

INVESTIGATIONS ON CETACEAN SONAR IV, A COMPARISON OF WAVE SHAPES OF ODONTOCETE SONAR SIGNALS

H. Wiersma, *Laboratory for Information Theory, Delft, University of Technology, Mekelweg 4, Delft, the Netherlands*

Summary

High-quality recordings of a number of members of odontocete families are now available, thus permitting a more reliable and detailed analysis than ever before. Waveforms of four different species, *Tursiops truncatus*, *Inia geoffrensis*, *Sotalia guyanensis* and *Phocoena phocoena*, are compared.

As remarkable as it may seem, the signals of these species all have a basic wave shape in common. This can best be understood by realising that all these species use sonar signals which are simultaneously very short in time and very narrow in frequency. They approach the theoretical lower bound for the product of time duration and frequency bandwidth and are therefore neither 'broadband' nor 'narrowband', but merely 'small time duration bandwidth' signals. The same can be said for the individual components of those sonar signals that are actually composed of two separate components at different frequencies.

Another characteristic feature they all share is that the actual sonar signal itself is consistently followed by reverberations, which are probably due to multipath effects inside the animal's head. For some species it is very pronounced, while for others an excellent signal-to-noise ratio is required to be able to notice it at all. Either way, it is far from a random process; on the contrary, it is very systematic and typical of each species.

A review of the material available, which now includes data from *Sotalia guyanensis* not earlier published, clearly demonstrates how alike the sonar systems of entirely different species and even of families are.

Introduction

Analysing recordings of sonar sound is generally a time-consuming operation. It requires the inspection of large amounts of data recorded with a bandwidth of at least ten times the human upper-frequency limit. A preliminary inspection by the human ear therefore requires the recording tape speed to be decelerated by at least that factor. Besides, even at the highest pulse-repetition rates more than 95% of the time is consumed by interpulse intervals and silences between pulse trains. Unless the recording was made during a carefully controlled experiment, using a fixed measuring setup, a preselection procedure must be carried out to sort out those pulse trains that are not too corrupted by noise and reflections from obstacles or walls. To be able to accomplish this, precise knowledge of the position and direction of the dolphin with respect to the recording hydrophone is indispensable.

Furthermore, it is essential to account for the directionality of the emitted sonar beam. Although up to now the exact location of the source(s) of the sonar sound is unknown, it is obvious that the transfer function of the dolphin's head depends upon the position of the hydrophone within the sound field emitted. Moreover, reflections from the skull partially interfere with the actual sonar signal, resulting in seemingly different waveform patters within the aperture of the sonar beam.

For these reasons averaging signals in order to obtain some average waveform is clearly not a suitable analysis method when the detailed structure of individual sonar signals is the subject of investigation. Since it is practically impossible to design automatic procedures to locate uncorrupted echolocation signals, the analysis to be done remains elaborate. Manual inspection of all the available data by listening and viewing on a slowed-down time scale while simultaneously watching the animal's behaviour recorded on videotape has proved to be worthwhile. A typical waveform is thus easily identifiable and minor variations can be followed during an entire pulse train. The data given here were selected according to this procedure.

Instrumentation and methods

All data were obtained from captive animals confined to the rather small dimensions of indoor tanks, most of the animals having been in captivity for a number of years. Standard recording equipment was used, consisting of Bruel & Kjaer hydrophones, types 8101 and 8103, amplifiers, types 2608, of the same manufacturer and a Racal $\frac{1}{2}$ inch instrumentation recorder, STORE 7D, running at a speed of 30 ips. Hum and rumble were sufficiently eliminated by a high-pass filter with cut-off at 200 Hz. The whole system covered a frequency pass-band up to at least 150 kHz.

This upper-frequency limit corresponds to a minimum wavelength of 1 cm, meaning that a dolphin emitting signals having a frequency of 150 kHz would be able to resolve differences in range within an order of magnitude of centimetres. This upper-frequency limit is presumably high enough to ensure that all acoustical energy present in echolocation signals will be properly recorded. There is no indication whatsoever that higher frequencies are used by any of the animals. Occasionally recordings with a higher frequency limit were made, but no significant energy was ever found beyond that 150 kHz limit. This of course does not prove that there can not be any energy at higher frequencies. However, since it can never be proved that something is not there just because it has not been observed, we must assume that the frequency range of interest can be limited to 150 kHz.

The dolphin's behaviour was recorded on a video recorder by means of two black and white cameras under water. One of the two available audio channels of the video recorder was used to record the output of the 8101 hydrophone and the other to record a time code for synchronizational purposes. As the same time code was put at real time on one of the channels of the instrumentation recorder an off-line accuracy of 20 msec, the time lapse of one video frame, could be attained.

The recorded tapes were processed by using a PDP 11/40 mini-computer of Digital Equipment Corporation, equipped with special peripherals to facilitate signal analysis, such as an LPS system for 12-bit analogue-to-digital conversion and a VT11 graphics display for signal presentation. With a maximum usable sampling frequency of 25 kHz and a maximum tape deceleration factor of 32 a sampling frequency of 800 kHz could be obtained. For the spectral analysis a 1024-point FFT procedure was used on an interactively selected arbitrary number of data points, exactly covering the entire echolocation signal without including reverberations - if any.

The components of one signal at different frequencies were separated by analogue filtering by means of a Rockland filter with a slope of -48 dB/oct, the cross-over frequency of which was set at the geometric mean of the dominant frequencies of the components to be separated. Since we will be dealing with such concepts as dominant frequency, time duration and bandwidth precise mathematical formulas must be given, so that these quantities can be deter-

mined for a particular signal. Please note that in this section the text in small print contains additional more detailed information but it is not necessary for the outline of this paper.

For the *time duration* we use a quantity which is analogous to the standard deviation known from statistics. If an echolocation is known as a function of time, it will be concentrated around some point in time and the spread of the signal, computed as a standard deviation around this point, is a measure for the duration of time during which the signal values are significant. The same applies to the *frequency bandwidth*, which is computed, so to speak, as a standard deviation of the frequency spectrum relative to the *dominant frequency*. This in turn is taken to be the mean frequency of the spectrum. For most of the echolocation signals this is very close or equal to that frequency at which most of the energy is found.

Suppose an echolocation signal is known as a function of time $s(t)$, say, then its total energy is proportional to

$$\int_T |s(t)|^2 dt,$$

where T denotes that part of the time axis at which $s(t)$ is known. Without loss of generality we may assume that the signals are all normalised in such a way that the total energy equals unity:

$$\int_T |s(t)|^2 dt = 1.$$

Since the instantaneous energy $|s(t)|^2$ is non-negative and its sum now equals unity it has all the properties of a probability density function, so that we may compute a mean value \bar{t} :

$$\bar{t} = \int_T t |s(t)|^2 dt,$$

the point around which $|s(t)|^2$ is centred. The standard deviation relative to this value equals

$$\left[\int_T (t - \bar{t})^2 |s(t)|^2 dt \right]^{\frac{1}{2}},$$

a form analogous to the ordinary standard deviation in statistics. This quantity may be taken as a basis for the definition of the time duration Δt of a signal:

$$\Delta t = C_1 \left[\int_T (t - \bar{t})^2 |s(t)|^2 dt \right]^{\frac{1}{2}},$$

where C_1 is an arbitrary positive proportionality constant which shall be determined later.

Similar definitions can be derived from the spectrum $S(f)$ as a function of the frequency if $S(f)$ is determined as the decomposition of $s(t)$ into elementary harmonic waves of different frequencies

$$S(f) = \int_T s(t) \exp(-2\pi i f t) dt,$$

in which i is the imaginary unit ($i^2 = -1$).

This formula is also referred to as the Fourier transform of $s(t)$, named after the French mathematician who developed the fundamentals of this frequency analysis. Now the frequency around which $S(f)$ is centred - called the dominant frequency - is given by

$$f_d = \int_{\Omega} f |S(f)|^2 df,$$

where Ω denotes the halfline of positive frequencies. The bandwidth Δf is, quite analogous to the time duration, given by

$$\Delta f = C_2 \left[\int_{\Omega} (f - f_d)^2 |S(f)|^2 df \right]^{\frac{1}{2}}$$

and C_2 is, like C_1 , a proportionality constant.

The product of the time duration and the frequency bandwidth is a positive dimensionless number; it is known from communication theory (GABOR, 1946) that it can attain arbitrarily large but not arbitrarily small values. In fact, it is always greater than unity for all echolocation signals.

Due to the Fourier transform relationship between $s(t)$ and $S(f)$ the product of Δt and Δf has an absolute lower bound, in fact

$$\Delta t \cdot \Delta f > K(\Delta f_r) \frac{C_1 C_2}{4\pi}$$

The function $K(\Delta f_r)$ only depends upon the relative bandwidth Δf_r , defined as

$$\Delta f_r = \frac{\Delta f}{f_d}$$

Dependant upon a certain relative bandwidth it reaches values between 0.59 ... and 1.00. The proof of this can be found in Hilberg and Rothe (1971).

The constants C_1 and C_2 which appear in the definition of Δt and Δf may be chosen so as to meet the requirements of a particular application. For our purposes we set $C_1 = 2\pi$ and $C_2 = 2$. This choice is experimentally determined such that Δt corresponds closely to the time during

which $s(t)$ is significantly different from zero and such that Δf almost equals the -3 dB or half-power bandwidth. These are the figures usually given in literature on echolocation signals. Again, the actual choice is irrelevant for comparative purposes, but it does facilitate the interpretation of the values for the quantities duration and bandwidth.

For the relative bandwidths found for echolocation signals we may show that the value of $K(\Delta f_r)$ is very close to unity. Therefore, the inequality satisfied by the product $\Delta t \Delta f$ now reads

$$\Delta t \cdot \Delta f \gtrsim 1,$$

the so-called uncertainty relation of communication theory. It greatly resembles the uncertainty principle of Heisenberg in quantum mechanics.

Not only the *absolute bandwidth* but also the *relative bandwidth*, i.e. the ratio of the bandwidth and the dominant frequency, is an important figure from an engineering point of view. Systems with small relative bandwidths are usually referred to as narrow-band systems. Hence, the relative bandwidth can be considered to be a measure of 'narrowness' of the spectrum of a signal. For instance, in communication engineering values less than 1% are not unusual, and these signals are freely labelled 'narrow-band'. But from the table we can see that echolocation signals may have relative bandwidths in the order of 50%. It is left to the reader to decide whether this should be interpreted as narrow or broad.

Table 1. Mean and standard deviations of the time duration (Δt), the dominant frequency (f_d), the bandwidth (Δf), the relative bandwidth (Δf_r), the number of cycles (N_c), the time duration-bandwidth product ($\Delta t \Delta f$) and the number of signals over which the averaging was done (N).

	t (μsec)	f_d (kHz)	Δf (kHz)	Δf_r (%)	N_c	$\Delta t \Delta f$	N
Inia g.	72.4 \pm 2.4	45.9 \pm 0.7	18.8 \pm 0.7	41.0 \pm 1.8	3.32 \pm .11	1.36 \pm .06	20
Tursiops t.	59.7 \pm 3.6	41.0 \pm 2.3	22.6 \pm 2.8	55.5 \pm 8.1	2.45 \pm .21	1.34 \pm .10	10
Sotalia g. (HF)	27.7 \pm 0.7	94.7 \pm 1.4	39.6 \pm 1.3	41.9 \pm 1.5	2.63 \pm .09	1.10 \pm .06	10
Sotalia g. (LF)	98.7 \pm 3.7	29.2 \pm 0.7	15.2 \pm 0.5	52.0 \pm 2.9	2.89 \pm .17	1.50 \pm .06	10
Phocoena ph. (HF)	36.0 \pm 2.7	118. \pm 3.6	32.7 \pm 2.8	27.7 \pm 2.1	4.23 \pm .28	1.17 \pm .07	25
Phocoena ph. (LF)	201. \pm 9.4	21.6 \pm 0.6	7.55 \pm .63	35.0 \pm 3.3	4.34 \pm .29	1.51 \pm .12	25

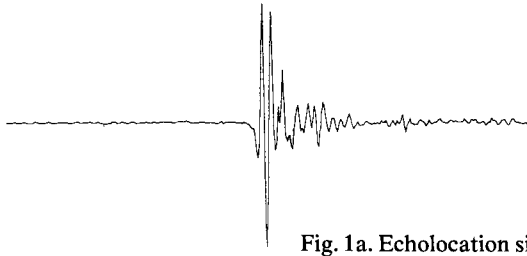


Fig. 1a. Echolocation signal of *Tursiops truncatus*

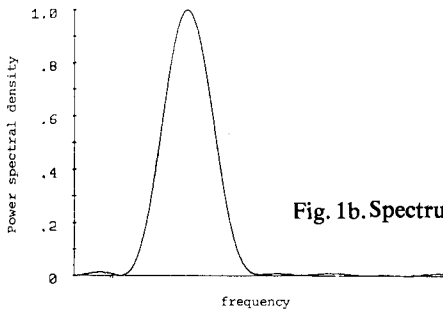
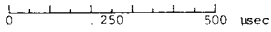


Fig. 1b. Spectrum of *Tursiops truncatus*

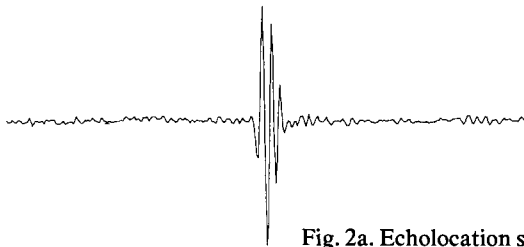


Fig. 2a. Echolocation signal of *Inia geoffrensis*

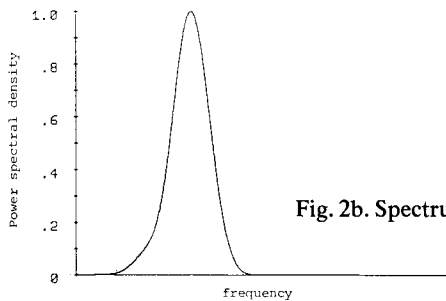
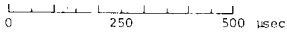


Fig. 2b. Spectrum of *Inia geoffrensis*

Finally, one other parameter can be considered to be a characteristic feature of an echolocation signal, namely the total *number of cycles*. Since it is often in practice not easy to count this number, averaging formulas may again be applied to estimate it. The number of cycles can simply be defined as the product of the previously obtained dominant frequency and the time duration. Note that this is indeed a positive dimensionless number.

The number of cycles N_c is defined as

$$N_c = f_d \cdot \Delta t,$$

and is clearly dependant upon the choice of the constant C_1 . It is therefore important to choose C_1 in such a way that N_c assumes values which make some sense for a class of signals. Specifically, if the number of cycles can easily be determined

by inspection, it is then desirable to let N_c assume values close to it. Of course, the problem of choosing values for the constants depends upon the properties of the signals within a certain class. Fortunately, almost all echolocation signals lie within a very restricted volume of the entire signal space, so that one and the same choice of constants applies equally well to all dolphin signals.

The quantities defined above are computed for typical echolocation signals of each species and compiled in a table as mean values with corresponding standard deviations. The number of signals of a species over which the averaging was done is included in that table as well.

Results

A female Atlantic bottlenosed dolphin (*Tursiops truncatus*), who had been in captivity since June 1968 in the Dolphinarium at Harderwijk, Netherlands, was trained to use her sonar system to locate plexiglass rings of about 15 cm diameter and a cross-section of the material of about 1 cm². To ensure the use of sonar she was completely blindfolded by rubber suction cups. In Januari 1980 echolocation signals were recorded, a typical one of which is plotted in Fig. 1a. Note that after the first 2½ periods the signal is corrupted by reverberations, probably the effect of multipath transmission inside the animal's head. This hypothesis is reinforced by the fact that in subsequent signals the first 2½ periods remain unaltered, while the appearance of the distortion before and after shows some variation.

The power spectrum $|S(f)|^2$ is plotted in Fig. 1b. Since only the waveform is of interest here the amplitude and frequency scale are normalised so as to facilitate comparison with signals of the other species. Consequently, no ordinate divisions are printed.

Echolocation signals of the Amazon river dolphin (*Inia geoffrensis*) have already been described (KAMMINGA and WIERSMA, 1981). Earlier data on *Inia geoffrensis* report peak energies of 60-80 kHz for animals living in the wild (EVANS, 1973), but in that study, as in all others, the computation of the dominant frequency is based upon the spectrum of the entire signal without any attention being paid to its internal structure. A sample waveform given there shows reverberation effects comparable to those in Fig. 1a.

The appearance of reverberation distortion depends upon the position of the hydrophone within the sound field emitted; occasionally signals are found without any reverberations at all. One example of this is Fig. 2a, which gives a typical echolocation signal of *Inia geoffrensis*, recorded during feeding (Duisburg Zoo, Duisburg, Germany, 1978). The spectrum is

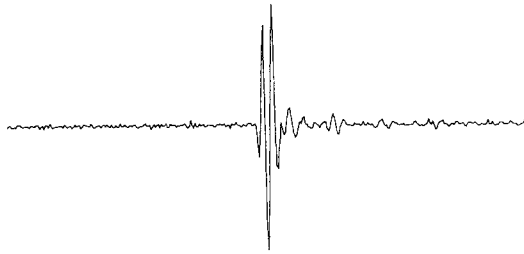


Fig. 3a. High-frequency component of *Sotalia guyanensis*

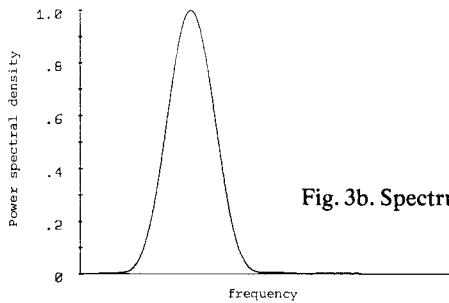
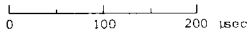


Fig. 3b. Spectrum of HF-component of *Sotalia guyanensis*

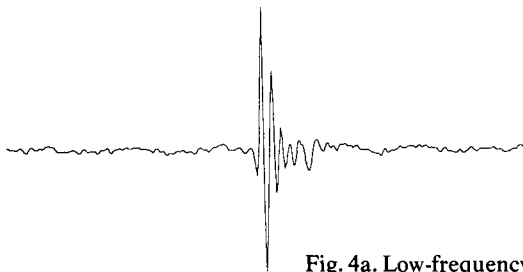


Fig. 4a. Low-frequency component of *Sotalia guyanensis*

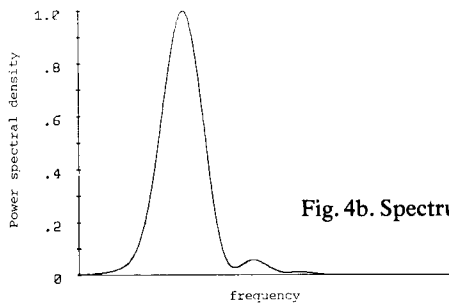
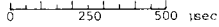


Fig. 4b. Spectrum of LF-component of *Sotalia guyanensis*

given in Fig. 2b. It must be added that further on within the same pulse train some distortion gradually appears. This is the effect of the movement of the animal's head when it was scanning to capture the fish.

Parameters such as dominant frequency, bandwidth and time duration should be computed from the signal as actually generated by the sound source, because the ultimate goal is to describe and characterise the source. Therefore, the recorded signal, being a result of the source signal passed through transmission paths inside the animal's head, the water and finally the hydrophone, has to be stripped of all the induced byproducts. If, for instance, reverberations are added to the original signal, the calculation should be based upon the undistorted part, should this be possible without severe truncation and - if not - by extrapolation of the undistorted but truncated part. Suitable extrapolation methods are available (WIERSMA, 1979), but will not be discussed here. Briefly, they use a mathematical model of the signal and then extrapolate the model beyond the undistorted part of the signal. This method had to be employed only for the signals of *Tursiops truncatus*.

Any other way of calculating parameters may lead to deviations as large as 20 to 30%. Consequently, by using the entire signal - as is usually done - a large variance in the parameters is inevitably induced, not by the data but by the way of defining them. This is procedurally wrong. It is for this reason that in literature one always finds intervals for the values of a parameter. Statements such as 'the signal has a frequency from 60 to 80 kHz' are the result of a wrong procedure and not of a wrong or variable signal.

The acoustical features of *Sotalia guyanensis* have, to the best of the author's knowledge, not been reported earlier. Although it was not really surprising to find that it had echolocation abilities similar to other species, this has not yet been extensively demonstrated. In Rhenen, Netherlands, two captive *Sotalia* were recorded (by courtesy of the staff of the Dolphinarium) in March, 1980. Without going into too much detail, only those results necessary for the comparison are briefly summarised here.

Sotalia guyanensis appears to possess a sonar signal with a high-frequency component at 95 kHz and a low-frequency component which is about 2 dB less in power at 30 kHz (see table). It is interesting to note the similarity in waveform between the individual components of one and the same sonar signal (Figs. 3a and 4a) and between the waveforms of *Sotalia* and other species. This similarity is even more clearly illustrated by the spectra (Figs. 3b and 4b).

One other well-known species, *Phocoena phocoena*, may be included in the comparison. For no other species have so many controversial data been published by so many different authors. This merely reflects the technical developments in the history of the recording of aquatic echolocation signals. A recent summary of the literature can be found in KAMMINGA and WIERSMA (1981).

It is recalled that *Phocoena phocoena* also possesses a two-component signal of 120 and 20 kHz, respectively (see table). Note again the striking resemblance between the two components (Figs. 5a and 6a) and their associated spectra (Figs. 5b and 6b). *Phocoena* signals appear to cover a few more cycles than the other species, but the individual parts can hardly be distinguished. The signals shown here are of a male animal stranded in June 1978 and taken care of in Harderwijk, Netherlands. The recording was done not long thereafter (December 1978).

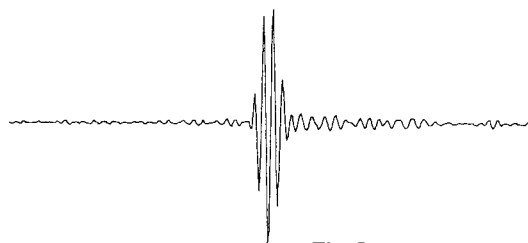


Fig. 5a. High-frequency component of *Phocoena phocoena*

0 100 200 μ sec

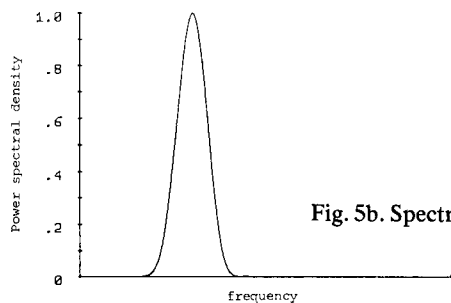


Fig. 5b. Spectrum of HF-component of *Phocoena phocoena*

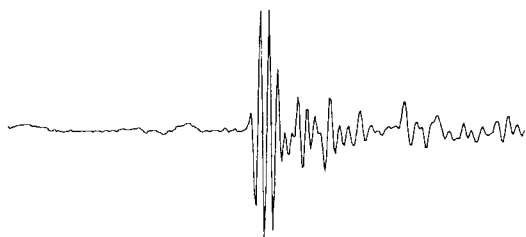


Fig. 6a. Low-frequency component of *Phocoena phocoena*

0 250 500 μ sec

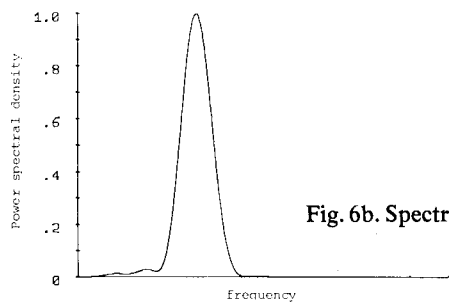


Fig. 6b. Spectrum of LF-component of *Phocoena phocoena*

Discussion and conclusions

From the table it appears that all species have sonar signals with very low time duration bandwidth products. Some of them even go as low as 1.03, compared to 1.00, the theoretically absolute lower bound. Of course, there is a question of how to interpret this number as a measure in the strict sense; nevertheless, qualitatively, it expresses that the volume of the entire waveform space is limited and that this limit decreases as the product decreases. Indeed, this volume can be made arbitrarily small by letting the time duration bandwidth approach unity. Theoretically, the lower bound can only be attained by a waveform that extends over the entire time axis, from the indefinite past to the indefinite future, which is of course not physically realisable. In this limiting case the volume would be zero, meaning that the variety in waveform within this volume is zero: there is only one point in the volume, one possible waveform. Therefore, since the product is small for all species it is not surprising that the waveforms all look alike.

For this reason these echolocation signals can better be characterised by referring to them as being small in time bandwidth rather than in terms of being narrow- or broadband. Moreover, on the basis of relative bandwidths it is hard to define what narrow or broad exactly is.

Finally, similarity remains a concept that is relative to the reader. However, objective measures can be applied to make statements of similarity more rigorous. But the human eye is a sophisticated comparator as long as the objects are presented in a suitable way. As the objective of this paper is to demonstrate the likeness of sonar signals of completely different species, even those belonging to different taxonomical families, it will be left to the reader to inspect the figures and the plots of the signals and judge for himself what the author means by similarity.

Acknowledgements

The author is indebted to all those who helped to make the recordings possible, not in the least to the staffs of the dolphinaria, where they were carried out. Special thanks are due to all those who contributed specifically to this research for their kind cooperation and inspiring discussion, in particular to B.M. van den Boom, Prof. Y. Bosma, Dr. W.H. Dudok van Heel and C. Kamminga.

References

- EVANS, W.E., 1973. Echolocation by marine dolphins and one species of fresh water dolphin. *J. Acoust. Soc. Am.*, 54: 191-200.
- GABOR, D., 1946. *Theory of Communication*. IEEE 93 (3) : 429-441.
- HILBERG, W. and P.G. ROTHE, 1971. The general uncertainty relation for real signals in communication theory. *Information and Control* 18: 103-125.
- KAMMINGA, C. and H. WIERSMA, 1981. Investigations on Cetacean sonar II. Acoustical similarities and differences in odontocete sonar signals. *Aq. Mammals*, 8 (2) : 41-62.
- WIERSMA, H., 1979. Intrinsic properties of the echolocation signals used by marine mammals. M.Sc. thesis, Delft University of Technology, the Netherlands.