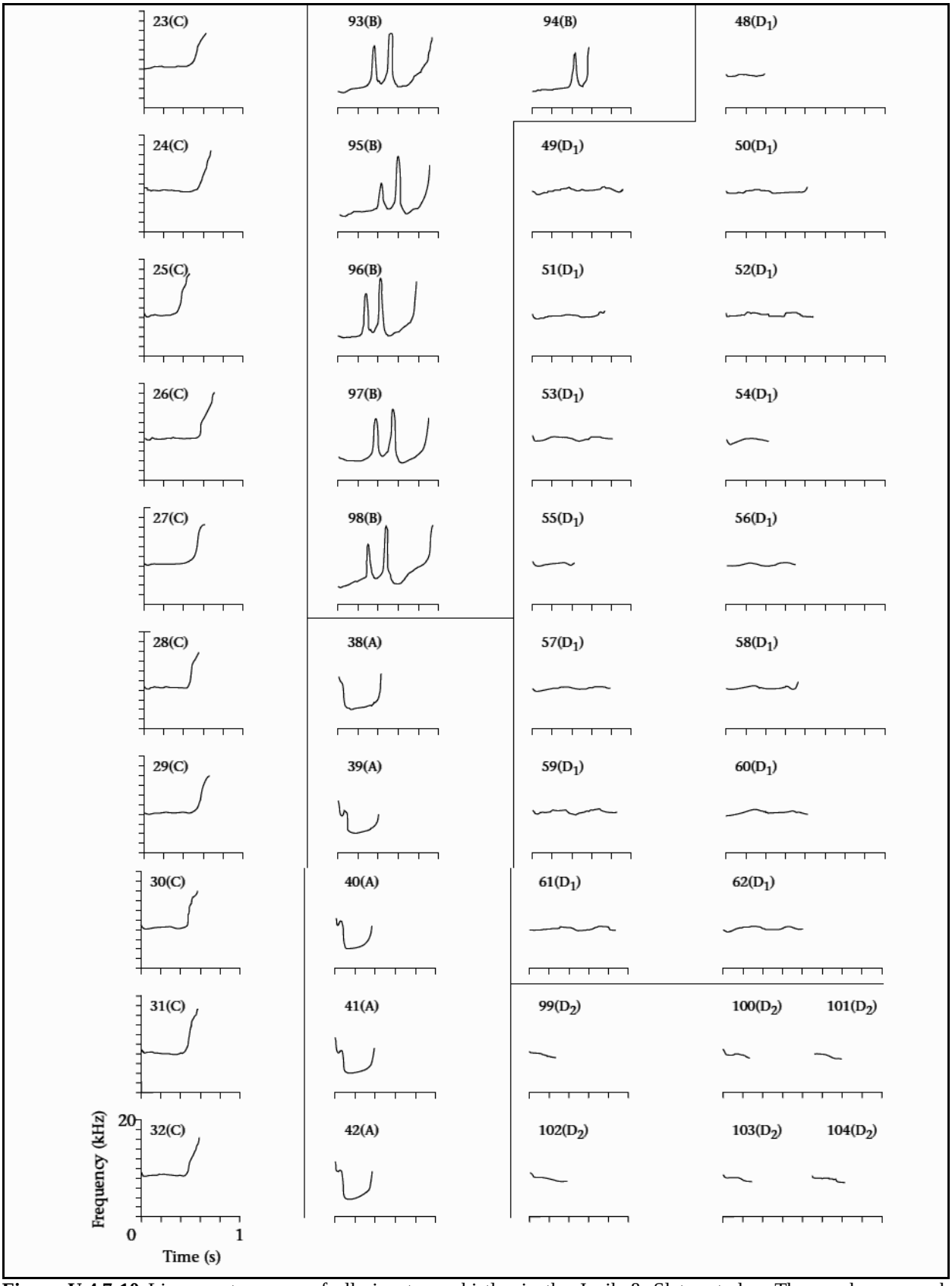

Figure U.4.7-10 Line spectrograms of all signature whistles in the Janik & Slater study. The number on each spectrogram is its identification number followed by a letter indicating to which whistle type it belongs. Note the maximum scale frequency of 20 kHz. From Janik & Slater, 1998.

Janik reviewed the procedures used by McCowan (1995) and noted the need to specify the default values employed in the software routines used to classify these spectrograms. He then describes his various comparison activities and the default parameters used. He then presented his results using only Caldwell's definition of a signature whistle.

Janik concluded that, "The results showed clearly that methods agreed to only a very limited extent. Signature whistles could be identified by human observers but none of the computer methods was capable of identifying them reliably. It is important to note that only after the whistle types had been defined by humans was it found that these whistle types were used almost exclusively by one animal and only if it was isolated from its group."After presenting a large number of dendrograms, Janik continued, "It is still possible that the computer methods, while failing to identify signature whistles reliably, could have discovered significant classes that were missed by the humans. It could be that dolphins use very different criteria for the classification of signature and nonsignature whistles."

Janik closed in 1999 with a very disconcerting position concerning the use of computerized classification techniques without excruciatingly careful design of the underlying protocols. In 2000, Janik reported from a bit more positive position[133] with an Abstract that said, "Dolphin communication is suspected to be complex, on the basis of their call repertoires, cognitive abilities, and ability to modify signals through vocal learning. Because of the difficulties involved in observing and recording individual cetaceans, very little is known about how they use their calls. This report shows that wild, unrestrained bottlenose dolphins use their learned whistles in matching interactions, in which an individual responds to a whistle of a conspecific by emitting the same whistle type. Vocal matching occurred over distances of up to 580 meters and is indicative of animals addressing each other individually." Thus, his studies surfaced many problems in analyzing *signature whistles* but confirmed the likelihood, often expressed, that dolphins were conversing in meaningful ways that imply a language protocol. His figure 2 provides a relative frequency of dolphin interactions versus distance in meters for both matching and nonmatching whistles. **Figure U.4.7-11** reproduces his figure 3 showing obvious cooperation among three dolphins repeating (or mimicing?) the same signal. They appear to be coordinating their positions using this signal. The meaning of the more complex trailing pulse-sound, with elements centered around both 6.5 and 13 kHz, was not discussed.

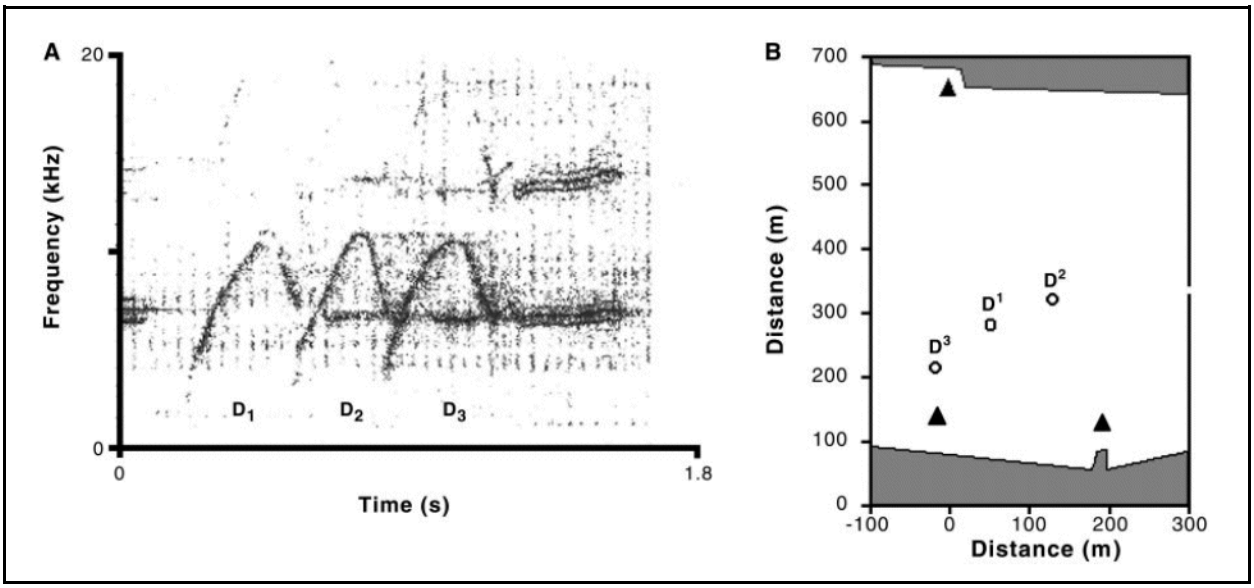

Figure U.4.7-11 A matching whistle interaction that involved three individuals. (A) Spectrogram of the produced whistles. (B) Plot of the array geometry with the locations of each of the dolphins that produced whistles D1, D2, and D3 in (A). Gray areas at the top and the bottom of the plot represent the shoreline. Circles, animals; triangles, hydrophones (25). From Janik, 2000.

U.4.7.3.2 Contributions of Savigh et al. and of Smith

[133]Janik, V. (2000) Whistle Matching in Wild Bottlenose Dolphins (*Tursiops truncatus*) *Science* vol 289, pp1355-1357

Sayigh et al. provided a large study on dolphin signature whistles in 2007[134]. This paper took strong issue with the theme of the McCowan & Reiss paper of 2001 labeling dolphin signature whistles a fallacy promoted by a wide variety of researchers. It offered a rare event in science; it actually replicated (correctly) the experiments of McCowan (1995) to demonstrate the weakness of her PCA, clustering, CS technique protocol. They described the weaknesses in much the same form as in, and in complement to, the above paragraphs developed independently of their paper.. They also repeated a visual grouping of a large number of unfiltered dolphin signatures using totally naive judges with statistically superior results to the McCowan results. The Sayigh recordings were made with equipment capable of reaching 48 kHz, thus capturing the complete scope of the proposed signatures (although not any laryngeal frequency due to the scale of the recordings). The time base was not modified from real clock time. The waveforms were cropped closely and do not show any prefix (adjective) phonemes at the start of their waveforms. They found the naive judges far outperformed 18.9 ± 1.6 out of a possible 20, $P < 0.0001$) at sorting the waveforms belonging to a specific dolphin and they detailed where their simple approach outperformed the McCowan et al. approach. "The only types of whistles that the McCowan method could potentially classify well are those with very simple contours (such as upsweeps)."

Oxnard has discussed the variety of solutions obtainable using PCA and cladogram techniques depending on the underlying parameters employed[135].

Sayigh et al. provided five examples of the signature whistles of each of 20 bottlenose dolphins in their paper. Because of space limitations, the scale of the individual spectrograms is not adequate to evaluate them in detail but variations within a set are recognizable. The sets vary significantly from those exhibiting only a single whistle frequency contour, to those formed by a harmonic group of whistle frequency contours to several signatures recognizable as voiced waveforms with sidebands accompanying the major contour and its harmonics (subject T in particular). In this work, the voiced waveforms are fundamentally different from whistles (see **Section U.2.3.4**).

Smith provided a paper in 2014 looking at estimating the size of N-gram structures in various species[136]. This paper compares multiple values for entropy related quantities published by multiple authors. His figure 3 plots the information graphs using the values from these authors and Table 1 tabulates the values, **Figure U.4.7-12**. The tabulation shows a remarkably low value of M for the baleen whale compared to its fellow *Cetacean*, the dolphin *Tursiops truncatus*. The M of the dolphin appears remarkably low relative to the lowly Robin that is not known for its large repertoire of sounds. These variants are probably due to separate investigators making different estimates. As noted earlier, the value of M =27 for the dolphin was arrived at by an unorthodox path. Smith notes this fact (page 537) and offers an alternate maximum of M = 112 and an alternate minimum of M = 111 using accepted formulas of entropy bias estimating. These values would give a theoretical $H_0 = 6.8$ for the dolphin, rather than $H_0 = 4.75$. A value of 6.8 appears more appropriate in comparison to the thrush, robin and gregarious starlings. His 2014 paper relied upon his earlier detailed paper of 2012 focused on human languages[137]. The 2012 paper included a Venn diagram describing conceptually the difference between entropy, mutual information and conditional entropy.

It is likely the Smith 2014 paper needs to be revised significantly based on the need to identify the McCowan descriptions of their symbol sets (references 27 and 33) as subsets earlier in this section.

---

[134]Sayigh, L. Esch, H. Wells, R. & Janik, V. (2007) Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus, Anim Behav* vol 74, pp 1631-1642

[135]Oxnard, C. (2009) Form and Pattern in Human Evolution. Chicago, Il: Univ. Chicago Press

[136]Smith, R. (2014) Complexity in Animal Communication: Estimating the Size of N-Gram Structures *Entropy* vol 16, pp 526-542

[137]Smith, R. (2012) Distinct word length frequencies: distributions and symbol entropies *Glottometrics* vol 23, pp 7-22

	Species Name	Reference	M	S	S_2	S_3	H_1	H_2	H_3
Dolphin	*Tursiops truncatus*	[27]	27	493	346	346	1.92	1.15	0.56
Humpback whale	*Megaptera novaeangliae*	[33]	6	202	195	N/A	2.15	2	N/A
Skylark	*Alauda arvensis L.*	[26]	170	10,000	10,000	10,000	7.05	1	0.29
Starling	*Sturnus vulgaris*	[24]	105	4,811	4,691	4,691	6.03	1.47	0.81
Thrush	*Hylocichla mustelina*	[13]	35	777	777	777	4.64	3.33	1.09
Robin	*Turdus migratorius*	[13]	44	2,700	2,700	2,700	4.03	2.74	1.95

Figure U.4.7-12 Collected information theory data for various animal species. References relate to original paper. Note the small value of M for the whale and the low value of M for the dolphin compared to the thrush and robin. S; sample size of symbols analyzed. S_2; the estimate (where available) of the number of 2-grams measured. S_3; the estimate (where available) of the number of 3-grams measured. Note the round numbers drawn from reference 13 and 26. See text. From Smith, 2014.

Based on a myriad of human languages (primarily European) and their ability to express endless ideas, it appears likely that any dolphin language would include on the order of 50 phonetic symbols, instead of 102 in its symbol set.

Both the normalized duration of the McCowan frequency contours (200 to 1000 ms) and the longer duration frequency contours of Janik (one second) are compatible with their subdivision into multiple phonemes that may be incorporated into a single or even multiple morphemes.

U.4.7.3.3 Analysis of the McCowan and Janik approaches

The efforts of McCowan and associates has surfaced many major features of dolphin communications that points strongly towards a recognizable language protocol for "Dolphinese." However, as shown in the previous paragraph, the McCowan approach can only be considered an interesting beginning in the analyses of dolphin communications. The omission of consonants, harmonics and the interval between consonants and vowels are serious limitations. Removing these features from a natural language like English would destroy the language as it is currently known and practiced.

A major finidng was that if Dolphinese exists, it is much more tonal in character than English and its language protocol should be compared to Chinese rather than to either English or Russian. Conversely, the lack of attention to consonants within dolphin communications suggests the data set is very incomplete. The prominence of voids between vowel sounds in dophin communications as shown in the earlier figure from Herzing (**Section U.1.2.3**), suggests these time intervals are important in any potential dolphin language protocol. These voids are of similar duration to those used in English.

Whereas certain comparisons with the log frequency of occurrence versus log rank of Zipf diagrams (1999, page 413) for dolphins versus human speech may be observed, the comparison is between apples and oranges until the consonant and other phonetic elements of dolphin communications are folded into the calculations.

She has introduced a scaling protocol and applied it to the whistles of a variety of dolphins and dolphin cohorts. The goal of her quantitative contour similarity (CS) technique was to normalize the whistles with respect to frequency, duration and the variances about the mean for these properties. Crucially, "most techniques do not control for the frequency/time expansions and compressions found in the whistle contours of some species, such as the bottlenose dolphin." This protocol may be appropriate for organizing the whistles of a broad range of dolphins regardless of age, gender and even the speed of communicating. However, in many cases, it folds many independent whistle forms (phonetic symbols?) of short duration into a single form of long duration, thereby reducing the phonetic symbol set artificially. Expansion/compression of the signal duration relates directly to the tempo or rhythm of speech, a significant variable in its prosody.

Expansion/compression clearly has a role to play in human speech when restricted to a factor of less than two. Compare the tempo of speech in Spanish versus English. However, when used at higher factors, it falls to the investigator to demonstrate that the compression has not forced multiple phonemes into a presumed single phoneme interval. Figure 4b in the McCowan, 1995, paper strongly suggests multiple phonemes on the order of 300 ms in a 1000 ms sound sequence may have been compressed into their standardized interval in their figure 4a. They describe their standardized interval as "frequency measurements (controlled for by whistle duration" rather than as relative time interval. As a result the expansion/compression process may obfuscate the separation of a call sequence into its phonetic elements. Compare the waveforms of figure 3 in the McCowan 1995 paper to the more complete set in McCowan & Reiss 1995b paper. Many of the potential

phonemes in the 1995b paper have been lost in the clustering process of the 1995 paper.

As noted in her figure 1, the technique ignores any consonants associated with a whistle and any harmonic content associated with a primary waveform (including the burst pulse sound forms of vowels). Using only 103 whistles from a small group, McCowan was able to organize those calls into a group of only 15 whistles that were distinctive as shown in her figure 3 and reproduced above.

These waveforms can be described in a series reminiscent of the original elements of a Fourier Series.

Fundamental waveform– Her waveform (f) essentially describes a constant frequency tone that is scalable among the cohort of dolphins.

Slow ramp waveform– Her waveform (a) clearly represents the same constant frequency tone modulated slowly by a ramp function. In this set the modulation was always positive (increasing the base frequency with time over the duration of the whistle.

Note carefully the base frequency was not shown to be the fundamental frequency generated by the larynx. It was the predominant frequency observed with a microphone external to the animal using a specific frequency analyzer bandwidth.

Half-sine waveform– Her waveforms (c) and (e) represent waveform (f) modulated by a half-cycle sinusoidal wave extending over the duration of the whistle. The modulation can be positive in waveform (c) and negative in waveform (e).

Full sine waveform– Waveform (b) represent waveform (f) modulated by a full sinusoidal wave extending over the duration of the whistle. Only a positive going sinusoidal waveform was captured in this data set.

Waveforms modulated at higher frequencies– The set is incomplete in this area but waveform (d) can be described as the waveform (f) modulated by a negative going 2 and ½ cycle sinusoidal waveform.

Waveforms modulated by damped waveforms– Waveform (n) can be described in a variety of ways, such as a fundamental waveform (f) modulated by an underdamped negative going step waveform.

Waveforms modulated by combined elements– Waveform (m) is an example of waveform (f) modulated by a combined ramp and a full cycle sinusoid over the duration of the signal. Alternately, waveform (m) can be considered an example of the slow ramp waveform (a) modulated by a full cycle sinusoid alone.

If these waveforms can be shown to apply to other cohorts, within the standard deviation demonstrated by McCowan, the question becomes what do these waveforms mean? Are they totally musical in character or do they represent unique meaning shared among dolphins?

McCowan summarized the results of her extensive tests to determine the validity of these waveforms on pages184-186 of the 1995 paper. She notes, "Whistle type (a) represents a predominant whistle type that was shared by all individuals. Whistle types b-f were shared by two or more individuals within the social group. Whistle types g-o represent the whistle types unique to individuals. The number of unique types (n = 9) accounted for 60% of the total number of whistle types but the percentage of total whistles within a unique whistle type ranged from 18.2% to 4.5%."

The subsequent 1995b paper of McCowan & Reiss extended the McCowan investigation to a larger group of dolphins, eight infants and ten adults from three different social groups. A total of 1281 whistles were quantitatively analyzed and assigned to 128 whistle types using the CS technique, *while the source of the whistle was observed and recorded on videotape*. Thus the individual whistle could be correlated with the activity of the dolphin . The infant population was observed during three separate periods; ages 1-4 months in Development Period (DP) 1, ages 5-8 months during DP 2 and ages 9-12 months during DP 3. The result was a matrix of 62 types of infant whistles during DP 1 (figure 1A) and 128 types of infant whistles during DP 2 & 3 (figure 1B). The patterns in their figure 1A and figure 1B necessarily include more complex waveforms. However, the waveforms can continue to be set in the framework presented above albeit with different identifying labels. In comparing the two data sets, they highlight the waveforms shared between the DP 1 infant and the infants of DP 2 & 3. Two whistle types (1 & 2 of each set) were addressed separately. They were predominant in the infant repertoire to the extent of 66% during the first year of their development. The paper included extensive mathematical analysis designed to interpret the waveforms and indicate the strength of the underlying mathematical techniques (based on the same methodology as the McCowan, 1995 paper).

"In general, whistle type 1, produced exclusively by infants, was produced in contexts of infant separation (M = 0.58, SD = 0.42) during Developmental Periods 1 and 2 and during more variable contexts during Developmental Period 3." Thus a positive going slow ramp waveform can be associated with distress in the infant probably

analogous to Ma in Mama and associated in this case with the meaning of "Help, Mother, where are you, I am lost." They note the same waveform is used among adults in social situations. It may be associated with: <salutation>, where are you? Observation of the videotapes would be needed to confirm this possibility.

The 1999 paper by McCowan, Hanser & Doyle explored dolphin communications from an information theory perspective. However, it presented a very confused picture of that communications in comparison to human communications (**Section U.4.7.3**). This resulted primarily from the methodology used to arrive at an estimate of the symbol set size, M, on page 413. They examined 600 "classified" whistles assembled from two groups of adult dolphins to define N= 102 whistle types. They then employed 147 whistle sequences of two to nine whistles in length from among 493 of the classified whistles from only the adult population to define 27 whistle types that appeared at least twice in a given sequence (Smith, 2014, page 537). A whistle sequence was defined by an inter sequence interval of 1600 ms with typical inter whistle intervals within a sequence of 300 ms.

> ***The calculation of the zero-order entropy based on a requirement that an individual symbol must occur twice within a whistle sequence is not an accepted procedure. It leads to an artificially small symbol set and results in an underestimate of the zero-order entropy. The zero-order entropy is defined with respect to the zero-order n-gram, i.e., no consideration relative to the repeat appearance of any character in a sequence of two or more symbols.***
>
> ***The restriction of their dolphin sounds to only whistles processed through their contour similarity clustering procedure (to the exclusion of barks, brays etc. used in dolphin communications) also leads to an underestimate of the size of the symbol set used by dolphins.***

The intervals of 300 ms and 1600 ms appear too long based on the conclusions of Markov & Ostrovskaya. They consider the duration of a single identifiable whistle feature (a phoneme) to be between 18 and 80 ms in Section **U.4.7.1.4** rather than 300 ms. Such a difference in interpretation is major.

The path followed by McCowan et al. in 1999 and the subsequent paper by Doyle et al. in 2008 highlights their misunderstanding or mislabeling the concept of zero-order entropy. In the 1999 paper, they define the zero-order entropy using an additional limitation when exploring whistles, that the particular whistle must appear more than once within a symbol sequence of more than two symbols. They did not use the term n-grams but are describing n-grams for n > 2. For this paper, they need to make clear, they have defined (labeled) a unique entropy specific to their work (see **Section U.4.7.2.2**). The general n-gram is discussed in **Section U.2.4.1**.

> The unique n-gram used to determine the number of symbols in the whistle set, labeled here as W(O) with M members and their interpretation of the zero-order entropy, H_0, was defined in the 1999 McCowan et al. paper and in the 2008 Doyle et al. paper in a very unusual manner. They defined their whistle symbol set,W(O), based on an n-gram of the form, W_n(x,y,z...x) where n is the order of the n-gram and the term ...x indicates the symbol x had to reappear later in the symbol sequence to be considered an element of the n-gram. ***While this formulation significantly reduced their symbol set from either M=102 or 128 to 27, it cannot be expected to represent the complete whistle set W(O) and/or be used to calculate the zero-order entropy, H_0, for the total whistle symbol set or the larger complete symbol set of the dolphin, S(R)***.

In the 2008 paper exploring feeding patterns of whales, they define the zero-order entropy based on a very small symbol set (M = 6) related to feeding in shallow water and not including the rest of their whistle set employed when communicating in deeper waters.

At the current time, the linguistic symbol set of dolphin communications is undefined and is described as an "open symbol set" in information theory. *The descriptions of both McCowan et al. and Doyle et al. suggest they are using subsets of the complete symbol set and unusual n-grams of the form,* W_n(x,y,z...x). *They need to use an additional indicator or different label to indicate they are calculating a unique entropy or a restricted value of the zero-order entropy, H_0.*.

The 1999 paper presented a Table 2 of entropies comparing various natural languages and a proposed dolphin language that presents an inconsistent perspective. **Figure U.4.7-13** reproduces a significantly modified version of that table. Quoting Yablom & Yablom (page 219), "The list of phonemes for a given language is obviously not identical with the list of alphabet letters. The total number of phonemes considerably exceeds the number of letters, since one and the same letter can be sounded differently in different cases." The table has been modified to highlight these differences.

- The zero-order entropies shown are calculated using what is known as the zero-order letter approximation *using letters of equal probability of occurrence*. This is a theoretical limiting condition not found in practice.
- Specialized communications (such as control tower to pilot talk) invariably generate different entropies than shown.
- Throughout the series of McCowan papers, the number of whistles was given as 102, or on occasion, 128. No justification was given for using a zero-order entropy of 4.75 (corresponding to 26 symbols plus a space) or for the

values shown for the higher orders. The zero-order phonetic approximation for whistles of equal probability of occurrence would be 6.68 for 102 symbols.

• The phonetic symbol set for any dolphin language must be considered open at this time with no known number of total symbols. Therefore a zero-order phonetic approximation cannot be calculated at this time regardless of the probabilities of occurrence of the symbols.

• The origin of the slope and R^2 values given for the dolphin whistles also come into question.

• A line has been added to reflect the entropy of English phonemes.

• The entropies shown for music are very unrealistic. They assume a single octave musical scale of 8 semitone notes. Yablom & Yablom noted other investigators have used a 1 ½ octave set of twelve notes and a piano has 88 notes.

• The calculation of entropy based on musical notes instead of chords highlights the problem associated with the CS technique of McCowan and associates who do not accept the validity of whistles based on multiple individual tones.

Their observation that;

> "Our sample of dolphin whistles are presently undersampled for entropic orders of two or more. Undersampling has the effect of lowering the higher-order entropies (less statistical informational structure), and thereby indicating more communication complexity than could actually be present."

is correct but understates the problem. Their sample of non-whistle symbols is nonexistent and their calculation of the zero order entropy for an open symbol set is spurious.

The key point with regard to Table 2 is that the number of phonetic symbols in the symbol set used by the dolphin is totally undefined at this time. The set remains "open" at this time.

Signal system	Slope	R^2	Entropy: Zero order	Entropy: First order	Entropy: Second order	Entropy: Third order
New Russian, 32 letters	−0.500	0.93	5.00	4.35	3.52	3.01
English, 26 letters + space	−0.566	0.96	4.75	4.03	3.32	3.10
Arabic letters	−0.797	0.96	5.00	4.21	3.77	2.49
Music, only 8 semitone notes	−0.680	0.99	3.00	2.73	2.00	NA
Dolphin whistles, **102**	−1.334 (?)	0.86 (?)	~~4.75~~ **6.68**	~~1.92~~	~~1.15~~	~~0.56~~
Russian phonemes, 42 est.	−1.519	0.89	5.38	4.77	3.62	0.70
English phonemes, 41 + space			**5.35**			

NA: not available.

Figure U.4.7-13 Tabulation of parameters used for various dolphin whistles, human languages & music. The values for human speech are provisional at this time, and await standardization by the linguistic community. See text. Significantly modified from Table 2 in McCowan et al., 1999.

The slope used in this figure and in the McCowan et al. 1999 paper is different from that associated with the Zipf diagram. It is defined by Doyle et al[138]., "The entropic slope is a regression of the entropies against their order." As noted in **Section U.4.7.2**, the entropies are frequently shown using a vertical bar graph. The peaks of these bars do not belong to a common data set and . Smith noted several caveats regarding the information graph in 2014 (page 528). The order number in an entropy expression is a label and not a cardinal number. H_0 actually represents H_1(max) where all symbols have an equal frequency of occurrence, and is not a distinct entropy. Connecting the peaks of these bars must be done with care. The resulting line is not a continuum and a regression line associated with the line, or the small number of original bar heights, lacks precision.

Smith concludes on his page 531, "Using these parameters, it is an intriguing question if we can estimate the

[138]Doyle, L. McCowan, B. Hanser, S. Chyba, C. Bucci, T. & Blue, J. (2008) Applicability of information theory to the quantification of responses to anthropogenic noise by southeast Alaskan humpback whales. *Entropy* vol 10, pp 33–46

size of the repertoire of multiple symbols or sounds in non-human systems of communications."

U.4.7.3.4 Parameters related to humpback whale & bird communications

The number of phonemes in the humpback whale, *Megaptera novaeangliae*, repertoire (M = 6) in the above figure is also highly suspect based on the great amount of observational data on that animal. Reference 33 is from Doyle et al[139]. The low value for M may be caused by the same problem found in the McCowan et al. (1999) paper. Alternately, Doyle et al. associated their M = 6 with only the zero-order entropy among the "feeding calls" of humpback whales within Glacier Bay under high and low background noise levels.

The symbol set used during feeding is clearly a subset of the full repertoire of the humpbacks and cannot be used to calculate the zero-order entropy of the vocalization used by the species.

Humpback songs are known to vary between cohorts, between habitats and over long periods of time. Even a cursory examination of the spectrograms of their songs quickly identify more than six phonemes[140]. Figures 1 and 2 from Winn et al. (1981) compare these songs and illustrate their spectrographic variation. They claim to have identified three dialects among the global population. Their spectrograms contain at least two dozen phonemes. The even earlier recordings shown in Payne and McVay exhibit even more "unit sounds" within phrases, within themes, within sound sessions[141]. Payne & McVay began the arduous task of labeling elements within humpback songs beginning at the theme level. Their fig. 2 identifies more than four phonemes in just the theme of the whales near Vava'u Tonga. Those phonemes differ from those of the same species near New Caledonia, etc. They identified 17 unit level phonemes in Fig.11 using an ART 2-A network as the filter. "The F2 category layer had 20 elements (M = 20), giving the network the potential for forming up to 20 soung unit categories."

The use of the designation F1 and F2 in connection with the ART 2 filter is unrelated to their use in formant labeling. The ART 2 self-organizing neural network is described in greater detail in both Carpenter et al[142]. and in Helweg et al[143]. Helweg et al. introduce the use of state diagrams into the interpretation of whale songs and sequences.

Noad et al. have described the change in songs of a cohort of humpback whales over a period of three years, suggesting their songs are not genetically based but learned[144].

Winn & Winn have described the humpbacks songs[145], "The song of Megaptera novaeangliae evolves through a series of 6 themes: "moans and cries"; to "yups or ups and snores"; to "whos or wos and yups"; to "ees and oos"; to "cries and groans"; and finally to varied "snores and cries". Snores, cries and other sounds can be found in different themes from year to year; yet, invariably one finds a set pattern of changing themes, in a fixed order."

Nowicki & Podos have provided time lines for bird songs that can be quite useful[146]. **Figure U.4.7-14** shows their

---

[139]Doyle, L. McCowan, B. Hanser, S. Chyba, C. Bucci, T. & Blue, J. (2008) Applicability of information theory to the quantification of responses to anthropogenic noise by southeast Alaskan humpback whales. *Entropy* vol 10, pp 33–46

[140]Winn, H. Thompson, T. Cummings, W. et al. (1981) Song of the Humpback Whale - Population Comparisons *Behav Ecol Sociobiol* vol 8, pp 41-46

[141]Payne, R. & McVay, S. (1971) Songs of Humpback Whales *Science* vol 173, pp 587-597

[142]Carpenter, G. Grossberg, S. & Rosen, D. (1991). ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition *Neural Networks* vol 4, pp 493-504

[143]Helweg, D. Cato, D. & Jenkins, P. Garrigue, C. & McCauley, R. (1998) Geographic variation in south pacific humpback whale songs *Behaviour* vol 135(1), pp 1-27

[144]Noad, M. Cato, D. Bryden, M. Jenner, M-N. & Jenner, K. (2000) Cultural revolution in whale songs *Nature* vol 408, pp 537

[145]Winn, H. & Winn, L. (1978) The Song of the Humpback Whale Megaptera novaeangliae in the West Indies *Marine Biology* vol 47, pp 97-114

[146]Nowicki, S. & Podos, J. (1993) Complexity, coupling, and contingency in the production of birdsong *In* Bateson et al. *eds*. Perspectives in Ethology, Vol 10. NY: Plenum Press pp 159-186

depiction of the phonetic elements in bird songs. It provides conventional definitions of trills and other elements of bird songs. Note the syllabic organization of the songs, typically containing three phonemes in each syllable. Syllables are frequently repeated multiple times but with stand alone syllables between the repititions of different syllables.

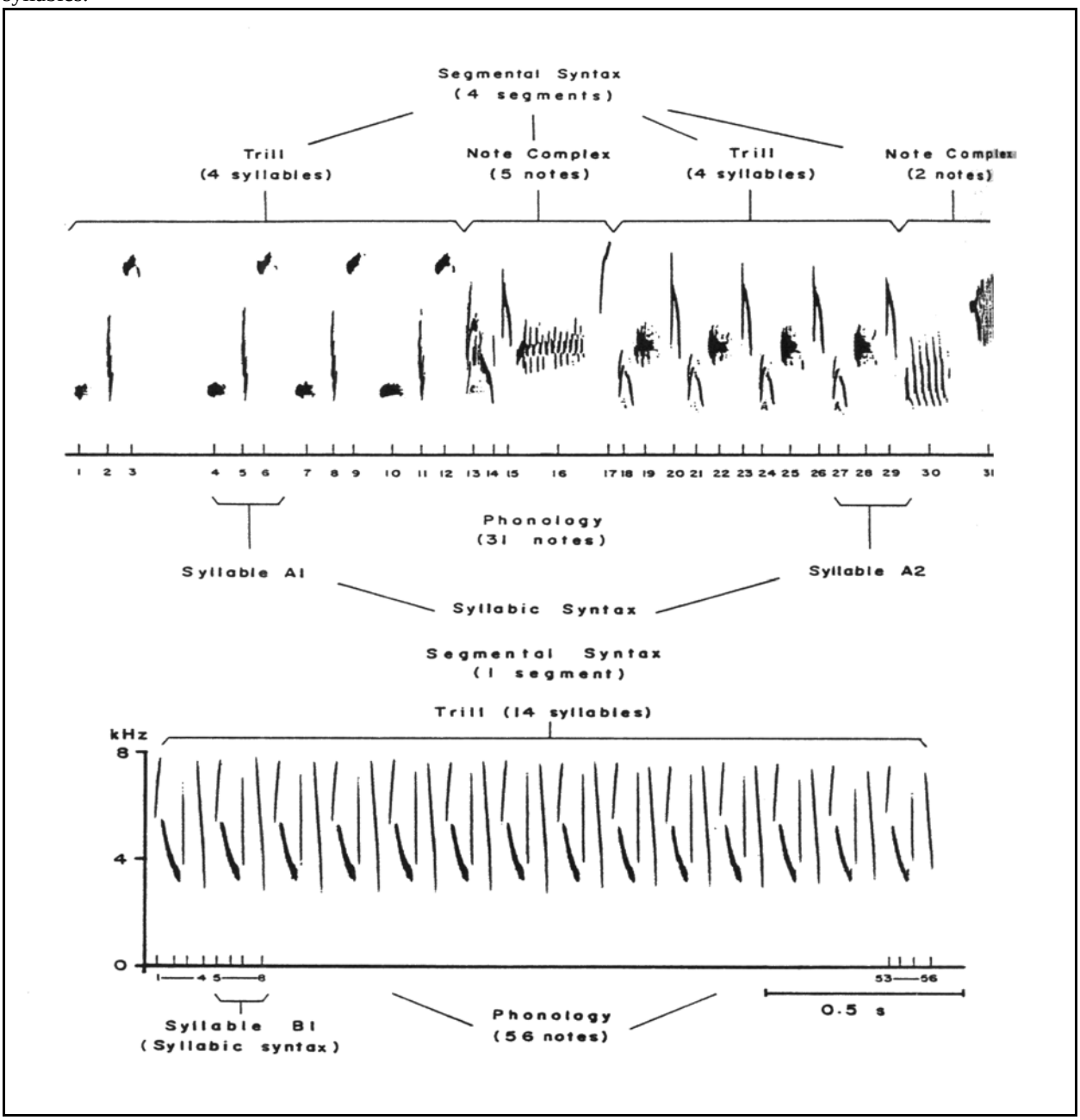

Figure U.4.7-14 Sonograms of typical song and swamp sparrow songs (300 Hz analyzer bandwidth). Top; This particular song sparrow song is made up of 31 separate notes (sonagram traces), which are in turn arranged syntactically into trills and note clusters. Bottom; The swamp sparrow song is syntactically less complex, with notes arranged into single trills. See text. From Nowicki & Podos, 1993.

U.4.7.3.5 Interpreting the utility of the CS technique of McCowan

The CS technique used by McCowan and associates limits exploration of the abilities of the dolphin in at least two ways. It limits the characterization of a dolphin communications symbol to a single frequency tone (whether modulated or not). It also segregates the individual whistle from any closely associated signals (symbols) that might be involved in creating a more complex morpheme. **Figure U.4.7-15** from Fletcher shows the situation involved in human speech is quite different. The left most column shows the commonly encountered phonetically defined

vowels. Note how large groups of the vowels exhibit virtually the same lowest frequency component but distinctly different higher frequency components, suggesting the isolation of the lowest frequency component using the CS technique does not correlate well with the phonetic symbols of human language. Fletcher notes, “As more data becomes available it is apparent there are considerable variations from these values by different speakers and also by the same speaker at different times. There is usually a third and sometimes a fourth format frequency in the high range but the variations from speaker to speaker and from word to word are so great that these higher components are considered to be characteristic of the speaker rather than of the vowel.

The right most column shows the combination of the consonant symbol *p* combined with the phonetically defined vowels. The result is frequently an overall waveform for the morpheme that includes a brief interval that allows for the isolation of the vowel from the consonant portion of the overall morpheme without difficulty (although such separation destroys the integrity of the morpheme).

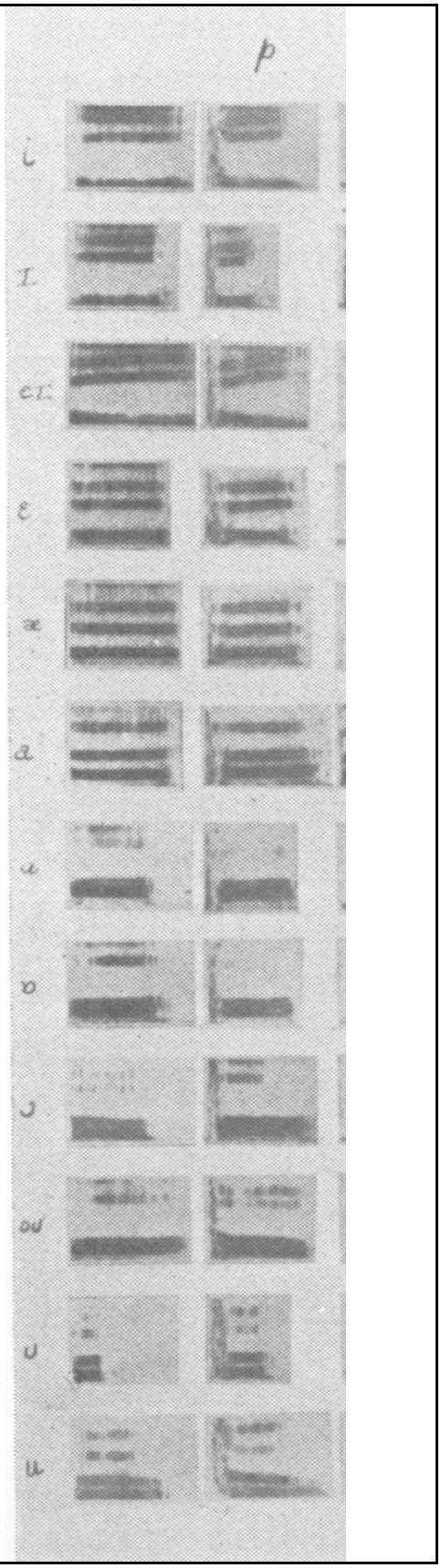

Figure U.4.7-15 Spectrograms of phonetic vowels and a set of consonant vowel pairs. From Fletcher, 1953.

Peterson & Barney[147] provided a possibly more convincing **Figure U.4.7-16**, reproduced in Lass[148], if an analogy with human speech is applicable to the speech of dolphins. It should be clear that selecting a single formant to describe a phonetic symbol is not adequate. At least two formants must be determined to define a single vowel uniquely. They also provided a second figure in different form but leading to the same conclusion. In human speech, the second formant is typically a 2nd through 10th harmonic of the frequency at which the larynx vibrates, the first formant. The third formant is usually a much higher harmonic, the 9th to 14th, 15th or 16th. The two vowel phonemes shown below the 2:1 ratio are largely unvoiced. They are largely dependent on the larynx generating noise instead of a tone for a pseudo F_1. That noise is then shaped by the oral cavity to form a pseudo but narrower noise bandwidth and more recognizable, F_2. There are no vowels reflecting only the basic F_1 generated by the larynx (at least in the major languages).

The detailed study by Peterson & Barney confirms the 1930 presentation by Sir Richard Paget showing each vowel consisted of two distinct frequency contours[149]. Peterson & Barney provide considerable background on their test equipment and its calibration.

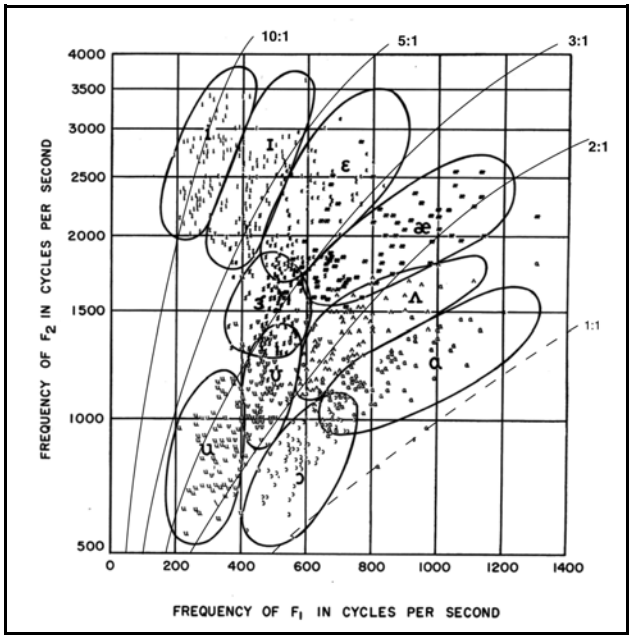

Figure U.4.7-16 Frequency of second formant versus that of first formant for 10 vowels spoken by 76 speakers. The ordinate scale is logarithmic while the abscissa scale is linear. The phonetic labels are from an archaic version of the IPA phonetic alphabet. Harmonic locii added to Peterson & Barney, 1952.

[147]Peterson, G. & Barney, H. (1952) Control methods used in a study of vowels *JASA* vol 24, pp 175-184

[148]Lass, N. *ed*. (1976) Contemporary Issues in Experimental Phonetics. NY: Academic Press

[149]Russell, G. (1970) The Vowel. College Park, MD: McGrath Publishing page xv

The requirement for more than a single formant to define a phonetic vowel in human speech is demonstrated by the development of various formant-based vocoders[150] (devices designed to mimic the human voice).

U.4.7.3.6 A vocoder modified to emulate dolphin vocalizations

Figure U.4.7-17 from Flanagan shows a common form of vocoder modified here to support the laminar flow path through the larynx of the dolphin as discussed in **Section U.2.3**. A more modern variant would provide a recording and playback capability between the central dashed line labeled speech defining signals. The resonators on the right are tunable by the musculatura of the nares. They are provided three types of flowing air via the larynx emulation at the lower center; turbulent air generating noise, laminar flow air that is essentially noise free, and pulsed air exhibiting a fundamental frequency adjustable over a limited range by the vocal chords.

The original vocoder of Flanagan and others should have also included a laminar flow path through the larynx to allow the proper formation of the s sound in the word “speech.” This sound is made placing the tongue near the palate and breathing out. This sound can be turned into a common whistle by bringing the tongue forward into a position behind the pursed lips.

[150]Flanagan, J. (1972) Speech Analysis Synthesis and Perception. NY: Springer-Verlag section 8.5

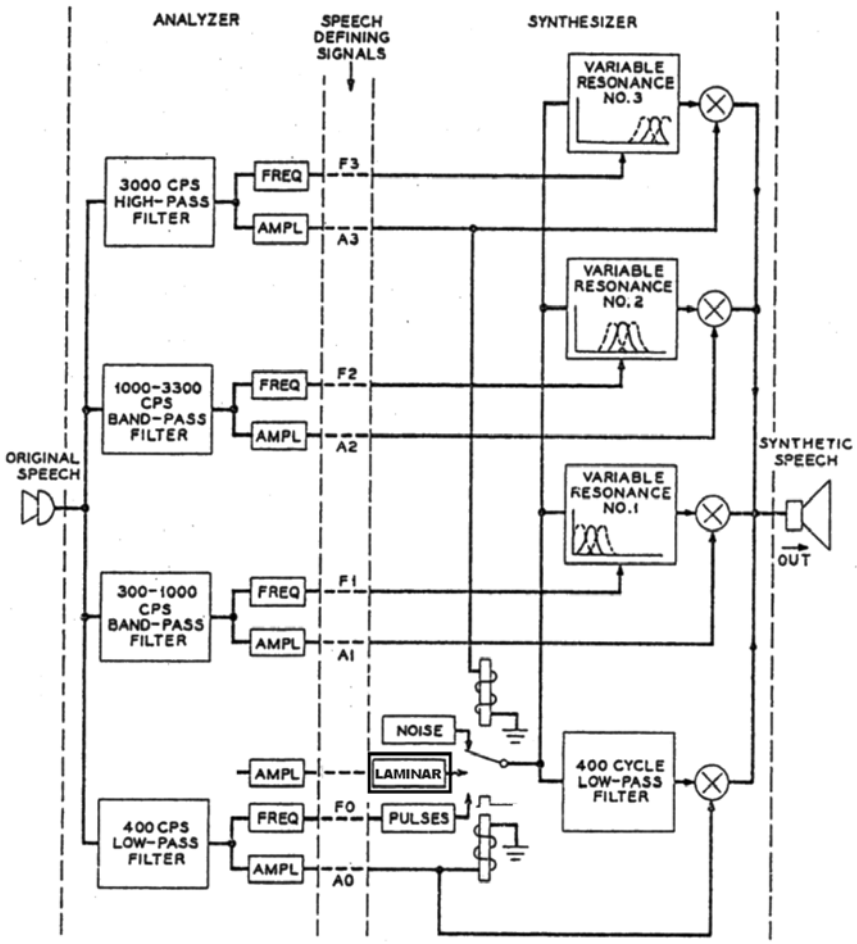

Figure U.4.7-17 Parallel-connected formant vocoder expanded for the dolphin. The switch at lower right is modified to support three positions, one allowing laminar flow air to flow through the emulation of a dolphin larynx. The bandwidths shown are for human speech, different values are appropriate for the dolphin. See text. Modified from Flanagan, 1972

U.4.7.4 Exploitation of information theory by McCowan (2000-2012)

During the first decade of the 21st Century, the McCowan team has been expanded to address the many sophisticated areas of research resulting from the work during the last decade of the 20th Century. In 2002, a paper that appears to have been written largely by Doyle appeared[151]. In this paper, the CS technique was modified to sample the normalized waveforms at 60 time intervals instead of 20, preserving more of the fine variations in frequency of the waveforms. It did note the undersampling present in their current database with regard to entropy calculations. Their Table 1 omitted the values for phonetic Russian incorporated in earlier papers ($M = 41$, $H_0 = 5.38$)and any row describing phonetic English (which uses 44 phonetic symbols rather than 27, i.e., $M = 44$. $H_0 = 5.46$). It left the phonetic investigations of dolphins and squirrel monkeys to be compared to the written alphabetical data of Russian, English and Arabic. The paper was not of great interest to the subject of dolphin speech.

- - - -

In 2002, McCowan, Doyle & Hanser explored how their earlier exploratory papers could be mined using the tools of Information Theory more effectively to quantify the characteristics of dolphin communications, particularly from the perspective of higher order entropies associated with semantic structures. In the course of this analysis, they reexamined their previously reported signal repertoires using a 60 sample, rather than 20 sample matrix of frequency levels for the baseline (not the fundamental) frequency waveform associated with a whistle. They presented a table of entropy values similar to that in their 1999 paper but for some reason deleted the values for phonetic Russian which were most comparable to dolphin whistles. In fact, their dolphin values are the same as those computed using only the 20 sample matrices of the 1999 paper. They did note that both their dolphin and squirrel monkey data remained under sampled. No data for Chinese was included in the table.

The ineffectual protocol used by McCowan et al. (See Sayigh et al. 2007) to separate signature whistles may have contributed to their arrival at a very small phonetic symbol set of only 27 symbols in both the 1999 and 2002 papers..

- - - -

In 2006, Stacie Hooper joined the McCowan team in order to explore mimicry between dolphin and both human and machine generated whistles[152]. The conclusion in the initial paper of this collaboration found, “Over a six-month study period, we experimentally exposed two infant male bottlenose dolphins and their mothers to six novel computer-generated whistles that were either unpaired or paired with specific contextual events (preferred toy objects). The results demonstrate that acoustic exposure alone was sufficient for spontaneous vocal imitation to occur but that context affects the timing, extent and quality of vocal imitation by bottlenose dolphins.”

During the first decade of the 21st Century, Ramon Ferrer-i-Cancho joined the McCowan team[153] and focused on interpreting the meaning of dolphin sounds (again constrained to whistle types). These two papers are highly mathematical but well supported.

The Abstract of the first paper begins, “We show that dolphin whistle types tend to be used in specific behavioral contexts, which is consistent with the hypothesis that dolphin whistle have some sort of ‘meaning.’ Besides, in some cases, it can be shown that the behavioral context in which a whistle tends to occur or not occur is shared by different individuals, which is consistent with the hypothesis that dolphins are communicating through whistles. Furthermore, we show that the number of behavioral contexts significantly associated with a certain whistle type tends to grow with the frequency of the whistle type, a pattern that is reminiscent of a law of word meanings.”

The care taken in exploring these relationships is shown in another quotation from the paper, “Notice that evidence of these positive or negative correlations does not constitute a proof of “meaning” (the problem is a particular instance of the fact that correlation does not imply causality). Correlation with context is a consequence of

---

[151]McCowan, B. Hanser, S. & Doyle, L.(2002) Using Information Theory to Assess the Diversity, Complexity, and Development of Communicative Repertoires *J Comp Psychol* vol 116(2), pp 166–172

[152]Hooper, S. Reiss, D. Carter, M & McCowan, B. (2006) Importance of Contextual Saliency on Vocal Imitation by Bottlenose Dolphins *Intern J Comp Psych* pp 116-128

[153]Ferrer-i-Cancho, R. & McCowan, B. (2009) A Law of Word Meaning in Dolphin Whistle Types *Entropy* vol 11, pp 688-701

"meaning" but not a sufficient condition. If we did not find evidence of contextual correlation we could conclude that dolphins whistle types have no meaning at all (assuming that the contexts considered are conveniently defined). If we found them, we could conclude that the hypothesis that they have meaning stands (by now)."

They conclude, "We have seen that the statistical properties of the mapping of dolphins whistle types into meanings is consistent with the hypothesis that dolphins whistles have some sort of "meaning" and that dolphins are communicating through them: whistle type use is constrained by the behavioral context and it can be shown that these constraints are shared by different individuals in some cases. We have found further connections with human language: the pattering of dependence the frequency of a whistle type and its number of contexts is reminiscent of a law of word meaning, something which is expected from standard information theory. These features of communicative systems challenge the claim that the finding of Zipf's law in dolphin whistle types could be the result of a simplistic die rolling experiment. Instead, the ubiquity of Zipf's law in natural communication systems could be explained by general principles of communication based on information theory. Statistical patterns offer a unique chance to a mutual understanding of language and other natural systems through a unified approach."

In the second paper, he broadened his focus to include inter whistle correlations[154]. "We have demonstrated that, for the majority of individuals, a dolphin whistle carries (on average) a significant amount of information about the next whistles of the sequence. Global randomization indicates that a whistle carries information about at least one of the next two whistles of the sequence whereas local randomization indicates that a whistle carries information about at least one of the next four whistles of the sequence. This is a property that is inconsistent with die-rolling, where a pseudo-word carries no information at all about the next pseudo-words of the sequence."

- - - -

A major point made in the second paper is, "We have demonstrated that, for the majority of individuals, a dolphin whistle carries (on average) a significant amount of information about the next whistles of the sequence (Tables 2 and 3). Global randomization indicates that a whistle carries information about at least one of the next two whistles of the sequence whereas local randomization indicates that a whistle carries information about at least one of the next four whistles of the sequence (Table 4). This is a property that is inconsistent with die-rolling, where a pseudo-word carries no information at all about the next pseudo-words of the sequence." This is very strong evidence for a language protocol supporting dolphin communications.

They conclude their discussion with, "The big question that future research on dolphins whistles must address is: what is the communicative complexity of a system whose units (e.g., whistles types) are distributed following Zipf's law for word frequencies [8], show a parallel of Zipf's law of meaning distribution [11] and form sequences with correlations that defy a simple explanation such as die rolling or Simon's model? We hope that our research stimulates further data collection to determine if the rather short range correlation discovered here are an intrinsic property of dolphin whistle communication or a consequence of the small size of our dataset."

> The assimilation of the above material from multiple papers and authors leads to a null hypothesis even in the absence of consonant sounds from the investigations. The null hypothesis is that *dolphins, at least the bottlenosed dolphin (Tursiops truncatus) employs a language protocol supporting intraspecies communications and therefore the dolphins employ a "natural language."*

To proceed further in the application of Information Theory to communications/language among dolphins or any other advances animal species will require more sophisticated statistical data gathering techniques, such as "capture-recapture." With this thought in mind, the text by White et al. is suggested[155].

U.4.7.5 Contributions of the Janik group during 2000-2014

The Janik team has focused on identifying the whistle component identifiable as the name of one or more members of a bottlenose dolphin cohort. Janik, Sayigh & Wells reported on only a few animals of the Sarasota, Florida cohort they have been working with over a long period[156]. They employed a complex procedure to create synthetic signals that were ostensibly stripped of the personal vocal features that might be associated with the signature

---

[154]Ferrer-i-Cancho, R. & McCowan, B. (2012) The span of correlations in dolphin whistle sequences arXiv:1205.0321v2 [q-bio.NC]

[155]White, G. Anderson, D. Burnham, K. & Otis, D. (1982) Capture-Recapture and Removal Methods for Sampling Closed Populations. Los alamos, NM: Los Alamos National Laboratory

[156]Janik, V. Sayigh, L. & Wells, R. (2006)Signature whistle shape conveys identity information to bottlenose dolphins *PNAS* vol 103(21), pp 8293-8297

whistle of an individual subject (without unfortunately having established the relevance of the features they sought to remove). They used a common amplitude envelope that was used to modulate the amplitude of a frequency modulated formant generally describable as of type 2 in human linguistic research. The original and synthesized signals were on the order of 1.0 seconds long (including when they involved two loops). The amplitude component was compressed or stretched over a range of 4:1 before being used to amplitude modulate a slowly frequency modulated tone. Little explanation for the justification of this procedure was given. Fortunately, the mechanisms within the cochlea may have suppressed much of this, probably extraneous, amplitude modulation.

Janik et al. (2013) provided a description of their methodology,[157] "We present a novel method, SIGnature IDentification (SIGID), that can identify signature whistles in recordings of groups of dolphins recorded via a single hydrophone. We found that signature whistles tend to be delivered in bouts with whistles of the same type occurring within 1–10 s of each other. Non signature whistles occur with longer or shorter inter whistle intervals, and this distinction can be used to identify signature whistles in a recording." Their SIGID method is a "human observer method looking at bouts to delineate signature whistles with a whistle duration of at least 100 msec.

They note the SIGID method had been tuned to be "extremely conservative" in order to avoid false positives while necessarily "missing about half of the signature whistles in the sample." Their table 1 summarizes the results. The feature of their filter is, "SIGID can identify signature whistles from recordings of free-ranging dolphins, even if many dolphins whistle at the same time, since the interwhistle interval is only measured between whistles of the same type."

King & Janik reported in greater depth on some recent work with a cohort to the East of Scotland, in Moray Firth and St. Andrews Bay[158]. The work dealt with at least twelve interactions with portions of the wild cohort in that area involving signals up to 20 kHz bandwidth. Their figure 4 and Table 1 are quite valuable in establishing the likelihood that they established the signature whistle of at least one dolphin successfully. They employed several little known statistical tests to demonstrate this conclusion using even small numbers of subjects. What they describe as treatments 11 & 12, where they played back natural recordings of apparently the same animal and that animal responded, are particularly relevant. Their discussion draws a number of conclusions based on logic that appears reasonable. Many of their synthetic signals were quite long in duration (typically 1 sec or longer) whereas their treatments 11 & 12 were on the order of 0.5 sec. Many of their control signals, from other dolphins both familiar and unfamiliar to the cohort were very complex and may have led to the low statistical scores related to the recognition of these signals. The signal in treatment 11 may well have been primarily a whistle. However, treatment 12 shows a more complex waveform incorporating both a whistle and what appears to be a voiced interval of about 0.1 seconds. This waveform needs to be analyzed using a better quality graphic.

Overall, the work during 2000-2014 shows real progress in identifying signature whistles on the assumption that the individual dolphin names are composed of only whistle elements.

U.4.7.6 Activities of Kassewitz, a trained linguist investigating the dolphin ADD

Kassewitz has a major program under way to learn more about the intraspecies communications among dolphins. His group has established a modern broadband recording system flat from a few hundred Hz to about 150 kHz. This provides a capability to observe all aspects of dolphin acoustic emissions. To support this program, the author has collected additional data about previous efforts to understand both intraspecies and interspecies (with humans) communications.

- - - -

Kassewitz defined a number of characteristics of dolphin communications (using the term language without a defined language protocol).

Fusional Language–A language in which one form of a morpheme can simultaneously encode several meanings. A morpheme is the smallest meaningful unit in the grammar of a language.

- Example: The word unladylike consists of three morphemes and four syllables.
- Morpheme breaks:
- un- 'not'
- lady '(well behaved) female adult human'
- -like 'having the characteristics of'

---

[157]Janik, V. King, S. Sayigh, L. & Wells, R. (2013) Identifying signature whistles from recordings of groups of unrestrained bottlenose dolphins (Tursiops truncatus), *Mar Mammal Sci* vol 29(1), pp 109-122

[158]King, S. & Janik, V. (2014) Bottlenose dolphins can use learned vocal labels to address each other *PNAS* vol 111 (>11) *Early Edition*

- None of these morphemes can be broken up any more without losing all sense of meaning. Lady cannot be broken up into "la" and "dy," even though "la" and "dy" are separate syllables. Note that each syllable has no meaning on its own.

Contoured Language– A language that has contours that are distinctive, rising and falling patterns of pitch, tone, or stress that affect meaning and understanding of this language. Acquiring the ability to use tone in language is child's play.

Tonal Language– A language in which different tones distinguish different meanings. Tonal languages fall into two broad categories: register and contour systems. Mandarin and its close relatives have contour systems, where differences are made not based on absolute pitch, but on shifts in relative pitch in a word. Register systems are found in Bantu languages, which more typically seem to have 2 or 3 tones with specific relative pitches assigned to them, with a high tone and a low tone being the most common (plus a middle tone for languages that have a third pitch).

An atonal language may exhibit changes in tone within a word but they are more associated with accents (a Texas twang) rather than changes in the meaning of the word.

Based on these definitions, Kassewitz et al. have asserted, “Fusional Contoured Tonal Language was established to exist in three different bottlenose dolphins at varying location in the Americas.”

Kassewitz has published a CD available via iTunes, “Dolphin Code,” where it is classified as New Age Music. It is primarily recordings of an organ, drums and other musical instruments with only a smattering of dolphin sounds thrown in.

U.4.7.7 A potential pulse based language by Preben Wik

Preban Wik has recently presented a thesis that contains useful insights[159]. However, it did not culminate in any actual data collection. He did not investigate the physiology of the dolphin in any serious way (and I do not believe the dolphin can reach 200,000 Hz in any meaningful way). He did not seriously discuss the frequency range of dolphin pulse signals (a maximum in the 150 kHz region) and tonal signals (a maximum in the 35---40 kHz region).

His focus on the pulse signals of the dolphin in the development of his protocol does not appear as useful as a focus on the combined pulse and tonal signals in ultimately establishing a closed symbol set for dolphin vocal communications. I would argue that the dolphin employs both the pulse and tonal portions of its repertoire in its language (roughly similar to consonants and vowels in English).

His paper did include some useful ideas, and compared the guidelines of Chomsky versus Hockett. He did note on page 53 that Hockett’s list may be too granular to fit the general case.

He did contribute the concept that dolphinese might employ pulses as *suffixes* to tonal formants, instead of the way humans use unvoiced *prefixes* to tonal formants. This approach might be more useful in a noisy environment.

The question of order also extends to dolphinese syntax. He noted, “There are six possible permutations of the units used for word order: SVO, VOS, OVS, VSO, OSV, and SOV. He noted correctly, every combination is used on one human language or another. However, he did not speculate on the order preferred in native dolphinese. Making this determination alone would be a key factor in breaking the dolphinese language code.

He defined a small set of SMU’s (symbolic message units). I do not believe he went far enough in establishing the complete range of potential morphenes (SMU’s ?) that his data collection device should address.

It would be interesting to run some broadband tapes through an expanded variant of his LVQ Tab program on page 98 (with the purpose of isolating the various morphemes and syntactic sequences). The result could be a set of statistical tables that could lead to a serious code breaking activity.

- - - - -

U.4.7.8 The role of bubblestreams in dolphin communications

Fripp has provided a significant paper concerning the relationship between bubblestreams and whistles particularly

[159] Wik, P. (2002) Building Bridges. Univ of Oslo in fulfillment of a Doctoral Degree

with respect to calving in dolphins[160]. The term “not representative” in her title may lead to argument. After presenting considerable statistical data, and a series of caveats, Fripp concludes,
“The reasons that dolphins use bubblestreams have never been clear.”
“The data do not strongly support the idea that bubblestreaming is involuntary or an indication of breathing problems.” and
“The data in this study suggest that bubblesteams may be used to indicate either distress or location.”

McCowan responded with an objection primarily on semantic grounds and one question of statistical procedure[161]. “While generally agreeing with Fripp,” McCowan argued that bubblestream whistles are a component of the overall dolphin repertoire, even if not a representative component, but instead a small component. Some imagery has shown adolescent dolphins emitting bubblestreams apparently during group play.

The precise role of these bubblestreams and their correlation with whistles at the time of calving remains undefined. However, they do appear to play a role during this period. It is possible the air sacs within the nares of the neonate have not fully inflated, or been purged of amniotic fluid at this time, necessitating the formation of some bubbles in order to vocalize.

U.4.7.9 The Telephone Transcription Project of Greenberg

Greenberg and Greenberg et al. provide an immense amount of information gleaned from transcribing telephone conversations from American English speakers around the USA using a phonetic alphabet.

Greenberg et al[162]. concluded, based on their transcripts in the phonetic alphabet of messages passing over a telephone system, “Detailed phonetic transcription of a spontaneous-speech corpus indicates that the spectral properties of many phonetic elements deviate significantly from their canonical form. Despite these spectral "abnormalities" such speech is almost always understandable, suggesting that other properties of the signal, such as segmental duration, may provide significant cues for intelligibility.”

The Switchboard Transcription Project paper[163] provides a very good overview of the difficulties in transcribing natural English speech over telephone circuits. “Often, only the vaguest hint of the "appropriate" spectral cues are present in the spectrographic representation. Formant transitions, usually associated with specific segments (such as liquids or nasals), are typically either entirely missing or differ appreciably from the patterns observed in more formally articulated speech. Such deviations from the "canonical" phonetic representation pose a significant challenge to current models of speech recognition.”

> “Although only 25% of the Switchboard lexicon consists of mono-syllable words, over 82% of the spoken words are in fact just one syllable in length (Table 6). This is the case both for the Switchboard corpus and for a comparable corpus collected and analyzed more than 65 years ago (Koenig, French and Carter, 1930).
>
> The reasons for this disparity between lexical potential (as indexed by the corpus dictionary) and utilization are likely to be several. First, most of the so-called "function" words, composed of such "closed" class entities as prepositions, pronouns, conjunctions, articles and auxiliary/modal verbs tend to be mono-syllabic. Second, even for such "open" class words as nouns, adjectives and non-auxiliary verbs, their is a decided preference for one-syllable options. This inclination toward monosyllabicity is illustrated in Table 7, which shows the 100 most common words for the Switchboard Transcription Corpus. Comparable frequency analyses for the Switchboard corpus as a whole, as well as for other spontaneous telephone discourse corpora (ibid) reveal a similar tendency towards monosyllabic lexical entities.
>
> The frequency of occurrence of these lexical items is shown in Figure 7. The frequency of occurrence for each word is shown on the left and appears to conform roughly to a "1/f" distribution consistent with Zipf's (1949) "law" (though his formulation was derived from written rather than spoken materials) and more importantly to Mandelbrot's extension of Zipf's formulation (Cherry, 1965) which explicitly incorporates

---

[160]Fripp, D. (2005) Bubblestream whistles are not representative of a bottlenose dolphin’s vocal repertoire *Mar Mammal Sci* vol 21(1), pp 29-44

[161]McCowan, B. (2006) Are bubblestream whistles unrepresentative of bottlenose dolphin whistle repertoires? *Mar Mammal Sci* vol 22(2), pp 492-495

[162]Greenberg, S. Hollenback, J. & Ellis, D. (1996) Insights into spoken language gleaned from phonetic transcription of the switchboard corpus, International Conference on Spoken Language Processing

[163]Greenberg, S. (1996) The Switchboard Transcription Project, International Conference on Spoken Language Processing ICSLP/JHU Workshop.

"information" (in a quantitative sense as explicated in Shannon and Weaver (1949)).

The right panel of Figure 7 shows the cumulative frequency for words in the Switchboard Transcription Corpus. The hundred most frequent words account fully for 60% of the lexical instances. This relation, in concert with the relatively small number of instances for words less common than the top hundred (Left panel of Figure 7) suggest that modeling these most frequently occurring words could be prove to be an effective strategy for modeling the pronunciation variation associated with this corpus (see below)."

In the summary, Greenberg notes,

"The present state of speech recognition focuses on the early stages of linguistic processing, in the hope that the higher level processing associated with semantics and syntax will eventually fall out of the wash. Words are assumed to constitute the basic building blocks of meaning, and phones are considered to serve as the basic elements of words.

The present transcription effort casts both assumptions into doubt, at least for spontaneous discourse. The exceedingly wide range of variation observed on the detailed phonetic level, both in terms of spectral features and in terms of acoustic realizations, suggest that the canonical set of phonetic elements upon which both our transcription system and orthography is based is inadequate to capture the linguistic information in the speech signal."

And finally, Greenberg notes with respect to his 3 kHz telephone channel,

"The detailed spectral properties of speech often appear less important than a much coarser spectral representation, combined with temporal properties unfolding over 100 - 400 ms. The latter perspective is commensurate with both the channel vocoder, developed in the late 1930's (Dudley, 1939) and with more recent work which associates speech intelligibility with modulation frequencies between 3 and 12 Hz (Houtgast and Steeneken, 1985; Drullman, 1994), distributed over a sparse spectral partitioning of the acoustic space."

The tabular data provided by Greenberg is of immense importance in defining a decryption program for dolphin speech.

U.4.7.10 Renewed focus on mimicry and copying

Following the work of Tyack et al. and of Richards et al. noted in **Section U.4.7.1** a new set of investigations have been reported beginning in 2006[164]. While these works have included a variety of spectrograms that provide visual comparisons, the majority of the analyses are statistical in character[165]. Esch, Sayigh & Wells focused on the interloop intervals and the maximum frequencies occurring within the signatures. They note the mean maximum frequencies for their subjects was as high as 27.3 kHz, considerably above the 20 kHz recording range of many earlier investigators. They also noted the various ways two loops could be connected. A summary of related investigations was included. King et al. provide a paper showing a variety of spectrograms where one dolphin "copied" the signature whistle of another[166]. Visual review of these spectrograms might suggest more variation than just attempts at coping. There may be elements of conversation among some of the spectrogram sets.

The use of the terms mimicry and copying used in textual discussion frequently suggests more precise copying than would be perceived by examining the actual amplitude profiles and spectrograms of particular signals. This assertion is supported by the images in Esch et al. (2009) and King et al. (2014). Sayigh et al. (2007) demonstrate the ability of humans to discriminate subtle features of spectrograms that are generally overlooked in textual discussions.

King et al. also introduced the use of multi-dimensional analysis (MDA) to this area, **Figure U.4.7-18**. However, they did not describe the parameters entered into the MDA in detail. It is clear that the underlying data set involves more that two dimensions. They do note the dotted lines with a and b suffixes refer to different relationships

---

[164]Esch, H. (2006) Whistles as potential indicators of stress in bottlenose dolphins (*Tursiops truncatus*) Wilmington, NC: University of North Carolina *Master's Thesis*

[165]Esch, H. Sayigh, L. & Wells, R. (2009) Quantifying parameters of bottlenose dolphin signature whistles *Marine Mam Sci* vol 25, pp 976-986

[166]King, S. Sayigh, L. Wells, R et al. (2013) Vocal copying of individually distinctive signature whistles in bottlenose dolphins Proc Roy Soc B http://dx.doi.org/10.1098/rspb.2013.0053

between the dolphin pairs than the others which were mother/calf pairs. If they had allowed for a third dimension and then plotted dimension pairs, it appears more meaningful results could be obtained. Rotating the axes in the paired presentations would have provided more easily interpreted results. The differences in the spectrograms for some of these subject pairs are shown in their figure 1. The spectrograms have been truncated at 20 kHz. The variation in the slopes of the lines suggests some of the dotted lines are not in the plane of the paper, suggesting the data set is more than two dimensional. Saying one dolphin copied the signature whistle of another may be an oversimplification. The first dolphin varied the context of its whistle and sometimes the number of loops in the overall "signature." Similarly, the copier approximated the original in a gross sense but not necessarily in a precise sense. There may be evidence of changes associated with conversation between some of these pairs. No identifiable voiced sounds appear in these signatures. The quality of the on-line and printed versions of figure 1 does not support reproducing it here.

May-Collado & Wartzok have reported on the variation of bottlenose dolphins from various locations within the Atlantic Ocean basin[167]. "The goal of their study was to provide a description of bottlenose dolphin whistles in 2 poorly known adjacent populations in the Caribbean of Costa Rica and Panama and provide insights on whistle variation by evaluating whether ambient noise, number of boats present, and zoogeographical relationships are associated with whistle variation between these 2 adjacent populations. We then summarize 8 published studies on bottlenose dolphin whistles from the Western and Eastern Atlantic to more broadly assess the role of distance on whistle variation." Their results were primarily statistical and includes un-calibrated dendrograms.

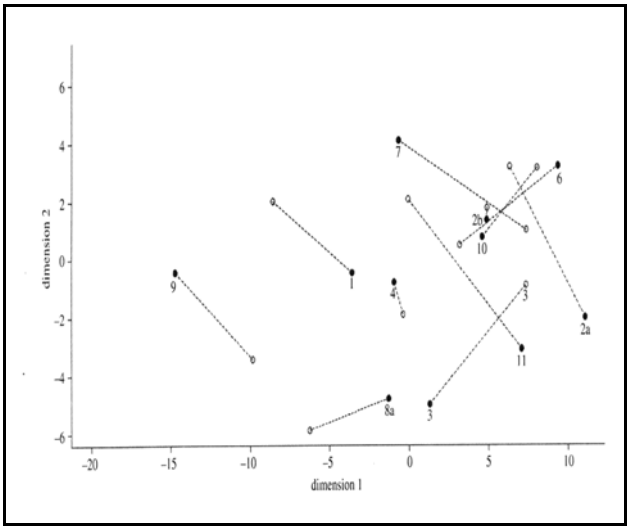

Figure U.4.7-18 Multi-dimensional analysis scaling plot between copies and originals based on all acoustic parameter measurements. Underlying data set appears to involve more than two dimensions. See text. From King, Sayigh, et al., 2013

A second paper by these authors showed that some dolphins employed whistle frequencies as high as 48.40 kHz[168]. "The study concludes that equipment with an upper frequency limit of at least 50kHz (150kHz for harmonics) is required to capture the entire whistle repertoire of the Guyana dolphin."

U.4.8 The probability of a dolphin language

Observers have been noting the actions of dolphins since antiquity. They have virtually all noted that dolphins are social animals usually traveling in pods. They are also known to cooperate in carefully choreographed (planned?) strategies when hunting. They are noted for their posturing and interpersonal rubbing (body language). Restrictly the discussion to vocal communications;

- Dolphins are noted for their large vocabulary of sounds that cover a spectrum much wider than that of humans (a few hundred Hertz to 40 kHz).
- Their sounds include a variety of voice and unvoiced sounds arising from their larynx and whistles arising within their nares (nasal passages).
- Their estimated phonetic symbol set has ranged from less than 100 (H_0 = 6.6) to up to 1000 (H_0 = 10.0) symbols, with a most likely size on the order of 50 symbols (H_0 = 5.6) based on the ability of humans to achieve a very large library of messages using less than 50 phonemes in a wide variety of configurations including a recursive feature).
- Their phonetic symbol set appears to contain phonemes with a duration as short as 18 ms (similar to humans).
- In terms of their vocalization capability, Markov and colleagues have asserted they have multiple sound producing mechanisms (compared to the single larynx of humans).
- They are noted for their considerable volume of vocalizations during a variety of group activities.
- Recently, it has been shown that dolphins can learn to correctly interpret complex semantics and even complex grammar when participating in one-way inter-species communications with humans.
- Dolphins quickly learn that the human may mislead them and they will adopt compensating strategies almost instantly.

---

[167]May-Collado, L. & Wartzok, D. (2008) A comparison of bottlenose dolphin whistles In the Atlantic Ocean: factors promoting whistle variation *J Mammalogy* vol 89(5), pp1229–1240

[168]May-Collado, L. & Wartzok, D. (2009) A characterization of Guyana dolphin (*Sotalia guianensis*) whistles from Costa Rica: The importance of broadband recording systems *JASA* vol 125, pp 1202-1213

• It has also been shown that they exhibit considerable self-awareness (a property presumed to relate to intelligence and cognitive ability).

Based on this long list of traits alone, it is difficult to deny that dolphins have the capability for intra-species communications by auditory methods employing a language protocol and generally described as language (the word used here as a noun). Whether the dolphin is able to employ recursion in its speech is impossible to determine until its symbol set and language protocol are deciphered.

When the apparent physical capacity of their brains are explored, they seem to have at least the potential of humans in cognitive ability. Their hearing capability vastly exceeds that of the human in frequency response and although its environment is fluid instead of gaseous, it seems to have excellent hearing sensitivity. The dolphin appears to have individually steerable external ears that are closely tied to its hydrodynamic form at any given instant. Its visual capabilities appear to exceed those of the human in the since they are comparable to that of the human, in both air and water. Besides these cognitive abilities, they are also well known for their apparently purposeful help to other dolphins and to humans in distress (even to the extent of protecting a human consort from potential shark attack).

Recently, a newborn calf named Wilson was found to be generating a wide variety of sounds (including multiple sequential whistles), and ranging with its echolocation capabilities within tens of minutes of its birth (Wilson first sound, second hour.wave at 1.5 sec). The ranging sequence was preceded by at most a few independent ranging pulses during the previous hour. The ranging sequence coincided with physical closure on its mother. These activities would suggest that dolphins have a considerable genetically-encoded acoustic repertoire at the time of their birth.

The only rational conclusion is that dolphins have all of the tools necessary to carry on intra-species communications using a rich and completely adequate auditory and body language. These tools are available, at least in primitive form, within minutes of birth. The challenge is to determine the characteristics of the auditory language, its phonology, vocabulary (at least the most common words), and its grammar. Killebrew et al. provide similar data on another neonate although the bandwidth of their recording equipment was severely limited at 12 kHz[169].

U.4.8.1 Initial categorization of dolphin phonetic symbols

Figure U.4.8-1 shows an attempt to organize the great number of labels used to describe dolphin vocalizations based on the fundamental sources and properties of these sounds rather than on their easily observable frequency parameters. The overall family of vocalizations, including the ultra high frequency sounds used in echolocation, fall into three sets based on th origin of the sounds. The identifiable origins are;
• the larynx,
• the soft tissue of the nares and
• the very hard tissue of the “ultrasonic lips” within the nare passages.

The major terms in this figure include,

Consonant– the label is used to describe unsustainable sounds with a typical length of 4 to 30 ms with a mean of about 18-20 ms. They may be voiced or unvoiced. The voiced variants generally lack sidebands but may be rich in harmonics.

Vowels– are sustainable sounds originating in the larynx and shaped within the nares with a typical length of about 300 ms. They are predominantly voiced and contain multiple harmonics. At least some of the harmonics (beginning with F2) exhibit sidebands related to the originating laryngeal vocalization (F1).

The question remains open as to whether the term vowel is appropriate here. In human speech, vowel implies the presence of a laryngeal formant, F1, and an associated oral cavity formant, F2.

In the context of vowels, “modulated” refers to the slow change in frequency of a contour with time and not the modulation associated with sidebands. Unmodulated refers to the absence of such slow changes in frequency.

Whistles– originate within the nares utilizing laminar air passing through the larynx (and are therefore labeled unvoiced). They generally lack sidebands but frequently are rich in harmonics. Some of the overtones are not true harmonics and show different contours than the fundamental whistle due to the change in tissue tension in specific sections of the nares.

In the recent literature, many long sequences of individual whistles, sometimes separated by consonants have been

[169]Killebrew, D. Mercado III, E. Herman, L. & Pack, A. (2001) Sound production of a Neonate bottlenose dolphin *Aquat mammals* vol 27.1, pp 34-44

labeled a single whistle. These are frequently described as signature whistles but are more precisely defined as signature sequences.

Dolphin vocalizations

Origin	Class	Unvoiced	Voiced	Sidebands	Harmonics	Common name	Ref.
Larynx	Consonants (short interval, unsustainaed)	✓				Pop/Bang	
			✓		✓	Bark	
	Vowels (sustainable)						
	Unmodulated						
		✓				Bray(1)	
			✓			Bray(2)	
					✓		
			✓	✓			
	Modulated		✓	✓ ?	✓	Chirp, at vowel freq.	
Nairs	Whistles	✓					
			✓	✓			
			✓		✓		
(Lips)	Ultrasonic (pulse streams)						
	Continuous					Buzz, high freq.	
	Gated					Buzz, low freq.	
	Variable pulse rate					Rasp	

Figure U.4.8-1 Proposed characteristics of dolphin vocalizations and the labels associated with them ADD. This table applies to one of two nares and possible one of two larynxes. The dolphin vocalization and echolocation capabilities may employ these sources in any combination. See text.

Ultrasonic pulse streams– used for echolocation and arise in either nare at the hard “lips,” sometimes labeled the “ultrasonic lips” or the “phonic lips.” Morris suggested “valvular flaps” of the nasal plugs as a more proper term. These pulse streams take on different forms during their utilization in the pursuit of pray as described below. The individual pulses consist of one or more cycles of a100 to 150 kHz carrier (a pulse width of about 0.12 ms is common). The ultrasonic pulse streams are included here because they frequently appear as overlays on the spectrograms of lower frequency vocalizations.

The spectral content of the individual pulses is a variable and may be diagnostic for a specific dolphin. The

content varies dependent on the detailed morphology and physiology of the lip and nare of a dolphin. It may also vary depending on the emotional state of the dolphin. See Appendix L and/or page 7 of Herman (1980).

Bangs– Described by Santos et al. as a loud broad band pulse of about 20 ms duration. Corresponds to the pop of this table

Bark– A short voiced consonant of duration similar to a pop. Used alone in a chain to form morphemes or used in conjunction with other phonemes to create a wide variety of morphemes.

Bray–This label may be used in two distinct ways. In one way, (1) it can be associated with the word itself where the vowel is that of the unvoiced a, (ei in the ASCII version of the human phonetic alphabet). In another context, (2) the word bray could be used to describe any sound where the vowel could be any one of the voiced vowels (i:, ai, Ou, u: etc). In the former case, the vowel in bray is typically described using only a noise-like component at low frequency formed in the larynx and one obvious formant developed in the soft tissue of the nares from a more laminar stream of air passing through either larynx of the two present. In the latter case, the bray may be based on two distinct and identifiable formants, #1 and #2 as usually found in human speech. The lower one originating in the larynx and the higher one developed in the nares from the first tone created in the larynx.

Santos et al. (1989) described a bray as a squeak followed by a grunt for a total of ~390 ms. This definition is deprecated here.

Santos et al. have also described a chain of 150 ms “noise like” brays, or simple unvoiced vowels, extending over a 700 ms span. In such a case, the bray set has been labeled a “**blast**.”

Buzz–The label is associated with the pulse repetition rate for echolocation signals consisting of one or more cycles of a pulse stream in the 100-150 kHz range used for echolocation. The human ear cannot accept and or detect signals significantly above 20 kHz. This is probably due to the limited resilience of the tympanic membrane, the lack of rigidity of the bones of the middle ear or a combination of both of these mechanisms. As a result, the signal introduced into the fluid of the inner ear (within the vestibule of the cochlea) is the result of demodulating the impressed signal much as the modulation riding on a conventional amplitude modulated radio carrier is. The resulting buzz perceived by humans is typically in the 5 to 100 Hz frequency range for gated and continuous pulse streams respectively.

If the perceived buzz changes in frequency during a series of echolocation pulse trains, it is described as a rasp. The change in rate is indicative of a change in range between the dolphin and its target. This change in frequency is well documented among bats. When perceived by humans, the buzz with a changing frequency may be described as a “**rasp**.” The frequency spectrum of the rasp is shown clearly in **Figure U.4.8-2** and discussed in greater detail in **Section 1.2.1.2** of Appendix L. The *T. truncatus* was closing on an artificial target while swimming freely. The text of Tyack is unclear related to these waveforms.

Chirp– A continuous voiced waveform beginning at one frequency and smoothly transitioning to a different frequency that is sustained as in the case of other vowels.

Pop–a short interval unvoiced pulse (nominally 18 to 30 ms long) forming the simplest consonant. When standing alone it can be very loud. Frequently used in a solitary chain to form one or more morphemes. Also used in conjunction with other symbols to form a wide variety of morphemes.

Rasp– A waveform perceived by humans due to a changing pulse rate in the ultrasonic pulse stream of dolphins during echolocation, particularly during the terminal phase in the pursuit of food.

The above list suggests the minimum consonant duration is similar to the duty cycle of the mammalian brain (on the order of 20 to 30 ms and best described using the critical flicker frequency of human vision, sections 7.5.1 & 17.6.6 in PBV).

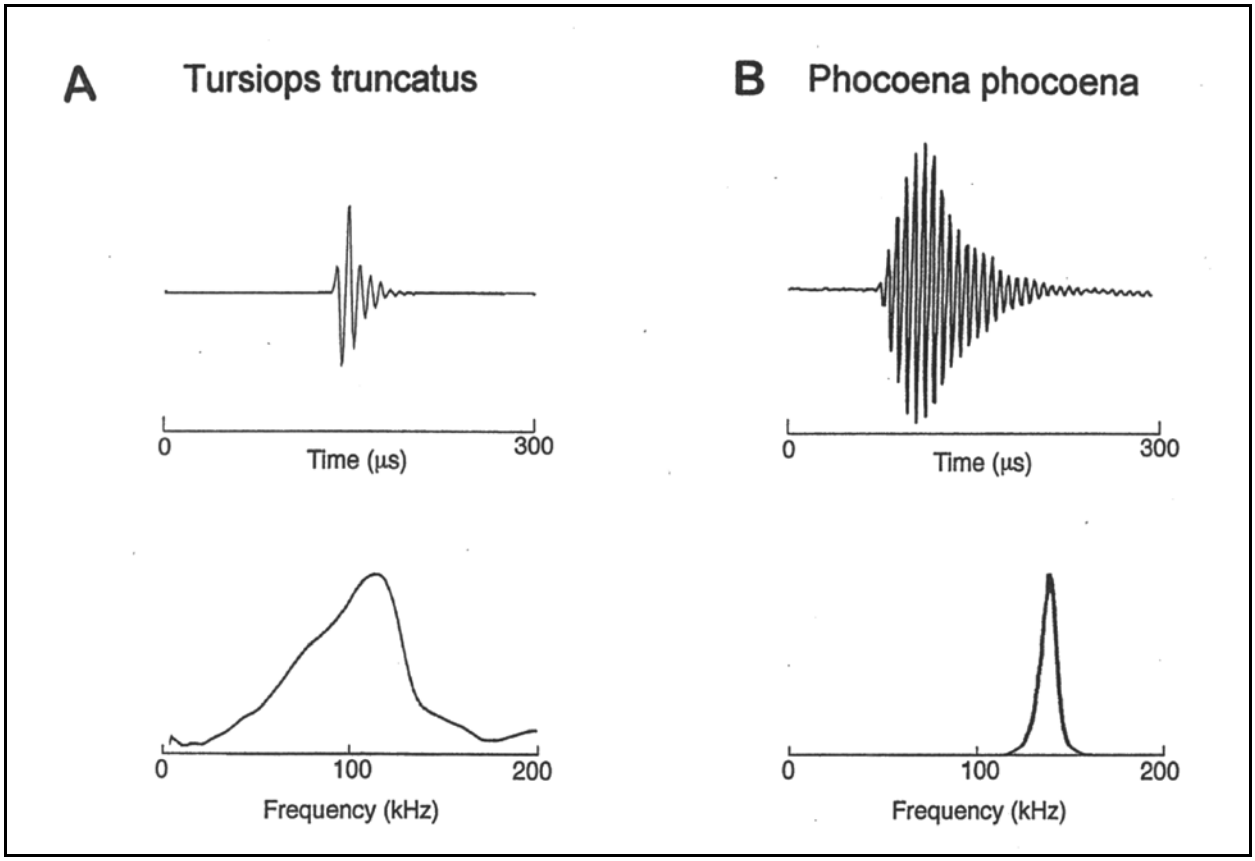

Figure U.4.8-2 Temporal and spectral waveforms of "rasp" and "buzz" ADD. Top; temporal waveforms. The *truncatus* waveforms represent the averages of the pulses in an entire click train with variable pulse interval, a rasp. The *phocoena* temporal waveform is for a single pulse from a similar pulse train of constant pulse interval, a buzz. The waveform at lower right is not a symmetrical Gabor function. See text. From Tyack, 2000; based on Au, 1980 (left) and Kamminga, 1988 (right).

U.4.9 Global view of task ahead to decipher potential dolphin

Several major challenges remain to deciphering dolphin speech:

- The analog character of dolphin vocalizations do not lend themselves to a closed symbol set.
    - Any entropy calculation based on an open symbol set is ephemeral if not misleading.
- With the frequency of occurrence of specific sounds (potential symbols) unknown, it will probably be necessary to artificially truncate the potential symbol set based on a minimum frequency of occurrence criteria.
    - Such truncation must be recognized in any description of an entropy value of any order.
    - Based on the results of Janik's team, in comparison to the results of the McCowan team, the truncation will require human review of the full signal spectrum of a large cohort of dolphins in order to translate the signals into a machine readable phonetic symbol set.
    - A recent estimate indicated it requires a ratio of 400:1 between human review time and dolphin signaling time.
- The success of limiting the symbol set to only 2nd formants of whistles (the dominant whistle frequency among the signals generated from laminar air flow through the larynx) is totally unknown at this time and eliminates all voiced and unvoiced (noise based) signals from the symbol set.
    - The available spectrographic recordings from dolphins indicate they do employ both voiced and unvoiced signals along with their whistle repertoires.

- The "signature whistles" recorded by many investigators appear too complex and too long to represent a single symbol; they appear to consist of one of a set of initial greeting symbol followed by additional symbols.

- The response of the dolphin receiving a communications (its pragmatics) will probably be crucial in determining the meaning of any putative language-based signals.

- The most likely context of any given dolphin communications suggests the communicants must be environmentally constrained during their observation.

- Initial success in deciphering any dolphin words will be (or may have been) based on the communications between two dolphins in separate pools connected by a telephonic channel that is recorded, thereby controlling the motivations and the dolphins and supporting close observation of their pragmatics..

The above is a long list of difficult problems to surmount at this time.

U.4.9.1 Boundary cues between phonemes and morphemes

Shoup & Pfeifer in Lass (1976, Chap. 6) have discussed the difficulty of differentiating between phonemes and morphemes in human language with several graphic examples. It is clear discrimination involves some training of the analyst. **Figure U.4.9-1** provides a clear example of the task using the word "chess." The separations only involve a few milliseconds but can be visually recognized by the change in character of the spectrogram components. The sinusoidal character of the vowel can be seen clearly in the amplitude profile but is largely masked in the spectrographic presentation. The need to make subtle differentiations between phonemes have been avoided when only examining dolphin whistles.

U.4.9.2 Estimate of the data processing problem

In 1989, Markov & Ostrovskaya (of the USSR Academy of Science) reported processing 300,000 signals to determine an initial estimate of the rank of the vocalization repertoire for a presumed multi-channel signal generator system within the bottlenose dolphin. Their figure 1 for one sound generator, using sketches, shows a much more complex set of signals at the word level than most other studies. Frames G & H display potential phoneme elements and their concatenation into a variety of morphemes. They noted, "On the average, one can identify 5–7 blocks in the signal, though their number can reach 12." These blocks were typically 17 to 80 ms long. They noted, "Actually, a whistle of any type amounts to an entire class of signals, since, while having the same shape of the contour, they can differ in duration, frequency range, register and rate of frequency modulation." They suggest a third level structure is common, essentially collections of morphemes. Their figure 3 shows the much more complex situation for two sound generators operating in parallel resulting in signals with identifiable overlaps in time. The relevance of this phenomenon is difficult to estimate, since such a capability is not known to occur in any other species. However, it clearly complicates the data reduction task attempting to catalog and interpret its significance. Their hypothesis is that each of up to four signal generators can operate in either the tonal or pulse environment. This paper restricts the signal generators to two "pure" whistle generators in the nares and one or two more conventional sound generators in the larynx. Their figure 4 sketches the operation of three generators simultaneously. They do not illustrate a potential four generator operation. They do state, "Signals whose structure is formed with the participation of four sound generators are observed rarely in bottlenose dolphins. While the share of signasl formed by one generator accounts for about 55%, by two generators - 40%, three generators - somewhat less than 5% and four generators - less than 0.1%."

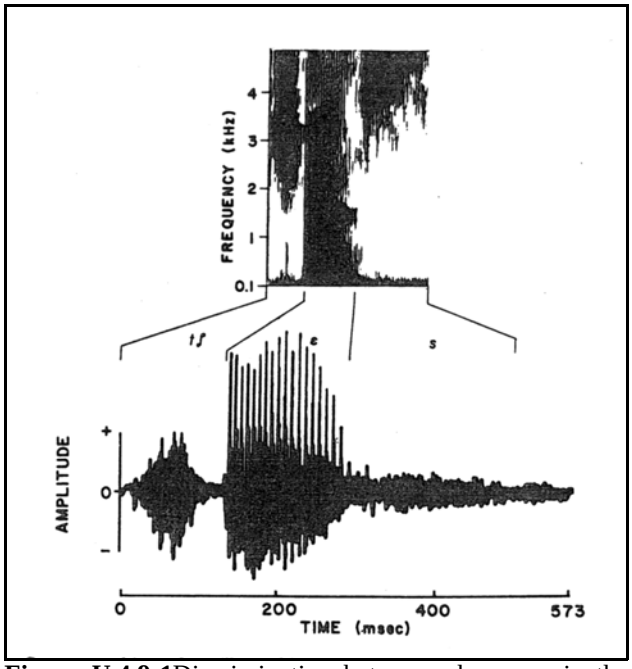

Figure U.4.9-1 Discriminating between phonemes in the word "chess" using the phonetic alphabet labeling. Note the prominence of the 120 Hertz sinusoidal component in the vowel portion of the amplitude profile that is masked in the spectrogram. See text. From Shoup & Pfeifer, 1976.

On page 613, they estimate the number of phonemes in a morpheme as 1 to 24 with an average of 5-7 and the

number of types of structural blocks (morphemes at well over 100, resulting in the potential from game theory calculations of 10^{12} distinct signals (phrases or sentences) in their complete corpus. Clearly, constraining the scope of the symbol set, and complexity of the morphemes and third level structures is important in limiting the data processing challenge in any cryptanalysis program envisioned in **Section U.5**.

The identification of first level (phonemes), second level (morphemes) and third level (phrases), the material for an entropy analysis on an open symbol set becomes possible. The results would be estimates based on a specific corpus of sounds.

Markov & Ostrovskaya also note, "It is rare that bottlenose dolphins produce single signals. As a rule, this is typical of very young or isolated adult animals. In normal communications, the intensity of signalization is very high, reaching sometimes 50 signals per minute." They also note the very important fact related to the planned cryptanalysis program, "In free dialogue (for instance, during communication of isolated animals through electro-acoustic communication link), *signals with different structures are combined into groups, the way human words are combined to construct phrases*."

Markov & Ostrovskaya also address the variation in prefixes and suffixes of apparently similar signals between dolphins during conversation and suggesting the complexity of their communications structures (syntax) and the organization of their communications content (semantics).

In 2005, Fripp of Woods Hole reported on 2100 minutes of data collected from 210 focal samples in 75 recording sessions. The data only involved bubblestreams and potentially associated whistles. She identified 200,000 cuts were examined from this collection yilding almost 20,000 whistles. Cuts containing overlapping whistles were eliminated from the data processing at this point. Following this process, the whistles were further processed to identify the most significant frequency contour. This was accomplished by removing the uppermost and lowermost 25% of the frequencies present in a given spectrogram. Six parameters were then measured relative to the remaining cuts. They also calculated a frequency asymmetry parameter (defined as the (median-min)/(max-min). The data was then subjected to a Hierarchical Cluster Analysis.

U.5 A formal investigation leading to the deciphering of the dolphin language(s)

[xxx Edit this combination of material under U.5]
The material within **Section U.4.7** shows that the dolphin exhibits all of the cognitive, vocal and hearing capabilities to carry on sophisticated intra-species communications using all of the major elements of a language protocol. Although they may be compared to the language protocol of only a child of the human species, the framework appears complete and may contain elements not found in the human language protocol (related to the multiple source generators in dolphin vocalization).

The immediate challenge appears to be one of cryptanalysis. It is likely there is communications going on among dolphins using a language protocol. However we do not have knowledge of the elements of that protocol or even the phonetic character set. Under these circumstances, it is useful to review Kahn's overview of the cryptoanalytic task (pages 435-450).

- - – -

It is important to establish a framework for this undertaking that is loosely based on the terminology of cryptanalysis. Referring to Kahn's original hard cover edition[170] of "Codebreakers" leads to a modified list of definitions that will be used here. His 1973 soft cover edition leaves out most of the technical material present in his 1967 edition. It is the technical material that is of major interest here. Clearly, the intent here is not to discover the plain text message (in the dolphins native language?) but to discover the equivalent text message expressed in understandable English text. In the framework of this book, we are interested in determining the instructions issued by the cognitive elements of stage 5 of the *Cetacea* by learning the structure of the signals created from the stage 6 (motor elements) of the *Cetacea*.

> In this context, it is a major goal to accept the acoustic signals of the *Cetacea* (that were generated by its stage 6 engines in response to its stage 5 cognitive activity and translate them into a (probably written) form that can be provided to stage 4 of the human for translation into ideas meaningful to stage 5 of the human cognitive system. [xxx may want a figure here]

Undertaking an activity of this scope is labor intensive and one must evaluate whether the various militaries of the world have already completed a similar mission. Based on the amount of published material from both government laboratories of both the USA and the Russian Federation in recent years, it appears safe to say the intensive efforts of

[170] Kahn, D. (1967) Codebreakers. NY: Macmillan

earlier decades were probably premature in their attempts in these areas. Conversely, success in this endeavor will require the application of a wide range of personal skills, personal motivations and recent advances in technology.

U.5.1 Developing the highest potential approach to decipherment

The goal is to achieve decipherment of the dolphin (at least bottlenose dolphin) language protocol as efficiently as possible.

- The approach outlined below adopts an action plan based on a modification of the outline of Reznikova.
- It focuses on a more general assumption about the dolphin phonetic symbol set than common in earlier studies.
- It accepts the use of captive dolphins as a plus during the initial stages of the plan.
- It accepts that broadening the decryption strategy will require working with dolphins in their natural environment.
- It assumes that nouns (both proper and personal) will be the first to be recognized through cryptanalysis.
- It seeks to limit the range of subject matter during observation of messaging between captive dolphins.

A review of the literature is needed after the following plan of attack is finalized. Marino et al. (2008), including a total of seventeen (primarily active academician) authors from a wide range of scientific fields, provided a broad description of virtually all of the information available about *Cetacea*, including dolphins and particularly the bottlenose dolphin. It included six pages of citations covering virtually all phases of academic research as well as behavioral studies on this family. Section IV demonstrates that nearly every facet of communications/language known to man was usable/used within the bottlenose community. Most of these capabilities were demonstrated during interspecies communications with humans, a presumably more difficult task than interspecies communications. The ability to use some 40 words, various syntactic arrangements and semantic forms were demonstrated during these experiments.

It is likely that the first success in cryptanalysis at the word level will involve nouns alone. As in the case of personal names, it is highly likely that dolphins would identify various food sources (using nouns) encountered during hunting and feeding in their conversations. With a collection of nouns identifying things, it is highly likely that dolphins would have created symbols to indicate the actions (verbs) associated with these identifiers.

U.5.1.1 The baseline block diagram of dolphin communications skills

To establish a starting point in exploring whether the dolphins use language, a number of propositions need to be addressed.

U. The difference between language and communications needs to be addressed in regard to oral signaling. It can be argued that animals communicate but do not employ language where language is identified by having a grammar. Some investigators attempt to define a dividing line between communications and the use of language based on the number of words a species can understand and/or the richness of the concepts that can be conveyed. Sometimes the presence of past and future tenses are considered a criteria.

2. A grammar can take many forms. It consists of arrangements of words representing a thought. The words are generally subdivided into those relating to a subject (S) performing some act, those describing the act and known as verbs (V), and those describing the object (O) acted upon. Clearly, there are less than a dozen orders in which these terms can be used. In English, the dominant grammatical form is a subject followed by a verb, followed by an object (S-V-O). In German, a more common form is a subject followed by an object followed by a verb (S-O-V). Professor McWhorter suggests this order is more common world-wide. Virtually all other combinations can be found in the languages of the world, including languages that consist of only single words combining S, O & V and expressing a sophisticated thought.

3.Human languages can be described using the expressions;

- tonal, those varying relatively continuously in intensity of the tones.
- stressed, those varying abruptly in intensity with definitive separation between syllables forming a single word.
- inflected, those varying in tone (frequency) between syllables.

To understand the variations in signaling that the dolphins might use to communicate, it is useful to review the recorded lectures of Professor McWhorter on human language[171].

Professor McWhorter discusses why languages developed, as well as why they differentiated over time. Many of

[171]McWhorter, J. (2004) The Story of Human Language. Chantilly, VA: The Teaching Company (36 lectures over about 18 hours)

the features of a language can be related to the environment and the culture of the speaker. As he pointed out, Proto-Indo-European had no word for palm tree, vine or Oak because of its origins on the plains (Steppes) of Southern Russia.

It is important to develop a scenario as to how and why dolphins might use language. A particular feature is their habit of hunting in packs in order to corral large numbers of fish before consuming them. The practice of river dolphins to proceed individually up river estuaries before coming together in a pack would suggest communications over distances of several miles and at least rudimentary language might be useful in achieving their goal. They clearly have the necessary communications skills. Whether they have the language skills is a subject of this discussion.

Many people have attempted inter-species communications with dolphins, typically using the audio bandwidth of the human, 200-15,000 Hz. Modern spectrograms of dolphins show their native audio bandwidth is approximately 2000-40,000 Hz. Initial wide band recordings of intra-species communications among dolphins suggest they use the entire bandwidth available for intra-species communications.

To attempt to discover whether dolphins use language in communications is a major code-breaking activity, not unlike some of the activities of the Second World War, including the Code-Talkers comprised of American Indians speaking a previously unwritten language and using substitutions of meaning as an additional degree of encryption.

The tree illustrated in **Figure U.5.1-1** based on Crystal helps understand the elements of communications needed here[172]. Prosodic refers to the metrics of speech and singing. Different languages adopt different rhythms. There are four sources of human sound as listed at upper right. They affect the timbre or quality of the voice.

[172]Crystal, D. (1987) The Cambridge Encyclopedia of Language. Cambridge: Cambridge Univ Press

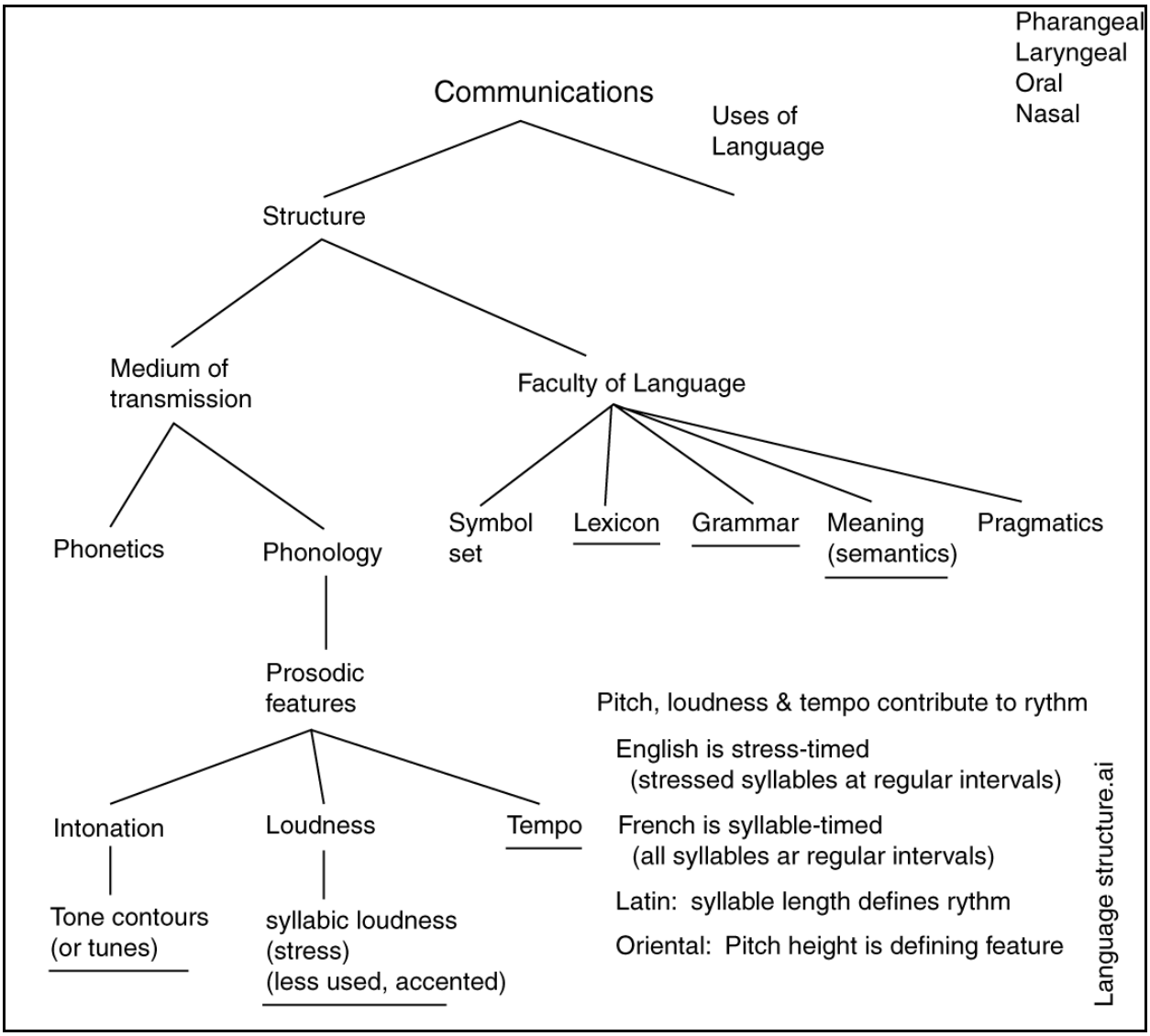

Figure U.5.1-1 Technical aspects of communications. Underlined characteristics are those identified at least initially in observations of the communications of *Tursiops truncatus*. The faculty of language has also been labeled the language protocol. See text. Modified from Crystal, 1987

Each aspect of dolphin communications explored and identified in the laboratory or field have been underlined. These observations strongly suggest that *Tursiops truncatus* (and most likely other species within *Tursiops* are using a language protocol supporting their vocalization in conspecific and interspecies communications.

A means of decoding dolphinese has recently been studied in the human context by Park & Glass[173]. The technique is described as unsupervised in that it records isolated speakers on the same subject verbatim. It then cross correlates their speech sounds to determine similar phonetic sequences in both recordings. These sequences are then analyzed with respect to their adjacent context.

- - - -

The long standing debate over whether the dolphins, and particularly the bottlenose dolphin, *Tursiops truncatus*, communicate within their species using a language protocol appears to be near solution. The animals are notoriously gregarious and well known to form temporary associations within their "fusion-fission" type of society. Second, they have been shown to have the intellectual curiosity and capability to support speech involving a language

[173]Park, A. & Glass, J. (2008) Unsupervised pattern discovery in speech IEEE *trans audio, speech, and lang proces* vol. 16(1), pp 186-197

protocol. Third, their vocalization mechanisms have far greater capability than needed to sustain "language." Fourth, their abilities at mimicry, from birth, are widely recognized. As shown using Crystal's diagram in **Section U.4.9.1**, research has demonstrated that the bottlenose dolphin has employed all of the elements (in at least rudimentary form) of a language protocol. The two papers of Marino et al. summarize all of the above capabilities. By careful concatenation of these elements, it has been possible to demonstrate two dolphins speaking to each other directly, or via electronic relay of their vocalizations between isolated enclosures. ***The result can only be described as the use of language as a communications skill***. The remaining challenge is to expand our documentation of this capability, and then to determine what to do with it!

For the last 25 years, the academic community has focused almost exclusively on the whistles of dolphins to the exclusion of pulse-sounds and various consonant forms. Until the 21st Century, the investigations were limited to only about one-half of the frequency spectrum used by the dolphins for communications/conversations (0 to 22 kHz out of about 40 kHz known signaling capability). As a result, they have been studying a fraction of the total symbol set used by dolphins and estimating entropy values based on this *open* (read incomplete) phonetic symbol set. These estimates have generally been based on only the first baseline frequency of the whistle, which is insufficient to delineate the vowel sounds of everyday English (Fletcher, 1953, page 60).

The goal of the following material is to address more than the delineation of signature whistles. It is to delineate to the greatest extent possible the lanuguage protocol of dolphins, with a focus on Tursiops truncatus.

In 2007, Baars & Gage discussed the question of "Speech versus Language' (page 325) once again, "a hard-and-fast distinction between speech and language is not justified."

U.5.1.2 Preferred method-operation totally in the marine acoustic environment

Reznikova has discussed the three primary potential methods of discovering the features of language in different species and given examples of the fragmentary cases where humans have discovered scattered elements of those languages[174] The paper is in a language familiar to linguists, "In this review, three main experimental approaches for studying animal language behaviour are compared: (1) direct decoding of animals' communication, (2) the use of intermediary languages to communicate with animals and (3) application of ideas and methods of the Information Theory for studying quantitative characteristics of animal communication. Each of the three methodological approaches has its specific power and specific limitations."

"The main difficulties in the analysis of animal 'languages' seem to be methodological. At least three main approaches to a problem of animal language behavior have been applied recently.

The first approach is aimed at direct decoding of animal signals. Although it is intuitively clear that many high social species have to possess complex communications, only two types of natural messages have been decoded up to the present. The matter concerns the symbolic honeybee 'dance language' and acoustic signals of danger, which were decoded for vervet monkeys.

The second approach is based on the use of intermediary artificial languages. Being applied to apes, dolphins and one grey parrot, this method has revealed astonishing mental skills in the subjects. It is important to note that this way to communicate with animals is based on adopted human languages.

The third approach to study animal communication has been suggested based on ideas of Information Theory. The main point is not to decipher signals but to investigate just the process of information transmission by measuring the time duration that animals spend on transmitting messages of definite information content and complexity."

The last approach is only a minor element, an element of traffic analysis, in the repertoire of Information Theory.

The list of Reznikova does not appear to be complete. The option of communicating between man and dolphin using computer generated acoustic language recognizable to the dolphin but appearing to be in a "foreign" language must be considered. The key to this approach is the total elimination of any human language from the investigation. It is basically the approach pioneered by Herman et al. as reported in 1984 and implemented more fully through the more extensive use of a vocoder of the type described in **Section U.4.7.3.6**. The human acts only as a test protocol planner an observer of pragmatics when appropriate. As part of the protocol planning duties, the human also supervises the data reduction activity.

The later data reduction activity, actual decryptanalysis, follows the earlier discussion of Reznikova. Reznikova presents an excellent survey of examples where animal communications having specific meaning for different

---

[174]Reznikova, Z. (2007) Dialog with black box: using information theory to study animal language behaviour Acta Ethol vol 10(1), pp 1-12

phonetic expressions have been decrypted. She even notes the preparation of a number of species specific dictionaries in recent years. She notes, "It is likely that a bottleneck for decoding animals' signals is low repeatability of standard living situations, which could give keys for cracking animals' species specific codes. In fact, animals sometimes behave similarly in repeatable situations, and if these repeatable behaviours are elicited by the distinctive repeatable signals, these behaviours can serve as keys for cracking animals' species specific codes. Decoding the function and meaning of wild communications is a notoriously difficult problem."

U.5.1.3 Probable range of variables based on a wider symbol set than whistles

There is a basic bifurcation in any cryptanalysis of an unknown language, particularly where the symbol set is unknown, the phonetic word-based approach and the phonetic symbol-based approach. Once either approach provides a degree of success, the intermediate concept of syllables becomes tractable.

Yip[175,176] has addressed the problem of decrypting the language protocol of the dolphins. He concludes his introduction with, "Not surprisingly, it turns out we have very few answers to these questions, so I then turn to how we might go about designing experiments to test for phonological skills. There I stop (in this paper), in the hope that someone might find these ideas sufficiently interesting to actually carry out these or similar experiments, and give us some answers." The Yip papers bring much to the discussion below. However, the text appears to be in a language tailored to the phonologist as opposed to the more general linguist or the field biologist. He discusses primarily bird and primate messages, but much of the discussion appears to apply to dolphins as well. He does note,

> "The duration of a motif is long compared to a human prosodic word (up to one second long, whereas a typical American English syllable is around 170 ms, with a range from 107-260 ms (Greenberg et al. 1996)."

> In an associated paper, Greenberg noted, "Overall, the transcription effort required nearly 400 times real time, which is far too time-consuming for transcription of an extensive corpus such as Switchboard. What is required is some form of intermediate-level transcription that captures the essential phonetic patterns and reduces the remainder to the bare minimum required for accurate word and higher level linguistic modeling." And, "Phonetic transcription has convinced many of us connected to the Switchboard project that much of the linguistically relevant information is not to be found in the detailed spectral properties of speech, but rather in the coarser temporal patterns, embedded within some form of syllable framework." And, " For this reason, we are about to explore a new form of transcription system in which different portions of the syllable are transcribed with varying degrees of granularity. *The tabular data provided by Greenberg is of immense importance in defining a decryption program for dolphin speech.* See **Section U.4.7.9**.

Greenberg, Chang & Hollenback have provided a comparison of human speech recognition approaches and implementations in 2000[177].

The Greenberg papers emphasize the necessity of narrowing the environmental conditions that are present when attempting to decrypt dolphin communications. *Restricting the dolphin communications to that over a wideband (preferably beyond 20 kHZ) telephone connection between two tanks containing small numbers of dolphins is probably an initial necessity*.

The phonetic symbol set of the dolphin will remain open for a considerable period of time. Based on the analysis of Smith's reading of the available literature, it appears a reasonable bound on the zero-order entropy for dolphin speech is in the $H_0 = 6.8$ area, nominally 112 symbols in its symbol set on the high side and 111 symbols on the low side. The first-order entropy, H_1 will be lower and is likely in the 6.0 or less area for a given cohort (64 or less high frequency of occurrence symbols). These numbers should be large enough to include all of the voiced, unvoiced and whistle sounds once a more orderly collection of phonetic symbols is arrived at (including the discernment of the individual whistle symbols included in the typically long whistles (treated as words) reported in the literature.

As McCowan et al. noted in 1999, "We found the literature contained neither sufficient information for us to assess the completeness of vocal repertoires nor the frequency of use of different vocalizations within repertoires, including

---

[175]Yip, M. (2006a) The search for phonology in other species *Trends Cogn Sci* vol 10, pp 442-446

[176]Yip, M. (2006b) Is there such a thing as animal phonology? Linguistics Research Center, Univ. California, Santa Cruz pp 311-323

[177]Greenberg, S. Chang, S. & Hollenback, J. (2000) An introduction to the diagnostic evaluation of switchboard-corpus automatic speech recognition systems. Proceedings of the NIST Speech Transcription Workshop, College Park, MD

the temporal integrity of vocalization sequences necessary for higher order entropic analyses."

A major simplification of the overall program would occur if a method of placing much of the recorded communications of dolphins under constrained conditions could be assembled in a common data base. Any recordings of constrained dolphins in separate tanks connected by telephonic circuits would be particularly valuable.

U.5.1.3.1 Assembly of initial symbol and word lists

A concerted effort to organize the available information of the words (and possibly symbols) used in prior work with dolphins is necessary. At this stage individual word lists from different cohorts (and possible individual trainers or training activities) from around the world should be segregated. The "foreign" language list of words from Herman et al. (1984) would probably form a good starting point. They involve no human vocalizations. Other lists established during experiments with synthetic dolphin sounds could be arranged in parallel columns with the Herman et al. material as a first step forward.

Verboom has offered several useful suggestions on standardizing the collection and presentation of the symbol data based on an expansion of the relevant human acoustic standards[178]. The tabulations must be expanded to account for the dolphins range of hearing. His notation relative to 1/3 octave intervals may not apply to the dolphin.

U.5.1.4 Initial operational methodology in the field–a constrained environment

Achieving a probable framework and initial vocabulary appears to be a tall order. As noted earlier, it will be much more difficult than standard code breaking where the character pool used is already known and traffic analysis can provide additional clues. The challenge appears to be to create a simple situation that is sufficiently stylized to involve only simple messages that do not involve a large potential for the appearance of synonyms or changes in grammatical order. This has come to be known as the "Rochefort test" in some communities. Rochefort suggests a scenario during the Second World War that caused the Japanese to use a specific expression in one of their encrypted messages that concerned their planned attack on Wake Island. Such a scenario could be arranged to cause one dolphin to send a message to another, such as "the portal to the other pool is open." If this message could be recorded on multiple relevant occasions, some of the syntax and words in the message could be defined.

The initial attempt in this work will be:
1. To adopt navigation as the initial basis for dolphin language as proposed by Kazakov & Bartlett.
2. To ignore the dictum of Wenger and assume the language includes nouns (and particularly personal nouns such as signature whistles).
3. To define a sufficiently restricted scenario that would be likely to contain a single sentence that would consist of a minimum number of grammatical variances and would be frequently repeated in the wild.
4. To design an experiment that would acquire sufficient communications records to allow a statistical search for this sentence.

Several potential scenarios can be developed through thought experiments.

Scenario A. This scenario will assume:
1. The presence of two dolphins in an estuary separated sufficiently that they cannot see each other but are close enough to communicate orally, and that they have not participated in any recent auditory communications.
2. The estuary exhibits a net positive flow toward the ocean even in the presence of tidal action.
3. The desire of the first dolphin to determine the location of the second dolphin and to proceed to arrange a rendevous.

Even this scenario involves a variety of options. The first dolphin can wait for the second dolphin to utter an irrelevant sentence or to issue its signature whistle, or it can call to the second dolphin to utter a relevant sentence.

Scenario B. This scenario will assume:
1. A first dolphin and a second dolphin from different pods meet in open ocean.

Scenario X. A potential scenario encountered by Jack Kassewitz.

Kassewitz has reported that at least one of his dolphins in Mexico has recently been observed uttering the same sound following a human whistle releasing them from a common acrobatic maneuver ***on multiple occasions***. This may be some sort of "OK" type of declarative statement, possibly like the "good day" used by flight controllers to

---

[178]Verboom, W. (1992) Bio-acoustics: standardization, reference levels and data notation for underwater sound measurement *In* Thomas, J. Kastelein, R. & Supin, A. *eds*. Marine Mammal Sensory Systems. NY: Plenum page 741+

terminate communications with a pilot. Since the maneuver involved a pair of dolphins, it may have been some sort of comment to its partner. Efforts are under way to check previous recordings and observe future exercises that might involve this same type of utterance. An effort will also be made to correlate this behavior with similar sounds from other dolphins (even within the potential range supported by different dialects).

U.5.1.5 Adopting a Program Management software compatible with the biological community

Wikipedia lists a great many program management and project management softwares that could be adopted for an effort of this size. However, most of the program management software are immature, not widely accepted or will disappear within the next few years. It is probably better to select a cloud based software available through a well established vendor. The project management software focuses on the more traditional activities associated with programs of the type envisioned here, including a focus on the Ghantt Chart, critical path management (CPM) and the rather archaic "Program Evaluation and Review Technique" (PERT). The Program Management Institute (PMI) publishes A Guide to the Project Management Body of Knowledge (PMBOK Guide), which describes project management practices that are common to "most projects, most of the time." The Guide is available in at least seven languages. While a membership organization, the fees are quite modest. The International Organization for Standardization (ISO) committee for project management, PC236, has completed its work with the publication of ISO 21500 as an International Standard.

U.5.1.5.1 The traditional approach

A traditional phased approach identifies a sequence of steps to be completed. In the "traditional approach", five developmental components of a project can be distinguished (four stages plus control):

Typical development phases of an engineering project
1.initiation
2.planning and design (The focus of this section)
3.execution and construction
4.monitoring and controlling systems
5.completion
Not all projects will have every stage, as projects can be terminated before they reach completion. Some projects do not follow a structured planning and/or monitoring process. And some projects will go through steps 2, 3 and 4 multiple times.

The proposed program lends itself to division into a number of line items within a Ghantt Chart.

The examples shown on the Wikipedia page, Project Management are clearly designed primarily for pedagogy and fail to highlight the importance of defining multiple task paths occurring simultaneously and interacting at a variety of points along the time line. A skeleton example of a more realistic Ghantt Chart will be developed below.

U.5.1.5.2 Suggested Ghantt Chart line items

- Description of a null hypothesis of maximum scope regarding a potential dolphin language protocol
    - Establish a set of guidelines supporting the most likely path(s) to establishing a dolphin language protocol
    - Coordinate definitions of voiced, unvoiced and whistle elements of the full symbol set
- Additional data mining to establish likely phoneme and morpheme designations
- Initial Experimental Phoneme symbol set determination
    - Establishment of a SAMPA phonetic alphabet for the dolphin symbol set
    - Establishment of a traceable database of phonemes & morphemes and the source of the entry
    - Coordination of mining and initial phoneme symbol set determination activity
- Data entry into an archive including the traceable database
- Development of individual frameworks for the elements of a language protocol
    - Assembly of an initial phonetic dictionary
    - Assembly of an initial descriptive syntax
    - Assembly of an initial descriptive semantics
    - Etc.
- Testing of the initial language protocol

- Establish a coordinated set of workshops/coordination meetings among community members
    - Revise the Master Ghantt Chart on a periodic basis
    - Establish new project elements as required and opportunities arise

U.5.1.6 The Null Hypotheses related to the dolphin language

It is necessary to promulgate a null hypothesis regarding the communications abilities of the dolphin when undertaking a task of this scope. It can be approached from a variety of perspectives, functional, operational, etc. and supported by several potential corollaries. To add specificity, it is also possible to focus the hypothesis on a specific species.

U.5.1.6.1 The Operational Null Hypothesis

Operational null hypothesis: The bottlenose dolphin Tursiops truncatus exhibits a fully functional, but minimally programmed, ability to converse at birth which can be expanded to a capability equivalent to at least a 3 year old human during its period of maturation.

Corollary one: The bottlenose dolphin is able to hear its mothers vocalizations with full fidelity prior to birth due to the unique fluid environment they live in.

Corollary two: The bottlenose dolphin is able to generate a variety of vocalizations at birth and achieve organized vocalizations associated with conversation within one hour of its birth.

Corollary three: The extraordinary ability of this species to mimic may obscure its actual vocal ability in response to its own cognitive activity during its pre-adolescent life.

Corollary four: The bottlenose dolphin is capable of elementary echolocation functions within days of its birth.

U.5.1.6.2 The Functional Null Hypothesis

Functional null hypothesis: The bottlenose dolphin is capable of conversation at an elementary level with its mother within hours to days of its birth and at an expanded level of structure within a few weeks.

Corollary one: The bottlenose dolphin exhibits a minimal language protocol at birth.

Corollary two: The bottlenose dolphin exhibits an identifiable language protocol, including lexical, syntax and semantic elements within six months of its birth.

Corollary three: The bottlenose dolphin is capable at maturity of employing recursive structures during its cognitive, vocalization and information retrieval activities.

U.5.1.6.3 The Structural Null Hypothesis

Structural null hypothesis related to language: The bottlenose dolphin converses with other conspecifics vocally using a sequential nonuniform code of weighted symbols that is instantaneously decipherable due to breaks between words.

Corollary one: The weighted symbols are drawn from a tonal bouquet of sounds including both long relatively steady whistles and short relatively abrupt voiced and unvoiced barks with a frequency range of 2 to 40 kHz.

Corollary two: The symbols used by the dolphin in conversation vary in duration with the emotional state and the cultural environment of the dolphin.

Corollary three: The specific spectral content of dolphin symbology varies among cohorts resulting in identifiable dialects.

Corollary four: The most frequent syntax of the dolphin is subject-verb-object based on **Section U.4.7.7**

Corollary five: The dolphin employs a language containing both tonal and amplitude stresses.

U.5.1.6.4 The Linguistic Null Hypothesis

Linguistic null hypothesis related to language: An animal species will develop an acoustics based communications capability compatible with its vocalization and hearing capabilities, its cognitive capabilities and its environmental needs.

Corollary one: The capability will arise spontaneously among conspecifics.

Corollary two: The capability will involve a language protocol that is extensible within the cognitive capability of the species.

Corollary three: The development of distinct languages, or major dialects, differing between isolated cohorts is to be expected.

The spontaneity of language development among isolated human children has been observed and reported repeatedly. The extent of these isolated spontaneous languages appears to be limited by the exposure of the children to more highly developed parallel languages.

The language capabilities of the bottlenose dolphin, as currently understood, appear to have advanced to that equivalent to that of human children of less than four years of age. It may be more advanced but the ability of humans to interpret the extent of the bottlenose dolphin is hampered by a lack of understanding of the symbol set currently in use.

End of the current plan, backup material below here

U.5.2 Review of terminology and pertinent elements of the language protocol framework

To do this, it is convenient to employ the following modified terms.

cipher alphabet– the exhaustive list of phonemes employed by Cetacea or to a species of Cetacea as the basis for their communications. Nominally the phonetic alphabet of the species.

Code– a large collection of words phrases letters and syllables used to substitute for the equivalent plain text message elements.

Code versus cipher– A more penetrating and useful distinction is that code operates on linguistic entities, dividing its raw material into meaningful elements like words and syllables, where as a cipher does not—a cipher will split the *t* from the *h* in *the,* for example. From about 1400 to 1850, a mixed system was common. It usually had a separate cipher alphabet with homophones and a codelike list of names, words, and syllables. This list, originally just of names, gave the system its name: **nomenclator**.

Complexity– Used variously and generally without precise definition. Examples from McCowan 1999;
- **Complexity** of the signaling system given by the product of bandwidth timed dynamic range.
- **Informational complexity**, represented by the values of individual entropic orders within a language.
- **Organizational complexity** of vocal repertoires across a diversity of species.
- **Social complexity** of a community of conspecifics.

Contoured Language– A language that has contours in its frequency components that are distinctive, rising and falling patterns of pitch, tone, or stress that affect meaning and understanding of this language.

Cryptographic messaging– the acoustic sequences carry the plain text message in unconcealed form that humans are able to record, given suitable broadband recording equipment, but are currently unable to understand.

Cryptogram– The final message as projected by the source, through the intermediate media to the planned receiver.

Fricatives– Unvoiced consonants involving the lips.

Fusional Language– A language in which one form of a morpheme can simultaneously encode several meanings. A morpheme is the smallest meaningful unit in the grammar of a language.

Homophones– acoustic phonemes that are equivalent references to the same plain text signaling element.

Key–Most ciphers employ a key, which specifies such things as the arrangement of letters within a cipher alphabet, or the pattern of shuffling in a transposition, or the settings on a cipher machine. No equivalent of a key will be needed in this initial analysis.

Nulls– signaling elements that have no meaning (and typically introduced to confuse the cryptoanalyst).

Personal dolphin signature– A phonetic signature, generally consisting of more than a whistle, used by a dolphin (at least within *Tursiops truncatus)* as a personal identifier. See signature whistle.

Phoneme– The smallest phonetic unit in a language that is capable of conveying a distinction in meaning, as the m of mat and the b of bat in English.

Plain text– is the message that will be expressed in acoustic form within the aquatic environment. The goal here is to provide an equivalent of this plain text message in written English that can be used to understand what the animals said and potentially prepare responses that can be communicated back to the animal in a form resembling conversation.

Redundancy– Used in plain text messages to provide a degree of protection against external noise by the sender and to provide a forward error correction capability by the receiver.

Signature whistle– historically described as individually distinctive and categorically different whistle types, See Abstract of McCowan & Reiss, 2001.

Transposition systems versus **substitution systems–** In cryptography, a transposition system merely rearranges the elements of a message (at the phonetic,syllabic or word level) or the equivalent terms in a written equivalent. A substitution system can be much more complicated as it expresses an idea in a totally different signaling form than that of the original plain text.

Unvoiced sounds– Vocal sounds with an internal random structure associated with noise.

Voiced sounds– Vocal sounds with a quasi-harmonic structure.

U.5.2.1 Definition of phonemes, lexonemes & morphemes

McCowan & Reiss (1995b) identified 62 whistles associated with neonate (1-4 months) dolphins. With more detailed analysis as appropriate, these whistles might be associated with finite duration tones (potentially vowel sounds). Little focus on potential consonants was shown in this paper.

Janik et al. took major exception to the McCowan & Reiss approach, indicating the use of PCA analysis involved too many hidden parameters in that procedure and showed that it did not provide results as accurate as visual sorting of the candidate spectra by humans.

The PCA technique as used at this time eliminates many subtleties that may be critical to identifying the phonemes of any dolphin language (such as non-harmonic perturbations during construction of a morpheme. In addition, PCA does not lend itself to cross-correlation of two formant phonemes.

To avoid confusion, several authors have used the terms lexeme and lexonemes to describe a fundamental unit (word) of the lexicon and the fundamental symbol (alphabetic letter)of a written language respectively. The contrast is with a fundamental morpheme (syllable) and the fundamental phoneme (individual phonetic symbol representing a sound) of the spoken language. The phoneme is also defined as the smallest element with a distinct meaning within a morpheme.

An aspect of human language that probably cannot be explored at this time among dolphins is the assignment of multiple meanings to a single morpheme within a language, thereby making it a *fusional language*.

Flanagan has addressed the obvious gender related differences in the phonemes and morphemes in human speech (page 1530. Such differences can be anticipated in dolphin speech.

Detailed interpretation of phonemes and morphemes requires close attention to the filter width used in the spectrographic analyzer displaying the waveforms (**Section 8.1.4**).

U.5.2.1.1 Initial description of obvious phonemes

McCowan & Reiss identify a type 1 whistle among neonates and a separate type 1 whistle among mothers with neonates. Based on the experience of this author, the type 1 neonate call relates to the typically emotional "Mama where are you, I am lost!" and the adult whistle relates to a similar state of anxiety, "Child, where are you, I do not see or hear you?"

From a broader perspective, the phonemes of any dolphin language must account for the many voiced, unvoiced and whistle sounds commonly recorded in dolphin spectrograms. Whistles alone are not likely to satisfy the needs of dolphins to express themselves vocally. It is also important that the ability of the dolphin to create more complex sound patterns than humans can based on their more complex sound generating system.

The option must also be maintained that significant phonemes in dolphin language may consist of multiple pairs of formants.

U.5.2.1.2 Initial description of obvious morphemes

Vocalization as used in dolphin conversation (like in human conversation) involves a continuum. It can be subdivided artificially into distinct intervals based on various spectral features; however, a large amount of "noise" accompanies such delineations. The problem is particularly difficult when exploring consonants. It isexhibited by the loose definitions of terms and the resultant overlap in cluster data employed in this field. Once subdivided, the subdivisions are frequently perceived as aggregated into smaller groups. Studdert-Kennedy discussed this situation in Lass (1976, Chapter 8). They note the difficulty in determining whether humans can, in fact, perceive the differences suggest by the symbology of the International Phonetic Alphabet. When discussing consonants, their discrimination has been labeled categorical perception (perception by categorization). It is generally found that observers can *discriminate* between phonemes better than they can *identify* an isolated phoneme. This is the common problem found within virtually all sensory modalities. When discussing vowels, the same problem has been labeled continuous perception. These difficulties are compounded when attempting to determine what another species perceives as a distinct phonetic symbol.

Shoup & Pfeifer, also writing in Lass (1976, Chapter 6) noted another difficulty. When exploring the secondary structure of various human consonants, they discovered certain labials (uttered with the participation of one or both lips) contained no high-frequency peak below 10 kHz. The labials exhibit similar characteristics to the whistles of dolphins (formed primarily from laminar air flowing through the larynx resonating in conjunction with soft tissue). This fact would suggest some dolphin "labials" might not contain any high-frequency peaks below 20 kHz. How such terms as alveolar, labial, palatal, sibilant and fricative are defined with respect to the pneumatic system of dolphins has not been found in the literature.

It is interesting to speculate as to whether the bubbles occasionally released by dolphins at play are closely associated with high frequency sounds generated as "labials" with the blowhole acting like the lips of primates.

As illustrated in **Section 4.9.1**, Shoup & Pfeifer note the difficulty in determining the boundary between many phonemes and the value of examining both a spectrogram and simultaneous amplitude versus time oscillogram. Thye also note the variation in spectrograms based on the bandwidth of the analyzing filter. They provide several examples to aid in defining these boundaries within a morpheme (or word as the case may be). The obvious advantage of isolating a vowel between two unvoiced consonants is clear. On page 203, they summarize their definitions of a wide variety of initial and final plosives.

The option of dolphin morphemes exhibiting multiple phonemes originating from separate sound generators, with potential overlap in start and stop times, must be considered at this stage of our knowledge.

U.5.2.2 Critical importance of adequate bandwidth in the recording equipment

Satisfying the conditions of Oswald et al[179]. for adequate recording bandwidth when collecting dolphin vocalizations is critically important. Many field investigators are not familiar with the significant change in the character of band limited recordings compared to identical recordings at full bandwidth. Some of the changes are subtle. Oswald et al. observed a variety of dolphin species in a large area of the eastern tropical Pacific ocean. In their abstract, they note, "The whistle repertoires of all four species contained fundamental frequencies extending above 20 kHz. Overall correct classification using discriminant function analysis ranged from 30% for the 20-kHz upper frequency limit data to 37% for the 40-kHz upper frequency limit data." They used eight variable in their discriminant set. Their figure 3 clearly shows the failure of a 18-kHz upper frequency limit to capture the characteristic signals of the striped dolphin. Even at 25 kHz, the upper formant of the spectrogram was obviously truncated. The caption for their figure 2 shows obvious truncation of the upper formants of the spotted dolphin vocalizations using a 45 kHz recording capability. Clearly a higher upper frequency limit or a more complex set of discriminants is required to achieve high classification success. Their method of determining a species from a set of variables was based on single vocalization analysis and they note some means of assembling multiple vocalizations from a given animal is desirable.

They note the use of a DAT recorder (without defining the term, "digital audio tape") sampling at its upper value of 48,000 kilo-samples/second only faithfully captures vocalizations of less than 24 kHz. It refers to a standard digital recorder adequate for satisfying the majority of listeners qualitative requirements for fidelity in recording *human speech and music*. The limit is not based on a scientific criteria applicable to dolphin research.

---

[179]Oswald, J. Rankin, S. & Barlow, J. (2004) The effect of recording and analysis bandwidth on acoustic identification of delphinid species *JASA* vol 116(5), pp 3178-3185

Many features of dolphin speech (see **Figure U.5.1-1**) may not be captured using inadequate recording equipment. Markov (1989, pp 617-619) address how the character of dolphin speech can differ with the emotional state of the animals. These differences may not be captured using inadequate recording equipment.

U.5.3 Implementation of the plan

McCowan & Reiss and colleagues have produced the largest amount of statistically significant information concerning dolphin communications beginning in 1993 with a paper by Reiss & McCowan[180]. Their focus is on learning but considerable information concerning the potential for a language protocol is included. The importance and bulk of this information cannot be understated. The significance to these studies as a group in identifying a potential language protocol among groups of dolphins is considerable.

The mechanisms of articulation used in speech do not lead to distinct separation between phonemes. A major requirement placed on the plan is to develop a selected set of phonemes; consonants, vowels, whistles and pulse-sounds that provide maximum delineation between these phonemes within the continuum of vocalization. Such delineation should be visible on both spectrographic and amplitude versus time displays.

U.5.3.1 Detailed plan of attack

By reviewing Yablom & Yablom, the steps in the cryptanalysis of an unknown messaging system can be described. Knowing that the purpose of any voice communications system is to achieve the highest entropy possible within a set of operational and functional constraints, the initial steps are to determine the type of coding used from the most global parameters first. The first choice is between the ideal situation, a block code system and the more common sequential code system. There is no indication that a block code is used by dolphins and the necessity of using a complex translation table that does not accommodate a minimal functional capability for a block code rules in favor of a sequential code. Among sequential codes, the choice is between a uniform code versus a nonuniform code employing a space between words. On examination of dolphin signals, it is clear that a nonuniform code appears to be in use. An additional choice is whether the dolphin might be using an equal frequency of occurrence code or a weighted code. Observation of dolphin signals would strongly suggest the dolphin is using a weighted code much like that used by humans. With the above choices, the remaining challenges are to discover the number of signaling symbols used and their gross characterization. Reviewing long intervals of communications using spectrographic techniques, it is clear the dolphin system of communications employs a variety of symbols including barks, brays and whistles. These appear to be analogous to consonants and vowels in human messaging, along with a variety of more specialized forms. All of these choices would lead to the conclusion, as illustrated by Herzing, that dolphin communications employ virtually the same mechanisms and signaling parameters as human speech except for the frequency range and medium employed.

Measurements clearly show that dolphins use a frequency range of about 2 to 40 kHz for purposes of communications and an even higher band, extending up to 150 kHz for echolocation. Within the 2-40 kHz band, the signals generated by dolphins look remarkably similar to those used by humans in their more restrictive operating band of 0.2 to 20 kHz. The challenge is to cryptanalyze a weighted sequential signaling environment using a 2-40 kHz signaling band and an unknown number of phonetic symbols.

There are a wealth of recordings of dolphin signals within a narrower 2-20 kHz band, but these may not adequately capture the essence or subtleties of dolphin signaling. Since 2000, more and more full bandwidth recordings of dolphin signals are becoming available. They clearly show a rich symbol set that includes much more than just whistles; they show a variety of voiced and unvoiced symbols suggesting a broad list of morphemes formed of multiple phonemes, i.e., barks and whistles.

Kassewitz et al. have specialized in the recording of broadband dolphin signals, up to 50 kHz for communications and up to 150 kHz for echolocation signals.

While it may not be necessary to record the higher frequencies in dolphins, it is necessary to establish academic credibility. The price for failing to appreciate the signals at high frequencies can be high.

> During World War II, the German and Japanese navies lost a large number of submarines to aerial attach because they were unable to detect the high frequency radars used by the Allied air forces.

> A recent United States Navy program recorded bottlenose dolphin signals at up to 390 kHz to be sure they

---

[180]Reiss, D. & McCowan, B. (1993) Spontaneous vocal mimicry and production by bottlenose dolphins (Tursiops truncatus): evidence for vocal learning. *J Comp Psych* vol 107, pp 301–312
312.

> were not missing significant signals generated by the dolphin[181]. They demonstrated significant energy generated in the 394 kHz range during echolocation activities. "Data analysis shows wide transmitting beam patterns at frequencies lower than 135 kHz contain a majority of the energy in the echolocation signal, agreeing with previously documented work. However, further analysis shows significant energy at higher frequencies. Early in the experiment, the dolphin steered narrow high frequency signals and adjusted the energy content in those different frequencies while scanning the target."

At a next level of complexity, Yip has provided a high level list of tests that need to be incorporated in any exploration of dolphin language[182]. He notes that many of his suggested investigations "that contribute to human phonology remains unexplored in other species."

U.5.3.2 Traffic analysis

There appears to be little argument concerning the ability of *Cetacea* to carry out long range two-way communications (conversation). Whales employ this capability over thousands of kilometers by using low attenuation acoustic ducts present in deep ocean waters typically associated with thermoclines. Marine dolphins appear to use it effectively in shallow waters when splitting into two or more pods and exploring the opposite sides of estuaries. They appear to nearly continuously report their individual positions and presumably details related to the local environment. In both cases, these conversations employ many unique signal sequences. If the signal sequences are repetitious routines, human observers have not been able to demonstrate that fact.

U.5.3.3 Naming conventions

> It is noteworthy that the names given to specific dolphins, and used extensively by their trainers (human associates) in trans-species communications are not limited to a single whistle. In fact many are two syllable words with each syllable beginning with a hard sound and ending with a vowel

Currently, the academic community has focused on a small group of signature whistles that is inadequate for supporting a large group of dolphins when fusing (such as occurs routinely in Shark Bay, Australia). Several researchers at that location have reported dolphins changing their signature sequence when interfacing with other cohorts at that location. This would be the equivalent of humans adding last names during the 17th through 19th Centuries as the population became more mobile. These last names frequently indicated parentage (Peterson or Ericksdoter) or original location (Gorbachev or Fulton).

Killebrew et al. recorded a variety of signals from neonate dolphins in 2001 using commercial equipment (with a catalog passband of 17 kHz at ±3 dB. Kassewitz recorded the first hours of life of a bottle-nosed dolphin using wide band sound equipment in 2007. He demonstrated that the neonate began using a name very similar (if not identical to) its mother's name immediately but soon modified it to represent its own permanent name. The new name is believed to have reflected its gender. Since the neonate was not likely to recognize its gender at that time, it is highly likely that the mother instructed the neonate to change its utterances to reflect its new name. Such a conversation between the neonate and its mother would appear to incorporate many signals reflecting the rules of a language, at least the concepts of a personal noun and/or pronoun and a verb (use this name as your permanent name).

- - - - -

Killebrew et al. also note the work of McCowan & Reiss[183] in the recording of eight neonates. The McCowan & Reiss paper is extensive and includes many illustrative waveforms. However, most of the waveforms have been drawn as idealized signals rather than spectrograms, or included in tabulations. "Whistles from captive-born infant bottlenose dolphins (*Tursiops truncatus*) were recorded during normal social interactions from birth over the 1st year of development and analyzed for acoustic structure, use, and context. Results indicate that 2 predominant whistle types (type 1, a rising ramp tone, and type 2, a rising ramp with a positive curvature) were shared by all the infants across all social groups." The values of their calculated coefficients of frequency modulation were not given in the paper. While their data set expanded to 128 waveforms, they continued to consider only the baseline tone of the

---

[181]Lemerande, T. (2002) Transmitting Beam Patterns of the Atlantic Bottlenose Dolphin (Tursiops Truncatus): Investigations in the Existence and Use of High Frequency Components Found in Echolocation Signals. Monterey, CA: Naval Postgraduate School

[182]Yip, M. (2006a) The search for phonology in other species *Trends Cogn Sci* vol 10, pp 442-446

[183]McCowan, B. & Reiss, D. (1995b) Whistle contour development in captive-born infant bottlenose dolphins (*Tursiop truncatus*): Role of learning *J Comp Psych* vol 109, pp 242-260

overall waveforms. The collection of waveforms included more complex modulated tones, including some consisting of baseline tones modulated by three to five cycles of a sinusoid. Also included were a variety of baseline tones modulated by an isolated pulse (positive or negative going) somewhere during the duration of the waveform. It is difficult to believe these complex waveforms did not have individual meaning within dolphin communications.

The results of these activities can be compared to the time lines developed by Frederici in **Section S.2.3.1** of Appendix S.

- - - -

More recently[184], McCowan & Reiss have added more fuel to the discussion of whether dolphins exhibit "signature whistles." They were specifically studying the whistles (only) of isolated dolphins. After denigrating the concept of signature whistle in their abstract, in the form previously proposed by others, they focused on one whistle in their documented repertoire. They conclude,

> "Results on the further classification of this predominant shared whistle type indicated that 14 subtle variations within the one whistle type could be partially attributed to individual identity."

However, in their text they modify this position,

> "Results also revealed that social familiarity, social group membership and capture site location hava a strong influence on spectral features of whistle acoustic structure. Therefore, this variability may be one mechanism contributing to 'regional' dialects in dolphins. Such acoustic variability within whistle types could also contribute to individual recognition."

While McCowan & Reiss (2001) turn away from the previously defined concept of signature whistles among conspecifics under conditions of isolation, they do not address the more complex situation of combinations of fricatives, burst-pulse elements and whistles under conditions of a shared social environment. Simultaneously, they suggest a potential dialectal aspect to a natural dolphin language protocol.

U.5.3.4 Secondary considerations

The hearing apparatus of Cetacea are designed to operate in a fluid environment. As such, a fetus of this family is able to hear effectively as soon as its auditory modality is functional. It becomes familiar with its mother's name at that time and generally is born uttering that name. According to Kassewitz et al., the neonate begins repeating its mother's name within an hour of its birth, but soon (apparently within hours) modifies that call and uses it to identify itself. The adopted identifier is believed to be gender specific. While dolphins are noted for their ability to mimic other dolphins, it is unlikely their vocalization mechanisms are sufficiently mature to mimic their mother's name with any precision at birth.

As in human communications, the use of a last (or formal) name is probably not used until a dolphin comes into contact with a larger community where more specific identification becomes useful. In Scotland, last names did not come into common use until large migrations became common in the 16th Century. Children do not commonly use their last names until they are introduced to larger communities as they begin schooling. When a large migration occurred from Fultounland (literally the low lands between current day Glasgow and the Clyde Estuary) toward the maritime provinces along the Irish Sea, the last name of Fulton evolved and became common. Turner reports a similar occurrence among the dolphins of Shark Bay when they intermingle with other groups from remote areas of that estuary or when nomadic dolphins enter the Bay.

U.5.5 Currently established characteristics of dolphin signal sequences

Some progress has already been made in establishing the characteristics of the voiced, unvoiced and whistle portions of the overall dolphin phonetic symbol set. As noted previously, dolphins employ a broad range of signal sequences that vary due to both the emotional state of the source animal, the local environment and the character of nearby members of the cohort. If they exhibit a language protocol, it is clearly tonal like Chinese (as opposed to atonal like English). The abrupt start of some of the tonal whistles would suggest a fricative type sound preceding the whistle.

When hunting in a group or when known by themselves to be alone, dolphins generate minimal signals that might be associated with conversation.

---

[184]Mccowan, B. & Reiss, D. (2001) The fallacy of 'signature whistles' in bottlenose dolphins: a comparative perspective of 'signature information' in animal vocalizations *Anim Behav* vol 62, pp 1151-1162

U.5.5.1 Background

It does not appear that enough similar signal sequences have been studied to identify any redundancies or homophones. Redundancies are generally introduced by sources to minimize the impact of extraneous noise on any conversation. Similarly, homophones, if used, provide a degree of forward error correction when messages are interpreted by receivers. [what about additional flexibility?]

U.5.5.2 Phonemes & morphemes available to the dolphin EMPTY

U.5.5.3 lexicon, syntax , semantics and pragmatics available to the dolphin

Section IV of Marino et al. (2008) has summarized the ability of the bottlenose dolphin to interpret correctly complex syntactic and semantic structures.

U.5.5.3.1 Language protocol elements involved in memory evaluation

While section IV(1) of Marino et al. is focused on memory and learning, it involves interspecies communications as a vehicle. As such, a variety of language protocol elements are involved. They have noted (and provided citations),

- the ability of dolphins to recall vocabulary items and syntactic rules for months without intermediate practice.
- the ability of the dolphin to identify a probe sound with any of a group of up to eight sounds previously provided in a sequence.
- the ability of the dolphin to remember the spatial location of named structures in an extensive habitat.
- the ability of the dolphin to define and later identify the individuals in one's immediate community.
- Dolphins understand representations of the real world (tv images of their trainer giving commands)
- Dolphins learn and master not only the semantic features, but also the syntactic features of artificial gestural and acoustic languages.
- Dolphins can learn more complex syntactic structures through inference rather than explicit instruction.
- Dolphins are one of the few species that can imitate both arbitrary sounds and arbitrary physical behaviors.
- Dolphins can reliably discriminate between a higher or lower in pitch than a reference tone.

Section IV(2) of Marino et al. focuses on vocalization and communication. They note, 'bottlenose dolphins are adept vocal learners, a trait rare among mammals." They also note,

- "There is evidence for individual-level variation in the whistle repertoires of dolphins."
- "Bottlenose dolphins produce individually distinctive whistles that they apparently use to identify conspecifics and may also be employed as a cohesion call."
- "Non-whistle 'bray' calls have been associated with salmon capture off eastern Scotland."
- "Non-whistle 'pop' calls have been associated with male aggression in Shark Bay, western Australia."
- "there is some evidence that the sequential order of whistles is an important feature of their communications."
- Dolphins are known to mimic a sound one octave higher or lower if the result is more within its preferred vocal range.
- The stages in the development of an individuals repertoire parallels that of humans and song birds.

- Among the killer whales, their sounds are so specific that other marine mammals have learned to recognize individuals.

These observations clearly describe at least the basics of the language protocol capability of dolphins and potentially other species of *Cetacea*.

U.5.6 Initial cribs for the putative the bottlenose dolphin language

A key element in attempting to decipher an unknown symbol set and unknown language is known as a crib in cryptanalysis. A few initial cribs have appeared based on the above analyses and the available literature.

Crib #1–The ability of the dolphin to create sounds equivalent to both human vowels and consonants as well as pure whistles strongly implies that what are frequently described as "signature whistles" by dolphin observers must be considered more complex symbol sequences. Each signature whistle can be defined more explicitly as a *signature whistle sequence* or more properly a *signature sequence* containing combinations of all of the available symbol types. Such a definition would be compatible with common human personal naming conventions. This interpretation allows the signals recorded for dolphins under emotional stress by Caldwell & Caldwell (1968) to be considered in a broader context than just a series of signature whistles.

Crib #2– The ability of the dolphin to employ multiple independent acoustic signal generators simultaneously and independently make cryptanalysis of their speech more complicated than the single source signaling associated with humans.

Crib #3–The auditory and ethological observations of Conner & Smolker at Shark Bay and reported in 1996 provide strong evidence that their "pop" is a single symbol phoneme that is used in strings to communicate with other dolphins, primarily addressed to females by herding males (typically pairs of males in alliance). The meaning of these strings appears clear with only the emphasis requiring further definition;

- 1 or 2 pops, "attention," in the vernacular, "hey" or "hey you" or "hey, look at me and give me your attention."
- 2 to 5 pops, "come to my location,"
  - in the vernacular of herding males to a female, "come to my location and follow my guidance!!"
  - in the vernacular of a female to her child, "come back to my location now!!"
- 4 to 10 pops, "stay in my immediate location,"
  - in the vernacular of herding males, "stay near my location, follow my guidance and do not attempt to escape!!"

Their comments on page 657 under the heading "The function of pops" are in accordance with this crib.

Crib #4–Credible reports that are both anecdotal and in the science literature suggest that mature dolphins change their signature sequence when interfacing with unfamiliar cohorts. By recording the animal's normal signature sequence and the sequence used in such encounters, significant clues concerning the meaning of the changes should be obtainable. The potential changes could indicate where the animal came from, what larger family it belonged to, or other information. By repeating this test sequence for different animals placed in similar situations, the pattern of signature sequence changes may provide important information about the structure of dolphin language. Changes at a specific location in the signature sequence of multiple animals of one cohort would be particularly valuable to the cryptanalyst.

Crib #5– [Reserved] xxx

U.5.7 Establishing a computer-implemented method for synthesizing dolphin speech

The Goldsmith team has provided a roadmap for generating speech in any language using a prosodic database[185]. It appears to be applicable to the generation of acoustic vocalizations by computer that would be understandable to the dolphin. The lexicon of the complete database could be established based on the experimentation carried out within the scope of the plan defined in this work. The prosodic component could then be established based on additional analyses of the vocalizations of different cohorts of dolphins

[185]Huang, X. Adcock, J. & Goldsmith, J. (1999) Prosodic databases holding fundamental frequency templates for use in speech synthesis US Patent 5905972A

Table of Contents, June 12, 2014

List of Figures

(Active) SUBJECT INDEX (using advanced indexing option)