



Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires

BRENDA McCOWAN*†, SEAN F. HANSER† & LAURANCE R. DOYLE‡

*California Regional Primate Research Center, University of California at Davis

†Marine World Foundation

‡SETI Institute, NASA Ames Research Center

(Received 29 August 1997; initial acceptance 10 December 1997;
final acceptance 21 July 1998; MS. number: A8015)

Comparative analysis of nonhuman animal communication systems and their complexity, particularly in comparison to human language, has been generally hampered by both a lack of sufficiently extensive data sets and appropriate analytic tools. Information theory measures provide an important quantitative tool for examining and comparing communication systems across species. In this paper we use the original application of information theory, that of statistical examination of a communication system's structure and organization. As an example of the utility of information theory to the analysis of animal communication systems, we applied a series of information theory statistics to a statistically categorized set of bottlenose dolphin, *Tursiops truncatus*, whistle vocalizations. First, we use the first-order entropic relation in a Zipf-type diagram (Zipf 1949, *Human Behavior and the Principle of Least Effort*) to illustrate the application of temporal statistics as comparative indicators of repertoire complexity, and as possible predictive indicators of acquisition/learning in animal vocal repertoires. Second, we illustrate the need for more extensive temporal data sets when examining the higher entropic orders, indicative of higher levels of internal informational structure, of such vocalizations, which could begin to allow the statistical reconstruction of repertoire organization. Third, we propose using 'communication capacity' as a measure of the degree of temporal structure and complexity of statistical correlation, represented by the values of entropic order, as an objective tool for interspecies comparison of communication complexity. In doing so, we introduce a new comparative measure, the slope of Shannon entropies, and illustrate how it potentially can be used to compare the organizational complexity of vocal repertoires across a diversity of species. Finally, we illustrate the nature and predictive application of these higher-order entropies using a preliminary sample of dolphin whistle vocalizations. The purpose of this preliminary report is to re-examine the original application of information theory to the field of animal communication, illustrate its potential utility as a comparative tool for examining the internal informational structure of animal vocal repertoires and their development, and discuss its relationship to behavioural ecology and evolutionary theory.

© 1999 The Association for the Study of Animal Behaviour

The comparative analysis of animal communication systems and their complexity, particularly in comparison to human language, has been generally limited by both a lack of sufficiently extensive data sets and appropriate analytic tools. Information theory measures, in their original application (Shannon 1948; Shannon & Weaver

1949), provide important quantitative tools for examining and comparing communication systems across species. These measures represent statistical information that is imperative for analysing communicative repertoires objectively, and for elucidating repertoires of unknown structure or function (Pea 1979). This statistical characterization is independent of the kind of communication system of interest and thus allows a comparative examination of the complexity of vocal repertoires and their organization, both with respect to human language (Zipf 1949, 1968) and, more generally, in relationship to behavioural ecology and evolutionary theory.

Correspondence: B. McCowan, Behavioral and Neurobiology Unit, California Regional Primate Research Center, University of California, Davis, CA, 95616, U.S.A. (email: bjmccowan@ucdavis.edu). S. Hanser is at the Marine World Foundation, Vallejo, CA, 94589, U.S.A. L. Doyle is at the SETI Institute, NASA Ames Research Center, Moffett Field, CA, 94035, U.S.A.

Information theory has been applied to animal communication by various authors since Shannon & Weaver (1949) established the field. Many applications have been conducted on the communicative interactions among individuals, within or between species, including aggressive displays of hermit crabs (Hazlett & Bossert 1965), aggressive communication in shrimp (Dingle 1969), intermale grasshopper communication (Steinberg & Conant 1974), dragonfly larvae communication (Rowe & Harvey 1985), and Altmann's (1965) study of macaque social communication. Preston (1978) used information theory in a unique analysis of the communication in a symbiotic relationship between a vertebrate, the goby, and two species of shrimp. Less frequent applications of information theory have been conducted on some structural aspects of communication systems including the waggle dance of honeybees (Haldane & Spurway 1954), chemical paths of fire ants (Wilson 1962), Chatfield & Lemon's (1970) brief application to the structure of songs in cardinals and wood pewees, and the sequential analysis of chickadee 'notes' (Hailman et al. 1985, 1987; Hailman & Ficken 1986; Ficken et al. 1994; Hailman 1994). Information theory has also been applied to specialized communication topics such as Beecher's analyses of vocal recognition in Mexican free-tailed bats, *Tadarida brasiliensis mexicana* (Beecher 1989).

All of these studies have brought important insights into the study of animal communication. However, we suggest that information theory can be applied to animal communication systems at the level for which the tool was originally designed: to examine mathematically the channel capacity and the structure and organization of communication systems. Shannon originally illustrated the concepts in information theory using statistically significant samples of human language. As we will demonstrate in this paper, large sample size is essential for the information theoretic approach to realize the full potential of its analytical capabilities. Most past studies have used information theory for modelling small samples of contextually restricted behavioural interactions between individuals within or between species. Due to limitations in sample size, few investigators have been able to apply these measures to sequential signalling by individuals as a means for deciphering the structure and organization of a species' communication system (but see early work by Chatfield & Lemon 1970). Continued advances in bioacoustic technology are beginning to permit the efficient analysis and categorization of large samples of acoustic data to be used in such sequential statistical analyses.

Furthermore, most past research has focused almost exclusively on the use of Shannon's measure for information (Shannon & Weaver 1949). Shannon's entropies examine the information content at increasingly complex levels of signalling organization. As an example from human speech, information content can be evaluated at the phonemic or letter level, the word level, and various levels of sentence organization. Each level can be represented by a series of increasing orders (e.g. zero, first, second, etc.) of entropy. Entropy is defined here as a measure of the informational degree of organization and

is not directly related to the thermodynamic property used in physics (although they do possess a common mathematical ancestor).

Few investigators of animal behaviour have examined the use of first-order entropic analysis known as Zipf's law or statistic (Zipf 1949; see Hailman et al. 1985, 1987; Hailman & Ficken 1986; Ficken et al. 1994; Hailman 1994). Zipf's statistic examines a first-order entropic relation and evaluates the signal composition or 'structure' of a repertoire by examining the frequency of use of signals in relationship to their ranks (i.e. first, second, third versus most-to-least frequent). It measures the potential capacity for information transfer at the repertoire level by examining the 'optimal' amount of diversity and redundancy necessary for communication transfer across a 'noisy' channel (i.e. all complex audio signals will require some redundancy). In comparison, Shannon entropies were originally developed to measure channel capacity, and his first-order entropy is conceptually and mathematically related to Zipf's statistic. However, Shannon's first-order entropy differs from Zipf's statistic in so far as Zipf did not specifically recognize language as a 'noisy' channel, as Shannon did. Shannon higher-order entropies (second order, third order, and so on) can also provide a more complex examination of communicative repertoires. They can be used to examine the organizational complexity of signal repertoires, for example, at the level of two-signal sequences, three-signal sequences, and so forth. Higher-order entropies provide a quantitative measure of the information content at each successive level and how many levels of informational complexity the communication system contains.

In this paper we return to the original application of information theory, that of statistical (objective) examination of a communication system's structure and organization. As an example of the utility of information theory to the analysis of animal communication systems, we applied a series of information theory statistics to a set of dolphin whistle vocalizations that were categorized quantitatively using the Contour Similarity Technique (McCowan 1995; McCowan & Reiss 1995a, b). Bottlenose dolphins, *Tursiops truncatus*, were chosen as our first animal model because dolphins produce a large and complex whistle repertoire (Dreher 1961, 1966; Dreher & Evans 1964; Burdin et al. 1975; Kaznadzei et al. 1976; Ding et al. 1995; McCowan & Reiss 1995a, b; Moore & Ridgway 1995). Little is known about dolphin whistle repertoire organization and, as a result, the communicative function of dolphin whistles. Whistles and vocal communication, in general, are probably essential to the dolphin's fission-fusion social structure and for negotiating complex social relationships (Wells et al. 1987; Connor et al. 1992).

First, we apply Zipf's statistic (Zipf 1949, 1968) to dolphin whistle vocalizations to illustrate its application as a comparative indicator of the structural complexity of vocal repertoires, as well as a potential indicator of acquisition/learning in animal vocal repertoires. We then demonstrate the need for more extensive data sets when examining Shannon's higher-order entropies (Shannon 1948; Shannon & Weaver 1949) of such vocalizations,

which can begin to allow the statistical reconstruction of repertoire organization. We introduce a new comparative measure, the slope of Shannon entropies, and suggest how it can potentially be used to compare the organizational complexity of vocal repertoires across a diversity of species. Finally, we illustrate the nature and predictive application of these higher-order entropies with our sample of dolphin whistle vocalizations (not yet sufficient for a full analysis of the higher-order entropies). Again, we emphasize the need for very large data sets that conserve the temporal sequences of signals as well as objective classification of such signals if these methods are to be applied correctly.

In summary, the purpose of this preliminary report is to re-examine the original application of information theory (i.e. channel capacity or complexity) and apply it to the field of animal communication (also see Steinberg 1977; Beecher 1989; Dawkins 1995), illustrate its potential utility as a comparative tool for examining the informational structure and organization of animal vocal repertoires and their development, and discuss its potential predictive application in light of behavioural ecology and evolutionary theory.

THEORY

Definitions of Information

Serious confusion in the nomenclature surrounding information theory has prompted us first to define our terms clearly. Most important is the definition of 'information', which can be defined in three seemingly conflicting ways. Perhaps most commonly, information can be defined in the 'contextual' sense, that is, what information is conveyed during a communicative exchange (i.e. meaning). Information also can be defined in the 'communicative' sense, that is, how much information can be transferred during a communicative exchange. Finally, information can be defined in the 'statistical' sense (as Shannon defined it), as the degree to which data are noncompressible (e.g. the word 'meet' can be compressed by 'e' \times 2; the word 'meat', however, cannot be compressed without losing information. Thus, there is more statistical or entropic information in 'meat' than 'meet').

These three definitions of information can result in very different measures of information. For example, one needs to know nothing about 'contextual information' to calculate 'communicative information' or 'statistical information'. Similarly, 'communicative information' differs from 'statistical information'. Statistical information is maximized when sequences of signals are truly random (or maximally diverse) in nature. Communicative information is maximized when a balance is reached between highly redundant or highly diverse signalling; a certain amount of redundancy is necessary for transmission in a 'noisy' communication channel. Therefore, to avoid any confusion, we define 'information' in the present study using Shannon's original use in information theory, that of statistical information, also known as entropy. Information in the 'communicative' sense will

be termed 'communication capacity' in the present study (i.e. complexity of the signalling system). Contextual information is not discussed in this study but we wish to stress that examining measures of communication capacity can lead to insights in contextual information.

Zipf's Law and the Principle of Least Effort

Zipf's law or the distribution of the logarithm (base 10) of the signal rank (i.e. first, second, third . . . most used) plotted against the logarithm of actual frequency of occurrence (i.e. percentage repetition of that signal) for human languages (Zipf 1949, 1968) has been invoked for a number of applications, including recently DNA sequences (Mantegna et al. 1994), with varying interpretations and reliability (Flam 1994; Damashek 1995; Bonhoeffer et al. 1996; Israeloff et al. 1996; Voss 1996). Such a function nevertheless remains a valid indication of both the nonrandomness of a system as well as the potential capacity for communication transfer of such a system. Zipf's law is based on what Zipf (1949) termed the 'Principle of Least Effort' in which he proposes that human speech and language are structured optimally as a result of two opposing forces: unification and diversification. To summarize his argument using a modern interpretation, if a repertoire is too unified or repetitious, a message is represented by only a few signals and, therefore, less communication complexity is conveyed (remembering that we are reserving the word 'information' for Shannon entropy measures). If a repertoire is too diverse or randomly distributed, the same message can be over-represented by a multitude of signals and, again, less communication is conveyed. These two opposing forces result in a 'balance' between unification and diversification. Zipf's Principle of Least Effort can be statistically represented by regressing the log of the rank of signals within a repertoire on the log of their actual frequency of occurrence. Balance is indicated by a regression coefficient (or slope) of -1.00 . Zipf subsequently showed that a multitude of diverse human languages (e.g. English words, Nootka varimorphs and morphemes, Plains Cree holophrases, Dakota words, German stem forms, Chinese characters, and Gothic root morphemes, Gothic words, Aelfric's Old English morphemes and words, Notker's Old High German, Pennsylvania Dutch, English writers from Old English to Present, Old and Middle High German and Yiddish sources, Norwegian writings), whether letters, written words, phonemes, or spoken words, followed this principle and the predicted slope of approximately -1.00 . This balance optimizes the amount of potential communication that can be carried through a channel from speaker to receiver. The structure of the system is neither too repetitive (the extreme would be one signal for all messages) or too diverse (the extreme would be a new signal for each message and, in practice, a randomly distributed repertoire would represent the highest degree of diversity). Thus, a system exhibiting such balance can be said to have a high potential capacity for transferring communication (which we term 'high potential communication capacity'). It only has the 'potential' to carry a high degree of communication, though, because Zipf's

statistic only examines the structural composition of a repertoire, not how that composition is internally organized within the repertoire (i.e. higher-order entropies). However, some vocal repertoires may have a higher 'potential' to carry a substantial degree of communication than others. Therefore, Zipf's statistic can be used as one comparative measure of the structural complexity of repertoires (another comparative measure might be repertoire diversity, a zero-order entropic relation: see below).

Shannon's Higher-order Entropies and Information Content

The Zipf's statistic can measure and compare the structural complexity of vocal repertoires, but it has limitations in evaluating the actual relationship between signals in a dynamic repertoire. For instance, Zipf's statistic cannot examine how signals interact or are internally organized within a communicative repertoire. We need to use higher-order entropic measures to measure signal system internal dynamics, for example, to determine the signal sequences within a repertoire. Zipf's statistic and Shannon entropic orders, however, are mathematically related. The Zipf relation is essentially an examination of the first-order information entropy, which was first defined by Shannon (1948; Shannon & Weaver 1949).

The zero-order entropy, H_0 , is the number of bits of information required to represent a particular sample of different events, that is:

$$H_0 = \log_2 N,$$

where N is the number of different events (i.e. number of letters, words, phonemes, notes, whistles, etc.). Base 2 is used because the 'bit' can be considered the most familiar logarithmic form for an information measure (any other base would do just as well and base 10 is used for the Zipf's statistic).

The first-order entropy takes into account the different probability of occurrence of each event as well, giving:

$$H_1(A) = -p(A_1) \log_2 p(A_1) - p(A_2) \log_2 p(A_2) \dots - p(A_N) \log_2 p(A_N),$$

where $p(A_1)$ is the probability (i.e. frequency of occurrence) of event A_1 , while $p(A_2)$ is the frequency of occurrence of event A_2 , and so on. An English letter analogy is the frequency of occurrence of the letter 'a' in a text, then the letter 'b', and so on. The uncertainties of events $A_1 A_2 \dots A_N$ add, while the probabilities multiply, so that a logarithmic function is appropriate to represent the uncertainties of each event (the logarithmic function allows products to add). Because, for H_1 , the probability of each event $p(A_1) + p(A_2) + \dots + p(A_N)$ is independent, the product of the log of the probability with the probability for each event is the uncertainty of the first-order entropy. The entropy therefore measures information content and varies most quickly for low-likelihood events (for example, being informed that a word ends in the

letter 'e' does not tell one as much as being informed that it begins with the lesser-used letter 'q'). Because probabilities are less than unity, their logarithms are always negative, but informational entropies are positive so that we have the final first-order entropic form of:

$$H_1 = \sum_{j=1}^N -p(A_j) \log_2 p(A_j)$$

The second-order entropy introduces conditional probabilities into the structure of the stream of events being studied (words, letters, phonemes, musical notes, or, in our present context, dolphin whistles). An English letter example would be the probability that the letter 'u' follows the letter 'q', which is very high, while the probability that the letter 'z' follows the letter 'x' would be very low. Thus the second-order entropy for events B_1 through B_N , given that the events A_1 through A_N have occurred, can be defined (after Yaglom & Yaglom 1983) as:

$$\begin{aligned} H_2(AB) = & -p(A_1 B_1) \log_2 (A_1 B_1) - \\ & p(A_1 B_2) \log_2 (A_1 B_2) - \dots - p(A_1 B_N) \log_2 (A_1 B_N) \\ & - p(A_2 B_1) \log_2 (A_2 B_1) - \\ & p(A_2 B_2) \log_2 (A_2 B_2) - \dots - p(A_2 B_N) \log_2 (A_2 B_N) \\ & \dots \\ & - p(A_N B_1) \log_2 (A_N B_1) - \\ & p(A_N B_2) \log_2 (A_N B_2) - \dots - p(A_N B_N) \log_2 (A_N B_N) \end{aligned}$$

Because here we are comparing the internal structure of events with each other, there are the same number of N events for both first occurrences of events A and second occurrences of events B (for an alphabetic letter comparison, for example, $A_1 B_1$ stands for the probability that the first letter of the alphabet will follow itself, $A_1 B_2$ that the second letter of the alphabet will follow the first, and $A_5 B_{10}$ that the tenth letter will follow the fifth, etc.).

Similarly, the third-order entropy of an event includes the conditional probability given the two preceding events, and therefore is defined by the expression:

$$H_3(ABC) = H_2(AB) + H_{AB}(C)$$

where $H_{AB}(C)$ is the entropy of event C given that events A and B have occurred (similar to the equation for the second-order entropy, and so on for the higher-order entropies; see Yaglom & Yaglom 1983).

To define these quantitative measures with respect to the analysis of animal vocal communication, we provide the following summary.

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

Therefore, as higher-order entropies drop significantly in value from one entropic order to the next, less

statistical information (entropy) and more organizational complexity is present at this order of entropy in the communication system (we note, however, that the communication system will also begin to lose complexity if it becomes too redundant). If one plots the slope of the entropic orders, a higher negative slope would indicate less statistical information but increasing importance of sequential organization (and thus higher communication capacity) at increasing entropic orders. For comparison, a truly random sequential system would show a slope of zero and would contain the most statistical information but less communication complexity. Thus, entropic slope can provide a measure of organizational complexity that can be used to compare the importance of sequential order in the communication systems of different species. Similarly, a comparison of slopes across different age groups within species, and the changes in the slopes from infancy to adulthood across species can provide a comparative measure of the developmental process underlying repertoire organization.

METHODS

Dolphin Whistle Data Sets

We categorized infant and adult whistle data sets independently into whistle types using a quantitative technique termed the Contour Similarity Technique which is based on k-means cluster analysis (McCowan 1995). We borrowed the infant data from a subset of infant data from McCowan & Reiss (1995b). We also borrowed a subset of the adult data shown from McCowan & Reiss (1995a). Detailed methods for recording and analysing the dolphin whistles can be found in McCowan & Reiss (1995a, b) and McCowan (1995). These data sets include individually identified whistles from four male infants (<1 month old: $N=23$ whistle types using 53 classified whistles; 2–8 months old: $N=73$ whistle types using 424 classified whistles; 9–12 months old: $N=60$ whistle types using 293 classified whistles) and eight adults (3 males, 5 females; $N=102$ whistle types using 600 classified whistles) from two different captive social groups of dolphins from Marine World Africa U.S.A in Vallejo, California. We excluded whistle type 1, an apparent ‘infant cry’ vocalization (McCowan & Reiss 1995a, 1997), from these analyses because infant cry was not included in the human infant data set for these analyses (Zipf 1949, 1968). It is important to note that these whistle data sets were statistically, not experimentally, categorized. Categorical perception and categorization experiments are necessary to confirm or modify our statistical whistle categories (for a full discussion of this issue see McCowan & Reiss 1995a, b). For whistle sequences, we included only individually identified whistles that occurred more than once in the data set and that also occurred within a sequence ($N=27$ whistle types from 147 sequences of two to nine whistles in length using 493 classified whistles from the adult individuals only). A whistle sequence was defined by an intersequence interval of 1600 ms. Typical interwhistle interval within a sequence was 300 ms.

Zipf Relation

We analysed the whistle data sets described above (McCowan 1995; McCowan & Reiss 1995a, b) from adults and developing infant dolphins for each of three developmental periods by regressing the log rank of whistles against the log of their actual frequency of occurrence, as conducted by Zipf (1949, 1968) for human languages. We compared the slopes generated from regression to those reported for human adults (Zipf 1949, 1968; Ridley & Gonzales 1994), human infants of 22 months of age (youngest age for which frequency of word usage could be obtained; Zipf 1949, 1968), and a randomly generated data set (10 000 randomly generated numbers of 102 types, a number equal to the number of whistle types found in our sample from adult dolphins).

Higher-order Entropies and Markovian Sequence Analysis

The calculations for the entropies for English letters, Arabic letters, Russian letters, Russian phonemes, and octave-music can be found in Yaglom & Yaglom (1983). We calculated the entropies for the dolphin whistles using Uncert statistical software (developed by J. Hailman & E. Hailman). We analysed whistle sequences from adult dolphins using first-order Markovian chain analysis (using the Uncert statistical program). The computer program generated a matrix of the probability that one whistle type immediately followed the same or a second whistle type for each whistle type in the data set. This test is equivalent to generating a second-order Shannon entropy test. For example, this type of Markovian analysis was used in Ficken et al. (1994) to analyse Mexican chickadee, *Parus atricapillus*, ‘note’ sequences.

RESULTS

Zipf’s Relation

Adult dolphin data (Fig. 1b, Table 1) showed a slope of -0.95 when log rank of signal was regressed against log frequency of occurrence indicating that the distribution of whistles in the dolphins’ repertoire is indeed non-random and, in fact, closely matches that found for words in dozens of different human language samples (slope = -1.00 ; Zipf’s slope is shown in Fig. 1a). Infant dolphins less than 1 month of age (Fig. 1c, Table 1) showed a slope (-0.82) closer to that of our randomly generated data (-0.087 ; theoretically a slope of zero; Fig. 1a, Table 1) than to that found for adult dolphins (-0.95). This is strikingly similar to the Zipf slopes found for human infants of 22 months (-0.82) and human adults (-1.00). During months 2–8 (Fig. 1d, Table 1), infants showed a higher slope (-1.07) than even adult dolphins, which also differed considerably from the slope generated from these same infants at less than one month of age. This distribution suggests a less diverse (i.e. random) and more repetitious repertoire during months 2–8 than during the first month of infancy as well as a less diverse repertoire in comparison to adults. In months

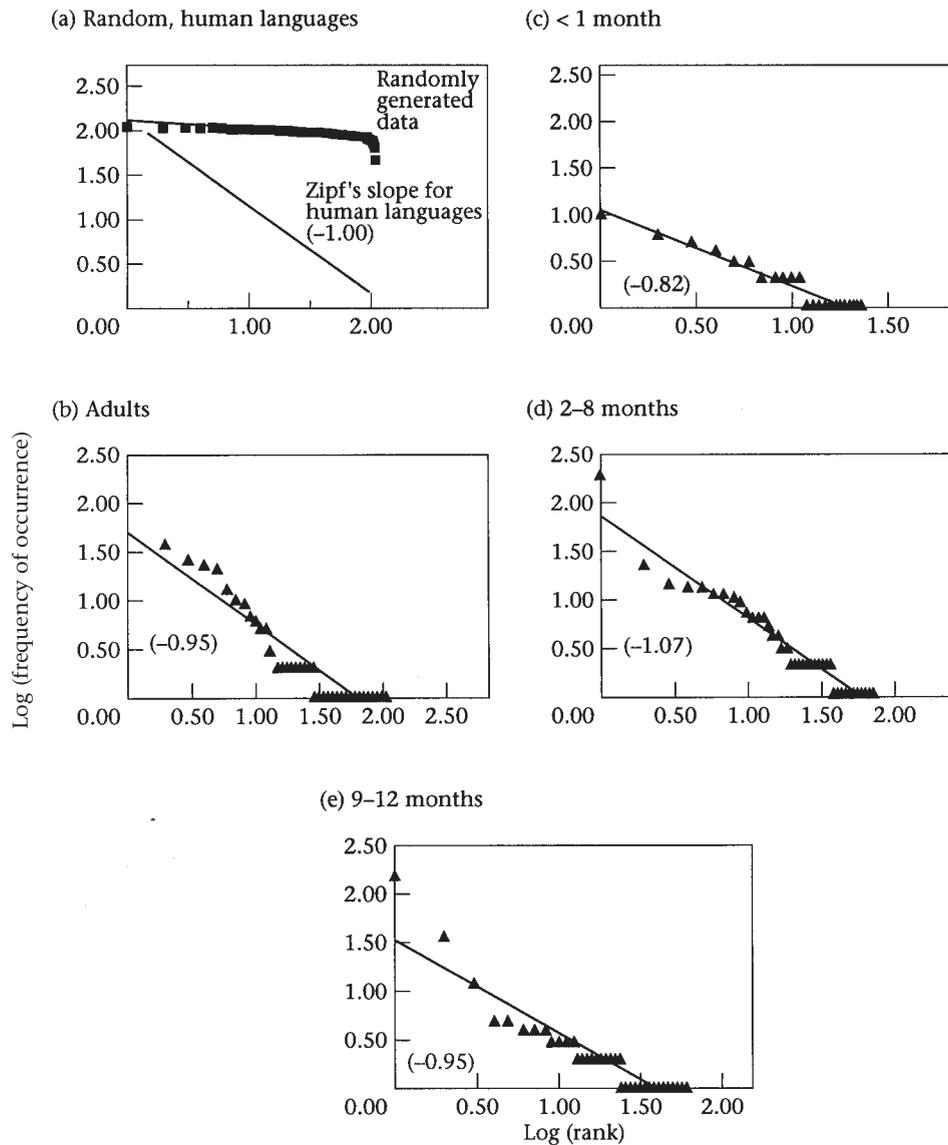


Figure 1. Regression of log (rank) versus log (frequency of occurrence) and lines of best fit for the whistle repertoires of adult dolphins (b) and infant dolphins at three developmental stages (c–e). Zipf's slope of human languages (-1.00) and randomly generated data (-0.09) also are shown (a). Slopes are shown in parentheses (also see Table 1).

Table 1. Regression summary of log (rank) on log (frequency of occurrence) for four age groups of dolphins

Age	Adults	<1 month	2–8 months	9–12 months
Regression coefficient	-0.95	-0.82	-1.07	-0.95
Standard error	0.04	0.04	0.03	0.05
Intercept	1.70	1.04	1.87	1.52
R^2	0.91	0.97	0.97	0.93
N (Whistle Types)	102	23	73	60
N (Whistles)	600	53	424	293

9–12 (Fig. 1e, Table 1), these infant dolphins showed a slope of -0.95 , essentially identical to that found for adult dolphins. These developmental changes in slope suggest that the distribution of whistles in the infant's repertoire becomes less random over development, eventually matching that of adults, and these changes

also suggest that infant dolphins acquire or learn the structure of whistle repertoires (Table 1).

Although the sample sizes of our dolphin analyses were much smaller than those reported for human languages, the dolphin whistles we used were adequately sampled for the Zipf plot (i.e. for a first-order entropic statistical

Table 2. Slopes of higher-entropic orders for various human languages, dolphin whistles and music

Signal system	Slope	R^2	Entropy			
			Zero order	First order	Second order	Third order
Russian letters	-0.500	0.93	5.00	4.35	3.52	3.01
English letters	-0.566	0.96	4.75	4.03	3.32	3.10
Music	-0.680	0.99	3.00	2.73	2.00	NA
Arabic letters	-0.797	0.96	5.00	4.21	3.77	2.49
Dolphin whistles	-1.334	0.86	4.75	1.92	1.15	0.56
Russian phonemes	-1.519	0.89	5.38	4.77	3.62	0.70

NA: not available.

analysis). A mean sample size of 10 times the number of signal types has been the suggested 'rule of thumb' for first-order entropy measures by some investigators (J. Hailman & E. Hailman; Uncert program). However, a sample size of four or five times the number of signal types is usually adequate. Sample sizes for each age class meet this latter criterion with the exception of the sample from the youngest infants (<1 month of age). However, 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. A much larger data set, however, would allow us to evaluate the reliability of our slope measures for all age categories.

In addition, 'power law noise', as discussed by some authors (Flam 1994; Damashek 1995) predicts that higher regression slopes can result from higher variance in the data. Thus the relationship between the variance and slope can result in a spurious coefficient (-0.6 in an example of random letters parsed into 'words' of random length; Mandelbrot 1953). However, our data show the opposite trend (that is, lower variance as the slope increases). Thus, our results are conservative in the sense that adjustments for a variance effect on the slopes would only increase the differences in the slope found, and then only slightly (see R^2 coefficients and standard errors in Table 1).

Higher-order Entropies

Table 2 presents the values and slopes of the higher-order entropies for English letters, Arabic letters, Russian letters, Russian phonemes, octave-system musical notes (H_0 was not available; Pierce 1961; Yaglom & Yaglom 1983), and our preliminary set of dolphin whistle vocalizations (McCowan 1995; McCowan & Reiss 1995a, b, 1997). Zero-order entropies (H_0) simply indicate sample diversity, while the existence of higher-order entropies (i.e. with H_n values less than H_0) indicate the presence of additional structure in the communication sequence (the entropy-order slope of a truly random number distribution would be zero). As predicted, human written languages as diverse as English, Arabic and Russian show similar values and slopes for zero- and higher-order entropies (Table 2). Unfortunately, the

sample shown of Russian phonemes (Yaglom & Yaglom 1983) as well as our sample of dolphin whistles (McCowan 1995; McCowan & Reiss 1995a, b, 1997) 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. As an example, the extent to which some words follow others (as in this written sentence) gives an artificially high representation of their statistical occurrence when only one sentence is considered. As more sentences are sampled, we would see, for example, that some words rarely follow others in a large English sample, even though they happened to in the one sentence considered. Therefore, the degree of complexity or slope would decrease. When the sample is sufficiently large, the value of the higher-order entropies is reliable (and these higher terms become substantially less than the initial entropies; see the equations for H_2 and H_3), thereby lowering the value of the overall entropic slope. Thus, as more data are added, one would expect the slope of Russian phonemes to approach that of the Russian letters. The determination of the well-sampled entropic slope of a statistically complete set of dolphin whistle vocalizations remains an intriguing near-term possibility. As we collect additional data, we will see what the entropic slope for dolphin whistles indicates about their organizational complexity, that is, how much 'communication complexity' their whistle communication system contains.

To determine how much additional sampling is needed to reliably calculate higher entropic orders, we can calculate how many values per cell (a cell being a combination probability) for the second-order entropy (first-order Markov chain) is required. The number of different ways probable events can be combined (the binomial equation) is:

$$N(r) = n! / r!(n - r)!$$

where r is the number of ways the samples (whistles, in our case) are taken and n is the total number of sample (whistle) types. In the case of second-order entropy, n is equal to 27 and r is 2 (for second-order entropy). Thus, we calculate: $27! / 2!(27 - 2)! = 351$. This puts our sample size per call at approximately 1.4 samples. For third-order

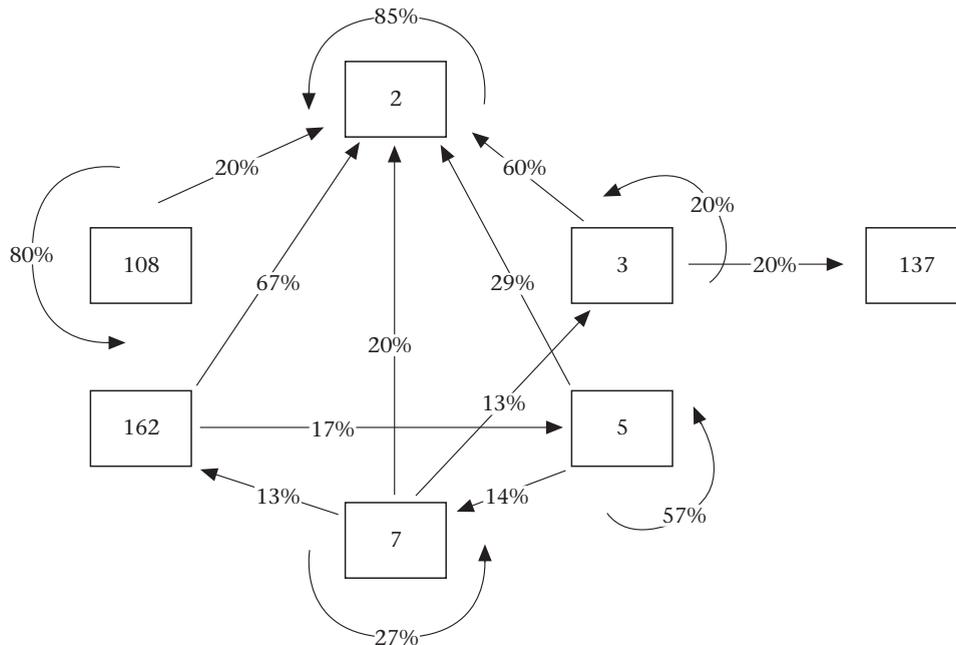


Figure 2. One set of two-whistle sequences shown as a probability tree based on a Markovian first-order (i.e. Shannon second-order entropy) analysis. Numbers in boxes represent whistle types. Percentages and direction of arrows shown represent the probability of one whistle type immediately following a second whistle type. A curved arrow indicates the probability that a whistle of one type immediately follows itself. Probabilities below 0.1 are not shown. The number of whistles for each whistle type (WT) included in the diagram were: WT2=188, WT7=15, WT162=12, WT5=7, WT108=5, WT3=5, WT137=1.

entropy, we should have: $27!/3!(27-3)! = 2,925$. This would put about 0.17 samples per cell for the 493 total samples (whistles) in our data set. A good sample size per cell would be 4–10 samples, but for an absolute minimum indicative value, one should have at least one sample per cell.

The number of different whistle types can, however, increase if the system has not been completely sampled (i.e. the system is still 'open'). Thus n will increase. By holding the number of types constant (e.g. 27) we can estimate the sample sizes required to obtain at least one sample per cell for different entropic orders (simply multiply by '10' for 10 samples per cell). As a rough estimate, using 27 different whistle types, we calculated the sample sizes required assuming a fully sampled (i.e. closed) system. For entropic orders 2, for example, a total of 351 whistles would be sufficient. However, for entropic order 7, a total of 888 030 whistles would be required to conduct such analyses.

Markovian Sequence Analysis and Probabilities for Shannon Entropies

We can examine one aspect of the second-order entropy of our dolphin whistle data set in more detail, looking at just the complexity of two-whistle sequences (Fig. 2). Probabilities generated from the Markovian-chain analysis give a preliminary indication of the sequential predictability of dolphin whistles. As an example from Fig. 2, consider whistle type 7. When this whistle occurs within a string of whistle vocalizations generated by one dolphin, the probability that the

whistle type immediately following is: for whistle type 162, 13% of the time; whistle type 3, 13% of the time; whistle type 2, 20% of the time; and whistle type 7, 27% of the time. We have not calculated the probabilities of a third whistle occurring in sequence, given that the other two have occurred within a string of whistles, as our data set is not yet sufficiently sampled for a reliable result. Thus, sequences of greater than two whistles within a string of whistles are not represented in Fig. 2. This analysis reflects an undersampled second-order Shannon entropy test. Further sampling of whistles from many individuals from several different social groups would allow us to validate this preliminary result and analyse higher-order entropic sequences (three- and four-whistle sequences, and so forth). This work is presently in progress.

DISCUSSION

Zipf's Relation: Comparative Measure of Repertoire Structural Complexity

These results suggest that Zipf's statistic can be applied appropriately to animal vocal repertoires, and specifically to dolphin whistle repertoires and their development. Both the adult and infant developmental patterns in comparison to human language and its acquisition suggest that Zipf's statistic may be an important comparative measure of repertoire complexity both across species and as an indicator of vocal acquisition or learning of vocal repertoire structure within a species.

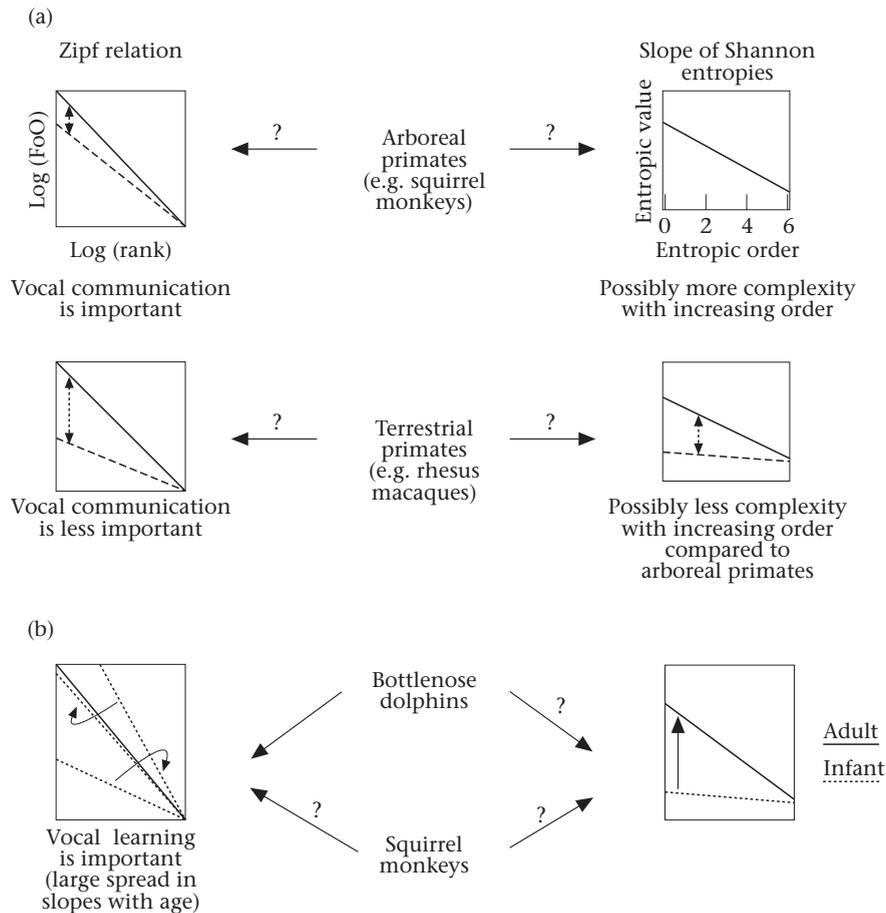


Figure 3. Predictions of Zipf's relation as a measure of repertoire structural complexity, and predictions of Shannon entropic slope as a measure of repertoire organizational complexity for comparison across species differing in (a) behavioural ecology and (b) the degree to which learning plays a role in development. Note that predictions presented here for vocal communication might be expected to be reversed if one applies this measure to another form of communicative signalling (e.g. gestural or facial communication).

Zipf's statistic measures the potential communication capacity of a repertoire using essentially a first-order entropic relation. This statistic, applied to animal communication systems, can be interpreted only on a comparative and/or developmental basis, not unlike the 'Encephalization Quotient' (Jerison 1973, 1982, 1985) used in studies comparing the potential information processing capacity of different species (Clutton-Brock & Harvey 1980; Macphail 1982; Connor et al. 1992; Marino 1996). A negative regression coefficient (negative slope) approaching -1.00 for a communicative repertoire suggests high potential for communication capacity in that repertoire (to be confirmed with higher-order entropic analysis), analogous to the relationship between a high encephalization quotient and high information processing capacity (Clutton-Brock & Harvey 1980; Macphail 1982; Connor et al. 1992; Marino 1996). As the regression coefficient moves away from -1.00 (both negatively and positively), the less the potential 'communication' capacity. For decreasing negative slopes from -1.00 , towards 0.00 , distributions become more randomly organized (more diverse, less repetitious). For increasingly steep negative slopes, they become more singular in content (less diverse, more

repetitious). In each case, they contain less potential 'communication' capacity. The Zipf's statistic can be used to test predictions on vocal repertoire complexity in light of behavioural ecology and evolutionary theory. Zipf's statistic could be examined in relationship to similarities and differences in social organization, life history characteristics, social complexity, and other socioecological parameters across a variety of species. For example, we might predict that species using vocal communication as a primary mode of social communication (e.g. forest-dwelling primates) will show a slope closer to -1.00 for their vocal repertoires than those species that do not use vocal communication as a primary mode of communication (e.g. savanna-dwelling primates; see Fig. 3a). Similarly, we might predict that infant vocal repertoires of species that learn vocally, such as humans and dolphins (Reiss & McCowan 1993; McCowan 1995; McCowan & Reiss 1995a, b, 1997), will show a lower slope (a more randomly distributed repertoire) at early developmental stages relative to conspecific adults than infant vocal repertoires of species that do not learn vocally (see Fig. 3b). However, the in-depth analysis of communication complexity lies in an analysis of the higher-order entropies.

Higher-order Entropies: Comparative Measure of Repertoire Organizational Complexity

These results suggest that dolphin whistle vocalizations contain some higher-order internal structure, enough at least to begin to predict statistically what whistle types might immediately follow the same or another whistle type. A more thorough knowledge of such higher-order entropic structures could eventually allow the reconstruction of dolphin whistle sequence structure, independent of any additional kinds of data input (such as actions or nonvocal signalling; such a reconstruction was first demonstrated by Shannon in 1949 for English). In addition, the relationship between entropic values, as represented by the slope of such values, could provide a new quantitative tool for comparing the organizational complexity of repertoires across species. As Zipf's statistic may provide an initial comparative indicator of repertoire structural complexity and its development, the slopes of higher-order entropies may serve as a comparative indicator of the deeper organizational complexity of animal vocal communication systems and its development. For example, one might expect the entropic slope to be, generally, more negative in species for which vocal communication plays an important role in mediating complex social interactions than in species for which it does not (see Fig. 3a). Similarly, one might predict that in species for which learning plays an important role in vocal development, the entropic slope would become more highly negative over development (see Fig. 3b). The utility of these comparative measures, however, depends upon the collection of extensive signal repertoire databases from many species that conserve the temporal integrity of signal sequences.

CONCLUSIONS AND FUTURE RESEARCH

Our analysis of dolphin whistle vocalizations using information theory measures has yielded potentially important tools for evaluating and comparing the internal informational structure of animal communicative repertoires in relationship to behavioural ecology and evolutionary theory. Yet this comparison will not be possible without extensive research and a broader perspective on animal communication systems. In addition to examining our own database, we conducted an extensive review of the literature of animal vocal repertoires in hope of including other species in this initial evaluation. We found that 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 the temporal integrity of vocalization sequences necessary for higher-order entropic analyses. We encourage all investigators studying animal vocal communication (and communication, in general) to collect comprehensive data that conserve temporal information and to categorize their data sets using objective (i.e. experimental and/or statistical) measures. In doing so, we may begin to quantitatively evaluate qualitatively different animal

vocal communication systems on a truly statistically comparative scale.

Although this report delineates the potential importance of information theory measures to the study of animal communication, further research is necessary to confirm its utility. First, using a larger dolphin whistle database, we need to evaluate the integrity of the higher entropic orders in dolphin development found in this study and their relationship to vocal learning. We also need to investigate how our data on whistle repertoires and development from captive dolphins compare to those from wild dolphins. Second, we need to study further how partially to correct mathematically for under-sampling of a repertoire for the higher-order entropies, so that we can still make trend predictions based on under-sampled data. Third, we need to evaluate whether other areas of repertoire organization and function can be examined using these information theory measures (for example, the correlation of frequency of signal use with signal duration as a possible measure of energy optimization). One might also examine components of the Zipf's plot for additional information relating to social group composition (e.g. socially familiar versus unfamiliar individuals might show a flatter slope at low rank values, that is for the most frequently used vocalizations, as in humans; Zipf 1949, 1968). We can also experimentally test the predictability of statistical vocal sequences and examine the correspondence between sequences of vocal and nonvocal behaviours in relation to their contexts. Most importantly, we must expand the application of these statistical measures of communicative repertoires to a diversity of species. We may then begin to determine how predictions of information theory measures of animal communication fit within the comparative frameworks of behavioural ecology and evolutionary theory.

Acknowledgments

We thank Jon Jenkins, Neal Heather, Sheldon Fisher, Diana Reiss and Lori Marino for their help and advice on this project, SETI Institute and Marine World for providing conference space, and Jack Hailman and two anonymous referees for their helpful comments. The research presented here was described in Animal Research Protocol No. D9002 approved on 15 April 1990 by the Institutional Animal Care and Use Committee of the Marine World Foundation at Marine World Africa U.S.A.

References

- Altmann, S. 1965. Sociobiology of rhesus monkeys. II: Stochastics of social communication. *Journal of Theoretical Biology*, **8**, 490–522.
- Beecher, M. D. 1989. Signalling systems for individual recognition: an information theory approach. *Animal Behaviour*, **48**, 248–261.
- Bonhoeffer, S., Hertz, A. V. M., Boerlijst, M. C., Nee, S., Nowak, M. A. & May, R. M. 1996. No signs of hidden language in noncoding DNA. *Physical Review Letters*, **76**, 1977.
- Burdin, V. I., Reznik, A. M., Skornyakov, V. M. & Chupakov, A. G. 1975. Communication signals in the black sea bottlenose dolphin. *Soviet Physics Acoustics*, **20**, 314–318.

- Chatfield, C. & Lemon, R. 1970. Analysing sequences of behavioural events. *Journal of Theoretical Biology*, **29**, 427–445.
- Clutton-Brock, T. H. & Harvey, P. H. 1980. Primates, brain and ecology. *Journal of Zoology*, **190**, 309–323.
- Connor, R. C., Smolker, R. A. & Richards, A. F. 1992. Dolphin alliances and coalitions. In: *Coalitions and Alliances in Humans and Other Animals* (Ed. by A. H. Harcourt & F. B. M. de Waal), pp. 415–443. Oxford: Oxford University Press.
- Damashek, M. 1995. Gauging similarity with *n*-Grams: language independent categorization of text. *Science*, **267**, 843–848.
- Dawkins, M. S. 1995. *Unravelling Animal Behaviour*. New York: J. Wiley.
- Deacon, T. W. 1990. Rethinking mammalian brain evolution. *American Zoologist*, **30**, 629–706.
- Ding, W., Würsig, B. & Evans, W. E. 1995. Whistles of bottlenose dolphins: comparisons among populations. *Aquatic Mammals*, **21**, 65–77.
- Dingle, H. 1969. A statistical and information analysis of aggressive communication in the mantis shrimp *Gonodactylus bredini Manning*. *Animal Behaviour*, **17**, 561–575.
- Dreher, J. J. 1961. Linguistic considerations of porpoise sounds. *Journal of the Acoustical Society of America*, **33**, 1799–1800.
- Dreher, J. J. 1966. Cetacean communication: small-group experiment. In: *Whales, Dolphins and Porpoises* (Ed. by K. S. Norris), pp. 529–543. Berkeley, California: University of California Press.
- Dreher, J. J. & Evans, W. E. 1964. Cetacean communication. In: *Marine Bioacoustics. Vol. 1* (Ed. by W. N. Tavolga), pp. 373–399. Oxford: Pergamon Press.
- Ficken, M. S., Hailman, E. D. & Hailman, J. P. 1994. The chick-a-dee call system of the Mexican chickadee. *Condor*, **96**, 70–82.
- Flam, F. 1994. Hints of a language in junk DNA. *Science*, **266**, 1320.
- Hailman, J. P. 1994. Constrained permutation in 'chick-a-dee'-like calls of the black-lored tit, *Parus xanthogenys*. *Bioacoustics*, **6**, 33–50.
- Hailman, J. P. & Ficken, M. S. 1986. Combinatorial animal communication with computable syntax: 'chick-a-dee' calling qualifies as 'language' by structural linguistics. *Animal Behaviour*, **34**, 1899–1901.
- Hailman, J. P., Ficken, M. S. & Ficken, R. W. 1985. The 'chick-a-dee' calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica*, **56**, 191–224.
- Hailman, J. P., Ficken, M. S. & Ficken, R. W. 1987. Constraints on the structure of combinatorial 'chick-a-dee' calls. *Ethology*, **75**, 62–80.
- Haldane, J. & Spurway, H. 1954. A statistical analysis of communication in *Apis mellifera* and a comparison with communication in other animals. *Insectes sociaux*, **1**, 247–283.
- Hazlett, B. & Bossert, W. 1965. A statistical analysis of the aggressive communications systems of some hermit crabs. *Animal Behaviour*, **13**, 357–373.
- Israeloff, N. E., Kagalenko, M. & Chan, K. 1996. Can Zipf distinguish language from noise in noncoding DNA? *Physical Review Letters*, **76**, 1976.
- Jerison, H. J. 1973. *Evolution of the Brain and Intelligence*. New York: Academic Press.
- Jerison, H. J. 1982. The evolution of biological intelligence. In: *Handbook of Human Intelligence* (Ed. by R. J. Sternberg), pp. 723–791. Cambridge: Cambridge University Press.
- Jerison, H. J. 1985. Animal intelligence as encephalization. *Philosophical Transactions of the Royal Society of London, Series B*, **308**, 21–35.
- Kaznadzei, V. V., Krechi, S. A. & Khakhalkina, E. N. 1976. Types of dolphin communication signals and their organization. *Soviet Physics Acoustics*, **22**, 484–488.
- McCowan, B. 1995. A new quantitative technique for categorizing whistles using simulated signals and whistles from captive bottlenose dolphins (Delphinidae, *Tursiops truncatus*). *Ethology*, **100**, 177–193.
- McCowan, B. & Reiss, D. 1995a. Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): a re-evaluation of the signature whistle hypothesis. *Ethology*, **100**, 194–209.
- McCowan, B. & Reiss, D. 1995b. Whistle contour development in captive-born infant bottlenose dolphins (*Tursiops truncatus*): role of learning. *Journal of Comparative Psychology*, **109**, 242–260.
- McCowan, B. & Reiss, D. 1997. Vocal learning in captive bottlenose dolphins: a comparison to humans and non-human animals. In: *Social Influences on Vocal Development* (Ed. by C. T. Snowdon & M. Hausberger), pp. 178–207. Cambridge: Cambridge University Press.
- Macphail, E. M. 1982. *Brain and Intelligence*. Oxford: Clarendon Press.
- Mandelbrot, B. 1953. Contribution à la théorie mathématique des jeux de communication. *Publique L'Institut à L'Université de Paris*, **2**, 5–50.
- Mantegna, R. N., Buldyrev, S. V., Goldberger, A. L., Harlin, S., Peng, C.-K., Simons, M. & Stanley, H. E. 1994. Linguistic features of noncoding DNA. *Physical Review Letters*, **73**, 3169–3172.
- Marino, L. 1996. What can dolphins tell us about primate evolution? *Evolutionary Anthropology*, **5**, 81–85.
- Moore, S. E. & Ridgway, S. H. 1995. Whistles produced by common dolphins from the Southern California Bight. *Aquatic Mammals*, **21**, 55–63.
- Pea, R. D. 1979. Can information theory explain early word choice? *Journal of Child Language*, **6**, 397–410.
- Pierce, J. R. 1961. *Symbols, Signals and Noise: The Nature and Process of Communication*. New York: Harper & Row.
- Preston, J. 1978. Communication systems and social interactions in a goby-shrimp symbiosis. *Animal Behaviour*, **26**, 791–802.
- Reiss, D. & McCowan, B. 1993. Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *Journal of Comparative Psychology*, **107**, 301–312.
- Ridley, D. R. & Gonzales, E. A. 1994. Zipf's law extended to small samples of adult speech. *Perceptual and Motor Skills*, **79**, 153–154.
- Rowe, G. & Harvey, I. 1985. Information content in finite sequences: communication between dragonfly larvae. *Journal of Theoretical Biology*, **116**, 275–290.
- Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423, 623–656.
- Shannon, C. E. & Weaver, W. 1949. *The Mathematical Theory of Communication*. Urbana, Illinois: University of Illinois Press.
- Steinberg, J. B. 1977. Information theory as an ethological tool. In: *Quantitative Methods in the Study of Animal Behavior* (Ed. by B. A. Hazlett), pp. 47–74. New York: Academic Press.
- Steinberg, J. & Conant, R. 1974. An information analysis of the inter-male behaviour of the grasshopper *Chortophaga viridifasciata*. *Animal Behaviour*, **22**, 617–627.
- Voss, R. F. 1996. Comment on 'linguistic features of noncoding DNA sequences'. *Physical Review Letters*, **76**, 1978.
- Wells, R. S., Scott, M. D. & Irvine, A. B. 1987. The social structure of free-ranging bottlenose dolphins. In: *Current Mammalogy* (Ed. by H. Genoways), pp. 247–305. New York: Plenum.
- Wilson, E. O. 1962. Chemical communication among workers of the fire ant *Solenopsis saevissima* (Fr. Smith): 2. An information analysis of the odour trail. *Animal Behaviour*, **10**, 148–158.
- Yaglom, A. M. & Yaglom, I. M. 1983. *Probability and Information*. Boston: D. Reidel.
- Zipf, G. K. 1949. *Human Behavior and the Principle of Least Effort*. Cambridge: Addison-Wesley Press.
- Zipf, G. K. 1968. *The Psycho-Biology of Language: An Introduction to Dynamic Psychology*. Cambridge: Addison-Wesley Press.