Marine Animal SOUND Database

ABSTRACT

The Marine Animal SOUND Database system encompasses (1) descriptive text databases cataloging the WHOI collection of underwater sound recordings from marine animals, (2) sets of files of digital sound sequences, (3) text databases organizing the digital sound sequences, and (4) software for analysis, display, playback, and export of selected sound files. The text databases index and sort the information on the sounds. The digital sound files are accessed directly from the text record, analyzed on screen, listened to, and compared or exported as desired. These databases provide comprehensive means for quantitative analyses and statistical comparisons of marine animal vocalizations.

The objective has been to develop basic tools for the study of marine animal sounds. The text database for cataloging the collection of recordings provides convenient sorting and selection of sounds of interest. Then, as specific sequences are digitized from these recordings, they become part of a second database system that manages these sound data. Once a digital sound is part of the database, several tools are available for interactive spectrogram display, sound playback, statistical feature extraction, and export to other application programs.

KEY WORDS -- Sound database, Marine animals, Underwater sounds, Animal vocalizations.

Marine Animal SOUND Database

Marine Animal SOUND Database

CONTENTS

Marine Animal SOUND Database

OBJECTIVES

The Marine Animal SOUND Database system has been developed to provide convenient access to acoustic data in the WHOI collection of recordings from marine species. The organization of database structures would follow that designed for convenient indexing and retrieval of data in our CETACEA database of literature references (Watkins, Daher, and Haley 1990). We wanted convenient access to representative sounds from the collections of marine animal sounds. Acoustic sequences from the various animal repertoires would be digitized and annotated in order to provide a representative selection of the sounds associated with the behaviors of the different animal species in the collection, as well as of sounds from a number of other non-biological sources. To facilitate analysis and sound comparisons, these digital sound files would be organized and accessed through companion text databases. We wanted to augument the functions of these databases to include spectrographic analysis and display, and playback of the sound sequences.

The objective, therefore, has been to develop basic tools for the study of marine animal sounds, for accessing, analyzing, and comparing the acoustic patterns. This would include exploration of means for characterization of sound features that can be used for statistical comparison, and then for automatic recognition and diagnosis of these marine animal vocalizations.

The SOUND databases described here fulfill these objectives, and they open the way toward comprehensive, quantitative analyses and statistical comparisons of marine animal sounds in ways that have not been previously possible.

INTRODUCTION

Database Organization --

The development of the CETACEA database system was described by Watkins, Bird, Moore, and Tyack (1988), and current format revisions were detailed by Watkins, Daher, and Haley (1990). The SOUND and CETACEA databases use the same systems for cataloging subjects, dates, geographic locations, and pertinent descriptive data. These notations are directly related to more than 150 species categories (Species List appended), referenced to detailed listings for cetaceans, pinnipeds, sirenians, and general notations for fishes and other vertebrates and invertebrates.

The text information about the recordings cataloged by the SOUND databases combined with the search and sorting functions of our database management program provide convenient, rapid mechanisms for selecting sound recordings or digitized data of interest. The databases are presently installed on PC and AT compatible computers running Microsoft DOS, and the current database program is based on INMAGIC 7.1 software (INMAGIC INC, Cambridge, MA).

The SOUND database information is currently indexed in 40 fields (date, location, recording data, sound class, species, number of animals, author, etc.), and provides for independent sorting and retrieval of each field, as well as unlimited subfields. The databases may be searched by any combination of indexed or unindexed alphanumerical notations, by any combination of terms or partial phrases or stems of these notations. Searches may be complex and mixed with Boolean search statements. The data record for the desired sound file is selected by the text databases, and then, the related digitized sound file may be immediately retrieved, analyzed and displayed on screen without leaving the text database. These databases continue to enlarge, and currently hold about 5000 records.

Marine Animal SOUND Database

Marine Animal Recordings --

The sound recordings that provide the basis for the Marine Animal SOUND Databases have been collected during studies at sea on acoustic behavior of marine species. This bioacoustic work was started in 1947 by William E. Schevill, and a continuous program of study of marine animal acoustic behaviors has been maintained at WHOI since that time. The initial collection of tape recordings organized by the SOUND database had approximately 2000 magnetic tapes. These tapes included underwater sounds from approximately 70 species of marine mammals, as well as fishes, crustaceans, and selected ambient noises.

The collection of marine animal sound recordings is constantly enlarging as new tapes are added from our own experiments at sea (for example, two Arctic and three Caribbean cruises in 1989-1990) in addition to recordings by others. This sound library includes a variety of magnetic tapes, reel-to-reel of different widths and reel sizes, VCR tapes, cassettes, phonograph records, computer digital media and optical disks. The combined SOUND databases now have about 5000 entries.

The animal sounds in the recording collection and its database system incorporate a number of the historically important scientific contributions in bioacoustics, such as the first science recordings from marine mammals at sea by William Schevill (beluga, Delphinapterus leucas) and the early sound records of echolocating bats (Eptesicus and Plecotus) and oil birds (Steatornis caripensis) by Donald Griffin.

These original underwater bioacoustic studies were underway before tape recorders were readily available, so that recordings were made on a variety of different types of equipment. For example, the earliest records of the acoustic behavior of beluga were made on a "Gray Audograph" dictating machine, and the signals were recorded on waxed paper disks. These sounds were then reproduced for analysis and distribution by cutting onto transcription phonograph records. Copies of such recordings made on magnetic tape are retained in the collection.

The recording spectra on the different sound recording media vary with each system that was used, from bandwidths of only a few kHz to 200 kHz or more. The data were made with single channel systems and with arrays of four or more hydrophones, with sensor depths varying from near surface to bottom mounted units. Recording equipment has varied considerably over the years, including standard audio systems, professional standards IRIG systems, and special purpose recorders. Most systems used "Direct" recording because of the wide bandwidths or high frequencies required to reproduce the natural animal sounds.

Marine mammals have been emphasized in the studies of animal acoustic behavior, therefore, much of the recorded data has focused on the sounds of cetaceans, pinnipeds, and sirenians. The recordings in the collection include underwater sounds from more than 70 species of marine mammals. Geographic locations for these recordings are scattered worldwide, with observations in many locations often taken over much of the year. Because of the long sequence of studies and the wide range of acoustic work represented by these continuing programs, the recording collection is the most comprehensive one available for many areas of the world.

In recordings made prior to 1970, the attribution for sounds heard underwater was not always immediately certain, but with time, more and more of the repertoire of the different species became recognized. As a result, there has been a change in the nature of the recordings over time. The use of multiple hydrophone arrays has increasingly allowed identification of the vocalizations and sequences of sounds from individuals. Sources for the recorded sounds have been positively identified by these means. Many of the recordings have been made especially to trace the usage or development of particular sequences of vocal activity. The more recently recorded tapes deal less with sounds that relate generally to species or population differences, and more with sounds that illustrate specific behaviors. Our later recordings are, therefore, more focused on specific scientific objectives (vocal identity, mimicry, shared signalling, ontogeny of learned signals) in selected species (sperm whales, finbacks, bowheads, bottlenose dolphins, etc). For example, our 1954-1968 sperm whale recordings led to specific studies of their activities, which in turn opened the way to work focused on coda signals produced by these whales, and then to analyses of distinctions between codas from individuals and those shared by the members a particular whale group.

ACKNOWLEDGEMENTS

Development of the Marine Animal SOUND Database system is a continuing program with joint effort and contributions from all in the WHOI bioacoustics program, including William Schevill, Peter Tyack, Nancy Haley, Terrance Howald, Laela Sayigh, and Cheri Recchia. This effort is also shared by previous researchers in our program, including Karen Moore, and James Bird. The data collection at sea also had expert contribution from all of those associated with the vessels and aircraft used for observations over these decades. The organization and interpretation of the marine animal acoustic data encoded now by the SOUND databases are a result of the collective ideas and experiences of this entire group. The authors' contributions have emphasized basic structural organization (Watkins), innovative computer tools for augmenting the database functions (Fristrup), and careful oversight of acoustic and annotation records (Daher). The central core of this program is the long standing collection of marine animal sound recordings -- tributes to the researchers that have made these important contributions over 44 years.

Support for the development of the Marine Animal SOUND Databases has been from the Ocean Acoustics Program (code 11250A) of the Office of Naval Research, Contract N00014-88-K-0273 and Research Grant N00014-91-J-1445, with supplementary support through NOARL (code 211). The program of bioacoustic studies that provided much of the previous work resulting in our acoustic recordings of marine life was also supported for a considerable period by the Oceanic Biology Program of ONR. A wide range of other research programs over this long period also have contributed to our work at sea and in the laboratory, providing understanding of the acoustic behavior of these marine species.

INTERRELATED DATABASES

The textual data on marine animals and their sounds are organized in a series of interrelated computer text databases. These are currently managed with INMAGIC 7.1 software (INMAGIC INC, Cambridge, MA) using PC and AT workstations. The separate, related databases for marine animal vocal behavior include: (1) CETACEA, a comprehensive database of references to the literature on marine animal sounds, (2) SOUND, a database for the descriptive information about each of the sound recordings, (3) SOUNDC, a database for the detailed information and acoustic descriptors for each of the separate digital sound cuts, (4) SOUND2 (SOUND2C, and other series), auxiliary databases for specific research emphases using these same data protocols, and (5) the matching set of digital sound cut files (for SOUNDC, etc.). The digitized sound cuts are stored as separate files on optical disk, and they can be accessed, analyzed (waveform as well as spectrographic portrayals), displayed, and played back from within the text (SOUNDC) databases.

The structures for all of the databases follow similar patterns, use the same systems of annotation, and relate to the same lists of species and geographic codes, etc. Separate, indexed fields within each of the database entries describe the recording situation, the equipment used to make the records, the geographic locations and recording dates, species that were present, behavioral notes, sound types, etc. Records with any of these data or any alphanumeric combination in any field can be rapidly searched and selected. Then, the matching digital sound file can be called up for analysis and display or access for other purposes. Auxiliary databases provide organization of data records and sound files for particular research projects on animal sounds (dolphin mother-calf signature whistles, beluga vocal behavior, sperm whale coda patterns, etc.).

A display-edit-export function for the sound files has been realized with a program designed (by Fristrup) for accessing these files from within the text databases. This is a pop-up utility for retrieving the stored digital sound file, and analyzing and displaying waveforms and color FFT spectrograms from within the text database. Cursors provide for measurement of signal time and frequency, and select portions for expansion or transfer.

Marine Animal SOUND Database

Each digital sound file has a header, containing a portion of the ASCII text from the related (SOUNDC) text database record. This includes text information from the original database for the library collection of recordings, SOUND, thereby allowing the sorting and retrieval of tape information in the SOUNDC and SOUND2C databases. The identified digital sound files are then retrieved from within these text databases, retrieved directly from storage, such as from optical disks.

Equipment --

Sounds from the library collections of marine animal recordings are reproduced on a variety of reel-to-reel, VCR, and cassette equipment. As possible, the spectra of the original recordings are reproduced so as to provide corrected frequency response within the original system bandwidths.

Sounds are analyzed with a variety of equipment, including Kay DSP Sonagraph 5500 (Kay Elemetrics, Pinebrook, NJ), WHOI VOICE analyzer (Martin, Catipovic, Fristrup, and Tyack 1990), and other special purpose routines adapted to specific requirements (most by Fristrup). Digital conversion for sound cuts is performed at more than twice the highest frequency of the sequences of interest. PC and AT work-stations (optimized 286, 386, and 486 computers) are used for the database operations.

The digital sound cut files are ultimately stored on optical disks (Maxtor OC-800), and large capacity hard disk drives are used for temporary storage of sound files during database manipulations.

CETACEA --

The comprehensive, background database of references to the literature, CETACEA, emphasizes information on marine animal sounds. The data in more than 4200 records (to date) are indexed and sorted by 300 subjects, 150 species categories, and a variety of other indexed notations including dates, locations, sound spectral characteristics, environmental observations, etc. A unique feature of these databases is the direct connection between species and all other indices, including subjects, locations, observation dates, notes, etc. In addition, codes have been adapted for ease in identifying and linking of data in various fields. Searching the database is rapid, using a wide variety of simple and complex Boolean strategies. The CETACEA database and its operations have been described in two reports (Watkins, Bird, Moore, and Tyack 1988; and Watkins, Daher, and Haley 1990).

The CETACEA database and the related SOUND databases use an adaptation of INMAGIC software (INMAGIC Inc., Cambridge, MA) for the text data indexing, sorting, and searching routines. For each database, the programs use three interactive files: structure, index, and data records.

SOUND --

The database for the marine animal library collections of analog acoustic recordings organizes the descriptive information about each of the sound records. SOUND databases (including auxiliary databases such as SOUND2, etc.) index and sort the data for approximately 2500 recordings to date. These databases continue to grow as the collections increase in size and complexity. Separate databases are established for particular research needs (for example, one database is currently used for separate organization of the collection of VCR and cassette recordings of dolphins, and other related databases are for specific data on digitized cuts of signature sounds, etc.). Unique retrieval numbers for each record also provide the date and sequence of recording (for example, 83093 indicates the 93rd recording cataloged for the year 1983), and these retrieval codes are used in indexing and shelving of the library collections. Data about the recordings are sorted by 150 species, 30 indexed subjects that are also directly connected to species, in addition to a variety of other information, including recording locations, dates, number of animals, sound types, ten categories of spectral data, etc.

SOUNDC --

The database for organizing the digital sound cuts taken from the analog library recordings. This is a text database with structure and design similar to the other SOUND databases, used in the same ways to index and sort the data appropriate to the individual digital sound cut files. To date, approximately 2200 sound cut files have been indexed in the SOUNDC (and SOUND2C, etc.) text databases. The sound cut databases relate specifically to the sequences of sound that have been excerpted from the library tapes. Unique retrieval numbers for each sound cut file identify the original recording as well as the sequence of digital cuts made from the original (for example, 83093034 identifies the 83093 library recording and this as sound cut number 34). Added details indexed in the SOUNDC database include particulars, such as reference to library tape cue, channel numbers analyzed, duration and sampling rate for the digital file, sound comparison characters, etc. These data are sorted by 150 species, 30 indexed subjects that are directly connected to species, in addition to recording location, dates, number of animals, sound types, etc.

The digital sound cut files selected by searching the SOUNDC database may then be analyzed and displayed on screen without leaving the database by means of the pop-up display utility, IPLOT.

Digital Sound Files --

The digital sound cut files are stored as independent files, identified by retrieval numbers. The first 512 bytes of each file include the indexing information as a header for that sound cut file. The related SOUNDC database provides the flexible search, selection, and retrieval capabilities for accessing these files from within the database. Digital sound cut files may be created for a number of purposes, including that of illustration of a particular type of sound, detailing the repertoire of individual species, analyzing sounds related to certain behaviors, comparing sequences of signals from different species, separating distinctive calls from different populations, or demonstrating variation with season and activity, etc. The SOUNDC database system permits indexing, sorting, and retrieval of the files which then may be immediately analyzed and displayed as desired.

ANALYSIS AND DISPLAY

The digital sound cut files identified by searching the SOUNDC database may be selected for later analysis, or they may be immediately analyzed and displayed without leaving the SOUNDC database. Immediate access to the digital acoustic data is provided by a pop-up utility, IPLOT, developed by Fristrup. IPLOT retrieves the digital files identified by the current search results from SOUNDC. Then, the digital file is analyzed with both waveform and spectrogram displayed simultaneously on screen. The analyzed signals may be edited, re-analyzed, and expanded as desired. Or, the text search results may be re-examined and other signals rapidly retrieved for analysis and display. These digital sound files may also be reformatted and exported by IPLOT as needed. A similar stand alone analysis program, NSIG, may be used to analyze and display digital sound files directly, without interaction with the database.

All of these databases, file handling routines, signal analysis and display programs are related. They utilize common data structures, field identifiers, search and sorting strategies, and display parameters. In addition to the primary databases and their analysis systems, other auxiliary databases are also maintained, such as SOUND2 and SOUND2C which organize data collected for specific bioacoustic projects and analyses of particular acoustic behavior.

INMAGIC PROGRAM

The SOUND databases are currently used with an adaptation of INMAGIC software (version 7.2, INMAGIC Inc., Cambridge, MA) for text databases. This is a flexible text database system that has proven to be relatively simple and easily searched, while retaining the needed complexity of relational association of indexed data. A unique feature of the data association for these databases is the direct connection between species and the other indexed subjects. The software allows every component or alphanumeric entry in the record to be available for searching by a wide variety of simple and complex Boolean strategies.

The INMAGIC text database system includes the following features: (1) permits records of any length, (2) relates 75 or more fields, (3) provides for unlimited numbers of defined subfields, (4) indexes and sorts fields and subfields, (5) indexes as each data record is entered, (6) allows independent sorting and retrieval of data in subfields as well as fields, (7) supports search strategies developed with Boolean operators (and, or, not) and nested arguments, (8) uses searches with qualifiers (greater than, less than, equal to, from/to), (9) provides for convenient right-hand truncation in search statements, (10) saves and combines search results, (11) allows use with user-defined formats for display or reordering of data, (12) prints any number of selected records in any of these formats, (13) lists any indexed terms or fields and subfields with their frequency of occurrence, (14) permits the use of extended characters in records, (15) provides for development of flexible on-line thesaurus of terms, search operators, and definitions for help in searching the records, (16) permits rapid copying of data records, and (17) allows importation of ASCII records created elsewhere. Records in the CETACEA database, for example, are indexed by more than 300 subjects, 150 species categories, and a variety of other indexed notations including dates, locations, sound spectral characteristics, environmental observations, etc. With approximately 5000 records in the CETACEA database, searching and retrieval of any records or combination of records are rapid, usually less than a second.

The databases may be searched by any combination of indexed or unindexed alphanumeric notations. Detailed searches may be made using specific indexed fields, such as genus/species (searchable by order/suborder and family as well). Searching may use any combination of terms and text words or even stems of words or partial phrases and parts of any alphanumeric entry. In addition, codes have been adapted for ease in identifying and searching species, subjects, geographic areas, etc. The alphanumeric coding of marine animal species allows indexing, sorting, and retrieval of most subject fields, geographic locations, dates, and events in direct relationship to specific species. Detailed searches may be made using genus/species, record number, identification, age, gender, observation date, geographic locations (including area names and latitude and longitude), sound type, sound spectral data, author of recording, etc.

Detailed descriptions of search operations and potential combinations of search statements appropriate to these databases are reviewed in abbreviated form in the CETACEA database reports (Watkins, Bird, Moore, and Tyack 1988, Watkins, Daher, and Haley 1990) and in more software detail by the INMAGIC MANUAL (Version 7.2 for MS-DOS, INMAGIC INC, 2067 Massachusetts Ave., Cambridge, MA 02140-1338).

These search features include simple and complex BOOLEAN relations (such as, equals, less/greater than, less than or equals, starts with, contains stem, from...to, etc.). Codes assigned to species and subject categories provide direct associations of most indexed fields with species, for example. This allows generalized or more and more specific searches by truncation of the codes (refer to Organization to Species List).

The design of the database takes advantage of INMAGIC's system for right-hand truncation, so that the placement of codes at the end of fields allows searches by codes or elements of codes, as well as by the other record data. Although a bit slower, searches are also possible using any alphanumeric combination contained in any record, whether in indexed or unindexed fields.

ORGANIZATION OF DATABASE RECORDS

The database records are organized for convenience in entering the data, for relative ease of utilization of the information in records, and for reorganization of the data for display.

The organizational detail is indicated by the list of field names in the database structure and descriptions of these fields (next sections). In brief, the records are organized as follows:

-- Each record is given a unique retrieval number (RECNO). This number gives the year of recording, the recording series, and the sequence of digital analysis. In addition, it serves as a pointer for identification of the recording in the library collection.

-- Separate fields give the recording cue for specific sounds, number and sequence of channels, equipment, author, sampling rate, and duration of the sound sequence.

-- Separate fields also are used for entering a variety of data about the animals and the particular vocal sequence, such as activity, animal identity, age, interaction, sound class and type.

-- Genus/species names and alphanumeric codes are entered for all species recorded -- order/suborder, family, genus, and species are indicated by each code.

-- Codes for genus/species are appended to data in many other fields for direct association with location, observation date and time, sound type, etc.

-- Location for the recording is given by name, by geographic code, as well as by latitude and longitude.

-- Notes, and annotations may be included, and are related to species.

-- Data on the recording situation are also entered, such as hydrophone depths, recording conditions, recording bandwidth, and received signal levels.

-- In addition, an array of measurements of sound features are indexed to provide statistical comparisons of sounds in the databases. These are variable, depending on the database and its intended uses (current entries in SOUNDC are classified by ten variables in temporal and frequency features).

Most fields in the SOUND database records are indexed and may be searched separately or in combination to provide rapid selection of these records. Notes and similar fields are not indexed, but these too may be searched (more slowly) for any words, phrases, or alphanumeric notation.

The List of Species is provided at the end of this report to assist in identification of interrelated codes and to aid in database searches.

STRUCTURE FOR SOUND DATABASES
(August 1991)
INMAGIC Program Format

LABEL	NAME	INDEX	SORT	EMPHASIS
RN	RECNO	T	3	1
CU	CUE	T	3	1
NC	NOCHAN	T	3	1
SR	SAMRATE	T	1	1
CS	CUTSIZE	T	1	1
PL	PLAYBAK	Y	7	1
SC	SIGCLAS	Y	5	1
ID	IDENT	Y	7	1
AG	AGE	Y	7	1
IA	INTERAC	Y	7	1
GS	GENSP	Y	5	1
GA	GEOA	Y	5	1
OD	OBSDATE	Y	4	1
NT	NOTE	N		
DA	DATE	T	4	1
IP	IDPRES	Y	7	1
AP	AGEPRES	Y	7	1
BH	BEHAV	Y	5	1
OS	OTHERSP	Y	5	1
NA	NOANIM	T	3	1
GB	GEOB	Y	5	1
GC	GEOC	Y	7	1
OT	OBSTIME	Y	7	1
SH	SHIP	Y	5	1
AU	AUTHOR	Y	5	1
LO	LOCATE	Y	5	1
HY	HYDEPTH	T	3	1
RC	RCOND	N		
RG	RGEAR	N		
RB	RBAND	Y	7	1
SL	SIGLEVL	Y	7	1
ST	SIGTYPE	Y	5	1
TD	TIMEDUR	Y	7	1
TS	TIMESIG	Y	7	1
TC	TIMECYC	Y	7	1
TA	TIMEAMP	Y	7	1
FR	FREREF	Y	7	1
FB	FREBAND	Y	7	1
FI	FREINST	Y	7	1
FT	FRETREN	Y	7	1
FS	FRESKEW	Y	7	1
SP	SIGPAT	Y	7	1

Abbreviated List of Fields -- SOUND Databases

RN	Retrieval number of record year/tape#/cut#)
CU	Cue or time on tape, buffer (B) size, sec in buffer
NC	Number channels recorded, digitized, chan/side ID
SR	Sample rate (convert Kay input freq. to sample rate)
CS	Cut size -- digital cut in sec (2 or 3 dec. places)
PL	Playback equip., filter setting low (L) high (H) kHz
SC	Signal class - Signature, Mimic..., qual 1-5, Overlap
ID	ID of individual vocalizing, species code
AG	Age, sex (M/F) prefix, birth year, ID, species code
IA	Interaction (MC = male-calf, etc.), ID
GS	Genus/Species animals producing sounds, species code
GA	Geographic location A = ASFIS codes, species codes
OD	Observation date of original recording, species code
NT	Note = Species code, observation or recording details
DA	Date of this record entry (latest modification)
IP	Identification of conspecifics present, species code
AP	Age conspecifics present - sex prefix, birth year, ID
BH	Behavior of the animals, species code
OS	Other species present, species codes
NA	Number of animals vocalizing, species code
GB	Geographic location B = name of area, species code
--/2	Location of birth/capture area, species code, (ID)
GC	Geographic location C = lat. & long., species code
OT	Observation time of original recording, species code
SH	Ship/cruise, aquarium, or recording platform
AU	Author, originator of the recording
LO	Location of original recording
HY	Hydrophone depth in m
RC	Recording conditions, weather, salinity, etc.
RG	Recording gear, equipment
RB	Recording bandwidth, lowest and highest (kHz)
SL	Signal level (dB received or source), species code
ST	Signal type codes, species code
--/2	Signal type names (coda, slow clicks), species code
TD	Time duration of analysis window
TS	Time signal duration (weighted RMS)
TC	Time cycle, duty cycle (area/peak)
TA	Time amplitude variable (attack and decay)
FR	Frequency reference center (aggregate)
FB	Frequency bandwidth (aggregate)
FI	Frequency instantaneous average (binsize)
FT	Frequency trends
FS	Frequency skew
SP	Signal pattern, occurrence sequence, repetition

SOUND Database Field Description --
(August 1991)

LABEL -- NAME -- Description with example.

RN	RECNO	Retrieval number of record (Year/tape#/cut#) Year (2 digits), tape no. (3 digits), cut # (3 digits, alphanumeric base 36). Example: 850430B6
CU	CUE	Cue or time (min:sec) on tape at signal end, (B) analyzer buffer size (min:sec), decimal time sec from start of Kay buffer to cursor at beginning of sound. Example: 542 B2:8 8.130 or 1:03:12 B2:8 8.130
NC	NOCHAN	Number of channels recorded (1st digit), number of channels multiplexed (2nd digit), channel ID letter(s) -- side one A & B, side two C & D. Example: 42CD
SR	SAMRATE	Sample rate (Kay input frequency converted).
CS	CUTSIZE	Cut size -- sec (2 or 3 decimal places).
PL	PLAYBAK	Playback recorder/filter type, settings in kHz (pass-band), L = low (HP) and H = high (for LP). Example: PEMTEK/KROHN-HITE L0.1 H12
SC	SIGCLAS	Signal class, letter codes (Signature, Mimic, Variant, Deletion, Uncharacteristic, Calf, etc.), quality 1 to 5 (best), O = overlap.
ID	IDENT	Identification of vocal animal, species code.
AG	AGE	Age, sex prefix (M or F), birth year (decimal for part year) of vocalizing animal, ID. Example: M3FB19 M1985FB19
IA	INTERAC	Interaction, male-male (MM)/ female-calf (FC), etc., with ID's. Example: FCFB10 FCFB19
GS	GENSP	Genus and species of vocalizing animal. Scientific names and species code.
GA	GEOA	Geographic location A, geographic area code from ASFIS map, species code. Example: ANWAB1A (ANW area, AB1A species)

Fields in SOUND Structure (continued)

OD	OBSDATE	Observation date for original, species code. Example: 20-Dec-1960 AB1A
	OBSDATE/2	Subfield for month and year, species code. Example: DecAB1A 1960AB1A
NT	NOTE	Species code, comments, recording details. ID of animals in parentheses. Not indexed.
DA	DATE	Date of this record entry, last modification. (Use F6 key) Example: 24-May-1991
IP	IDPRES	ID of conspecifics present, species code. Separate subfields for each animal.
AP	AGEPRES	Age of conspecifics present = sex prefix (M/F) with age (decimal) and ID, also with birth year, ID. Separate subfields each animal. Example: M0.3FB19 M1985FB19
BH	BEHAV	Behavior codes, species code.
OS	OTHERSP	Other species present, species codes. Separate subfields for each species.
NA	NOANIM	Number of animals vocalizing, species codes.
GB	GEOB	Geographic location B, area name, species code.
	GEOB/2	Location of birth/capture, species code, ID of animal (in parentheses).
GC	GEOC	Geographic location C, latitude and longitude, "N" or "S" and two digits for lat., "E" or "W" and three digits for long. Each with species code. Example: N70AB1A E020AB1A
OT	OBSTIME	Observation time for original, species code. Example: 1430 AB1A
SH	SHIP	Ship/cruise, aquarium, or other platform, (SMall boat, AQuarium, BEach, ICe, UNKNOWN).
AU	AUTHOR	Author, originator of the recording.
LO	LOCATE	Location of original recording.
HY	HYDEPTH	Hydrophone depth in m.

Fields in SOUND Structure (continued)

RC RECCOND Conditions, weather, salinity. Not indexed.

RG RECGEAR Recording gear, equipment. Not indexed.

RB RECBAND Recording bandwidth, lowest and highest kHz.
Example: L0.02 H30 (20Hz-30kHz).

SL SIGLEVL Signal level Received, Source, (dB), species code. E = estimated. Example: S172AB1A E

ST SIGTYPE Signal type, Long= > 0.1 sec; Short= < 0.1 sec
- BL Broadband long, noisy, many frequencies
- BS Broadband short, clicks, pulses
- NL Narrowband long, tonal, few harmonics
- NS Narrowband short, bell, short tones
- FM Frequency modulated, sweeps, "contours"
- CH Chirp, short FM sweep
- PU Pulsed, click bursts, sidebands
- SE Series, train, sequence of sounds
- SO Song, long, repetitive, patterned

-- with species code. Example: NLAB1A

SIGTYPE/2 Signal type, general: coda, slow clicks, etc
-- with species code.

[Remaining statistics for sound feature analysis.]

TD TIMEDUR Time duration of analysis window.

TS TIMESIG Time signal durations (weighted RMS).

TC TIMECYC Time duty cycle (area/peak).

TA TIMEAMP Time amplitude variable (attack and decay).

FR FREREF Frequency reference center (aggregate).

FB FREBAND Frequency bandwidth (aggregate).

FI FREINST Frequency instantaneous average (binsize).

FT FRETREN Frequency trends.

FS FRESKEW Frequency skew.

SP SIGPAT Signal pattern of occurrence, repetition.

Marine Animal SOUND Database

GEOGRAPHIC LOCATIONS

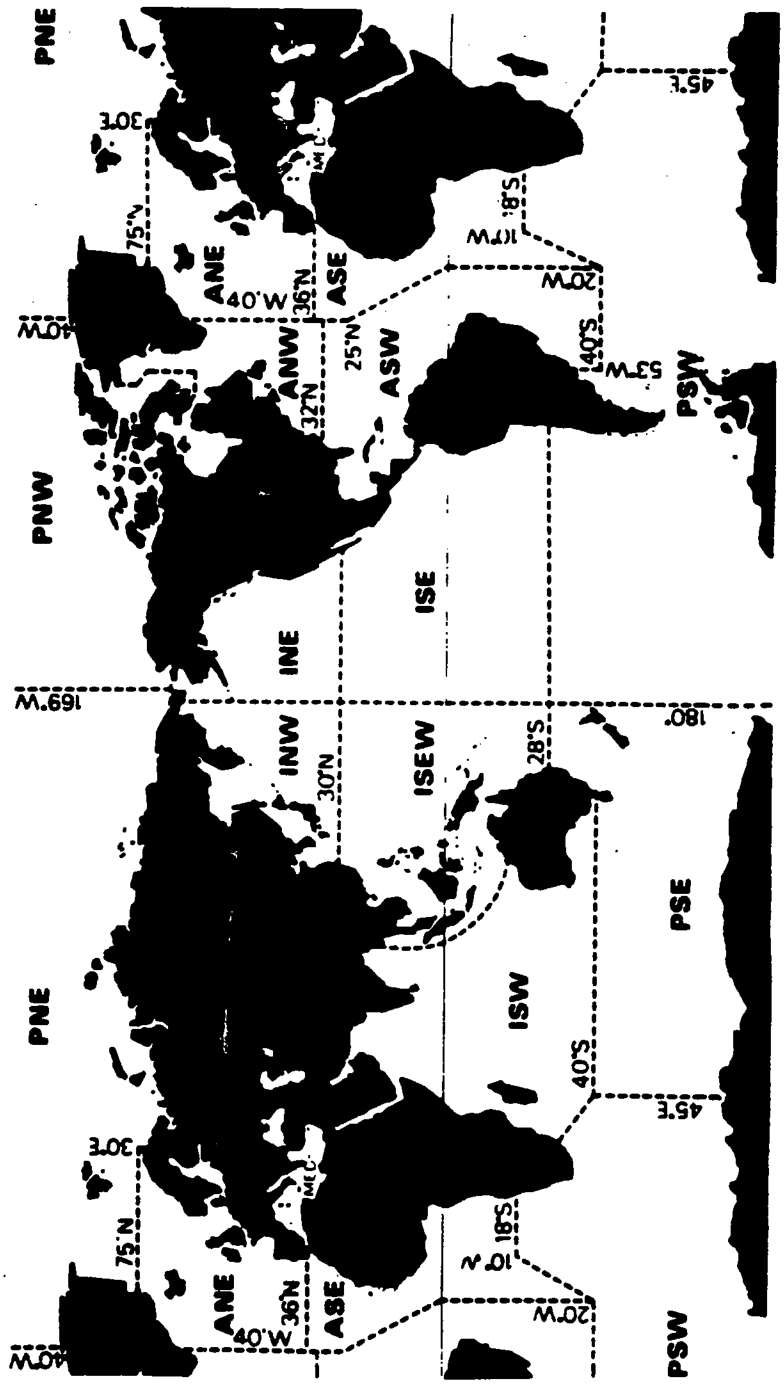
GEOA field -- The first of the geographic location fields in the structure listed above uses codes for the locations or places at which the original recordings were made. These location codes are adapted generally from the geographic codes used by the Aquatic Sciences and Fisheries Information System (Anon. 1980, ASFIS):

ANE	Northeast Atlantic
ANW	Northwest Atlantic
ASE	Southeast Atlantic
ASW	Southwest Atlantic
INE	Northeast Pacific
INW	Northwest Pacific
ISE	Southeast Pacific
ISC *	Southwest (central) Pacific
ISW	Indian Ocean
MED	Mediterranean
PNE	Eastern Arctic Ocean
PNW	Western Arctic Ocean
PSE	Eastern Antarctic Ocean
PSW	Western Antarctic Ocean
CSL *	Coastal Waters
FSR *	Freshwater
COS *	Cosmopolitan

* changes from the ASFIS codes

The species codes are combined with the ocean area codes. For example, a recording of <u>Megaptera</u> <u>novaeangliae</u> in the Indian Ocean is indexed with the codes ISWAC2A (ISW for the Indian Ocean, and AC2A for <u>Megaptera</u> <u>novaeangliae</u>).

ASFIS Map -- Aqua. Sci. Fish. Info. Sys. (Anon. 1980, ASFIS).



TEXT DATABASES

The use of text-based systems has allowed rapid association of our different existing data sets dealing with marine animals and their sounds. This includes the CETACEA database for literature references and the variety of observational data accompanying animal sightings, as well as the library of oceanic recordings of animal sounds and the SOUND cut databases. These text based databases perform admirably for organization, cataloging, and sorting of such discrete pieces and sequences of information.

Although these text databases have provided the descriptive detail that is needed for organization of the material, we anticipate future transitions to more interactive, relational systems for use in extensive comparisons of the sound data from large components of the acoustic files. Even so, these fundamental relationships between textual records about the sound recordings and the behavioral descriptors relative to the acoustic data will necessarily continue to be the same, regardless of the database engines that are employed. The text-based data will always be significant, and the means for intelligent organization of these data will continue to be important.

USING THE SOUND DATABASES

The following instructions for using the text SOUND database program provide both a quick, step-by-step guide to its use, and they show the basic simplicity of database operations. The current database organization is designed for use with software by INMAGIC INC. (Cambridge, MA).

1. Call up the directory containing the INMAGIC software and enter the database by typing INMAGIC.

2. Choose SELECT from the Main Menu for searching the database, for printing or displaying entries.
-- or choose MAINTAIN from the Main Menu for work within a database to edit old records or to add new ones.

3. Indicate the database that is desired, such as SOUND, SOUNDC, SOUND2, or SOUND2C.

4. Passwords may be indicated as required.

5. And then, searches and retrieval of records may be initiated.
-- or records may be called up for modification or new records entered into the database.

Any number of indexed fields may be sorted and retrieved rapidly (much less than a second), and even unindexed fields, such descriptive notes, can be easily searched for any alphanumerical string. The time required for searching unindexed fields varies with the amount of stored data, about a minute for a 12 Mbyte search on our current PC's.

6. The records that are selected may be displayed, sorted, printed, or exported as desired (sequentially through the list of selections). Digital sound files for each text record may also be accessed from within the search display, then analyzed by waveform and spectrogram as well as played back (digital to analog conversion) on loudspeakers.

7. It is important to EXIT properly during any maintenance work on the databases, because of potential disruption of indexing operations that could scramble the data.

These operations are similar to those described for the literature reference database, CETACEA (Watkins, Daher, and Haley 1990). See also the INMAGIC Manual for detailed descriptions of program functions and operations.

SEARCHING -- TEXT DATABASES

The SOUND database systems (with INMAGIC software, Cambridge, MA), provide for organization and rapid handling of the textual data associated with the library of acoustic recordings of underwater sounds from animals. Indexed database fields are rapidly searched by specifying field labels or names by simple and by a variety of complex search strategies. Records (in any sequence) may be displayed, printed, or exported in a wide variety of formats defined by the user. Unlimited subfields may be indexed and searched in the same ways to provide extremely flexible and detailed sorting and retrieval of these text records. The INMAGIC Manual (INMAGIC INC 1990) describes the commands and relationships that are available and that can be combined for searching these databases and for displaying the data.

Unindexed fields may be similarly searched for any word, phrase, or alphanumeric string, but at a slower rate.

Search commands may be combined with search relations to provide complex Boolean command sequences (including, excluding, broadening), word and word-stem searches, comparative searches (equals, starts with, greater/less than or equals, from...to), etc. Searchable, sorted data can be listed in a wide variety of ranges and formats.

The flexibility of the database software search system is combined with a system of coding assigned to the data. Codes are appended to each species, behaviors, sound types, etc. Therefore, the codes may also be used as a rapid search strategy. In addition, the species codes are used to make sure that different subjects, dates, places, etc. are always directly linked with species. Species codes have been assigned to each species by (1) order/suborder, (2) family, (3) genus, and (4) species. For example, a finback whale (*Balaenoptera physalus*) has the code "AC1F" -- indicating suborder Mysticeti, family Balaenopteridae, genus *Balaenoptera*, species *physalus*. See section on Organization of Species List (p. 41).

DIGITAL SOUND FILES

Sequences of sound (cuts) from the repertoire of vocalizations of the different species of marine animals are selected from the original underwater recordings and digitized. These 1 to 30 sec digital sound cuts are stored as separate files, and they are organized, indexed, and sorted by related text databases (SOUNDC, SOUND2C, etc.). The first 512 bytes of each digital sound cut file is a binary/ASCII header, including data from the SOUNDC database record for this sound sequence. This annotation includes such detail as the tape cue for the location of the sound sample in the original recording, sampling rate, length of cut, observation date and place, species, associated behavior, etc. -- usually all of the database fields through "Note" (see sections on Data Structure on p. 18 and on Database Fields pp. 19-22). The digital sound cut files are labeled with an extension of the tape number from the original library recording to maintain continuity of the data (including year of recording, tape number, and number of this sound cut from this tape), as noted below for the sound cut numbering convention.

Example: <u>8821120F</u>

88211	The complete library recording number.
<u>88</u>	The year of the original recording.
<u>211</u>	The recording number for that year.
<u>20F</u>	The digital sound cut number.
8821120F	The complete digital sound cut number.

All three places of the sound cut extension number use a 36-base character set -- employing numbers 0-9 and letters A-Z, in this order (20A follows 209, 5F0 follows 5EZ).

A different character is used for the first place of the sound cut extension (last three characters) to designate different series, channels, animals, sound types, etc.

The sound cut extension number <u>000</u> is not used (88211 and 88211000 are considered to be the same by INMAGIC).

Digitizing the Sound Cuts --

A variety of systems may be used to provide the detailed scrutiny of the analog library recordings, select sound sequences of interest, and convert the sounds to digital format. Two signal analyzers used most extensively for creation of the files for the SOUND database include VOICE (a spectrogram computer display package by Martin, Catipovic, Fristrup, and Tyack 1990) and the KAY DSP Sona-graph 5500 (Kay Elemetrics, Pine Brook, NJ). The analog library recordings are played back on appropriate equipment, and the sound is monitored aurally by good quality loudspeakers and visually by the real-time waveform and spectrographic displays of the signal analyzers. To provide an indication of the methodology for selection and digitizing of sound cuts, procedures and operations are listed using a PC computer, tape playback, and a KAY (DSP):

1. Computer -- call up the INMAGIC program.
   a. Choose SELECT from the Main Menu.
   b. Select the appropriate database (SOUND or SOUND2).
   c. Request the record number for the tape to be digitized. (Ex: q RN = 65005).
   d. Write (copy) this record into an ASCII file using the annotation format, ANN (Ex: w in 65005.asc u ANN).

2. Computer -- call up the KAY (K) File Management Program.
   a. From the Main Menu select "System".
   b. Choose "Data Path" and indicate Drive and Path for downloading of digitized signal files.
   c. Choose "Interface" and select "SCSI".
   d. Choose "I/O" and select "From 5500".
   (To open "Data Transfer Menu"):
      1. "File Format" = 5500.
      2. "Number of Bits" = 8, 10, 12 or 16.
      3. "Channel" = 1 or 2.
      4. "Source" = Buffer, Between Cursors, Screen, or Highlighted List.
      5. "Name of File" = Sound cut number.KAY
         Digital filename (Ex: 65005001.KAY).
         Leave cursor blinking until download.

3. Playback -- Mount and play selected recording.
   Locate sound sequence for digitization.

4. KAY -- set the analyzer for real time analysis.
   a. Copy sound onto KAY buffer and scroll through (waveform and spectrographic displays).
   b. Locate signal within parameters to be downloaded (Buffer, Between Cursors, Screen).

5. Computer -- still in the KAY File Management Program:
   a. Press [Ret] to accept the KAY file selected (named above with .KAY extension).
   b. Press [F2] to transfer digital signal from the KAY to the designated Drive and Path.
   c. Exit KAY program if no more are to be transferred or remain in KAY program for more signals.

6. Computer -- call up SideKick Plus (TSR editor utility sharing KAY program on screen).
   a. Access the ASCII text file created in #1 above.
   b. Rename this record, giving it the same number as the digital sound cut (in #2) but TXT extension.
   c. Update the text information appropriate to sound cut (digital file just downloaded #5).
   d. Save the changes made to this file (press [F2]) and exit SideKick Plus program.

7. Computer -- call up the MASSEDIT program (to attach header information to digital files).
   a. Indicate the ASCII file (made in #6) containing the updated text information for the sound cut files.
   b. Designate the Drive/Directory of the digital files (Ex: MASSEDIT 65005.TXT D:).

8. Computer -- call up the INMAGIC database program (to enter sound cut records into the database).
   a. Choose MAINTAIN from the Main Menu.
   b. Select the appropriate database (SOUNDC or SOUND2C)
   c. ADD the ASCII text file modified in (#6) SideKick (Ex: A 65005.TXT) to index the new edited data records into the sound cut text database.

9. Computer -- from within the INMAGIC database program:
   a. Search the SOUNDC (sound cut) text database for the sound sequence of interest (see Searching the Text Databases, p. 29).
   b. Display the text data for the sound sequence.
   c. Analyze this sound on screen (a key stroke) with both waveform and spectrographic analyses.
   d. Return to text data for this sound (a key stroke).
   e. Listen to this sound, D-to-A conversion and playback of the sound (a key stroke).

PROGRAMS FOR THE SOUND DATABASES

A variety of programs have been designed or adapted for work with the SOUND databases and their associated digital sound cut files. These are listed and described briefly below.

INMAGIC	--	Text database index and search.
SIDEKICK PLUS	--	Text data editor and transfer.
HEADEDIT	--	Edit annotation of individual files.
MASSEDIT	--	Annotation of one or more sound files.
KAYCHECK	--	Check data in file headers.
NSIG	--	Analyze digital sound file.
IPLOT	--	Access and analysis of sound.
DA	--	Access and D-to-A playback of sound.

INMAGIC is the current software for organizing, indexing, sorting, and retrieving records containing information about the library collection of sound recordings, and the digital sound cuts. Program by INMAGIC INC, Cambridge, MA.

SIDEKICK PLUS is a pop-up utility (TSR) program used to access and modify text records, for example, from the SOUND database to provide appropriate text data for the sound cut files, indexed in the SOUNDC databases. Program by Borland International, Scotts Valley, CA.

HEADEDIT program allows display and editing of ASCII text header on individual digital sound files (usage: HEADEDIT FILENAME.KAY). The binary data written by the download program for the KAY (DSP 5500 Sonagraph) are also displayed but are not available for editing, including number of samples, sampling rate, and number of bits per sample. HEADEDIT may be used to read the header data, change existing header text, or enter new data. New text data for the header may be entered as blocks of new data by means of a utility program such as SK PLUS, or text may be typed in place directly with HEADEDIT. Program by Kurt Fristrup.

MASSEDIT program is used to insert ASCII text into headers on one or a number of files at one operation (HEADEDIT works with one file at a time). This text data is inserted into the header space, the first 512 Bytes of each digital sound file. This header space and some identifying binary code is provided, for example, by the KAY download program of the Kay Sonagraph (DSP 5500). The header data typically includes the first fields of the text record through at least the NOTE field, allowing direct connection of the information about the recording to the appropriate digital sound file. MASSEDIT matches the record retrieval numbers and ASCII data sets of a specified file (with .TXT extension) to the appropriate files (with .KAY extension) in a specified directory (Usage: MASSEDIT FILENAME.TXT D:). See example in #7, p. 32. Program by Kurt Fristrup.

KAYCHECK is a comparison program that searches for every .KAY file (format for KAY download program) in the current directory, and checks the header text data (first 512 bytes on the digital file) for accuracy of filename, sample rate, file size, etc. Files that are found to have discrepancies or are without a proper header are identified, with date, time, and size of file indicated. The output of KAYCHECK may be written to another file (usage: KAYCHECK>FILENAME) for printing or editing. Program by Kurt Fristrup.

NSIG program provides access, analysis, and display of files of digitized sound (independent of databases). NSIG accesses the specified sound file, analyzes, and displays waveform (top of screen) and concurrent spectrographic analyses (bottom of screen). Frequency and time scales and variable cursors are provided. Portions of the displayed analysis may be selected and re-analyzed. All or any part of the data file may be exported. See section below on Operating Commands, p. 36. Program by Kurt Fristrup.

IPLOT is a TSR analysis program used for accessing and displaying the digital sound file from within the SOUNDC database. IPLOT provides the same types of analysis and display as NSIG, above. This permits immediate analysis of each of the sound cut files whose text data record is displayed on screen by the database. IPLOT is loaded as a TSR utility before database selections are made, then file access and analyses are initiated (a key stroke) from within the database. At anytime, the sound analysis display may be immediately returned to the text data record, re-analyzed, or any segment of the display expanded for further analysis. Any sequence or number of files selected by the database may be accessed, analyzed, and displayed in turn. All or any part of the data file may be exported. See section below on Operating Commands, p. 38. Program by Kurt Fristrup.

DA program provides access to the digital sound files, D-to-A conversion, and then playback of the analog signal over a loudspeaker system. DA is loaded and initialized in similar ways to IPLOT. Sound playback is initiated directly from the digital sound file or (by a key stroke) from within the SOUNDC database. DA adds another dimension to the sound analysis by allowing the sound sequence selected by the SOUNDC database to be listened to -- as well analyzed and displayed by IPLOT. Program by Terrance Howald.

Operating Commands for IPLOT and NSIG --

IPLOT loaded as TSR at start of database search.

From the text database display:

[Alt] [Esc] invokes IPLOT,
accesses the selected digital sound file,
and draws signal waveform and spectrogram.
[Esc] ends analysis display.
[R] re-displays the database text.

From within the analysis display:

[Ins] increases dynamic range 3 dB each time,
brings up relative level of background.
[Del] decreases dynamic range 3 dB each time,
reduces relative level of background.
[Ctrl] [Home] increases analysis "attenuator",
raises gain of signal.
[Ctrl] [End] decreases analysis "attenuator",
lowers gain of signal.

[Alt] [F] toggles spectrogram display.
[Alt] [N] toggles noise compensation.
[Ctrl] [PrtSc] saves data between cursors,
use FILENAME.TXT for ASCII.

[LeftArrow], [RightArrow] time cursors move
left or right.
[UpArrow], [DownArrow] frequency cursors move
up or down.
[Ctrl] [LeftArrow], [Ctrl] [RightArrow]
time cursors move 5X speed.
[Home] <keypad> toggle front/back cursors.

[PgDn] zoom to new analysis display,
bounds set by cursors.
[PgUp] step back to previous display bounds.

LITERATURE CITED

Anonymous 1980. Aquatic sciences and fisheries information system, Aquatic Sciences and Fisheries Abstracts Database User Guide, Aquatic Sciences and Fisheries Information System, NOAA, U.S. Dept. of Commerce, Rockville, MD, 65 pp.

Anonymous 1990. INMAGIC Manual, INMAGIC INC, 2067 Massachusetts Ave., Cambridge, MA 02140-1338, 314 pp.

Martin, Ann, Josko A. Catipovic, Kurt Fristrup, and Peter L. Tyack 1990. VOICE - A spectrogram computer display package. Technical Report WHOI-90-22, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, 95 pp.

Watkins, William A., James E. Bird, Karen E. Moore, and Peter Tyack 1988. Reference database marine mammal literature. Technical Report WHOI-88-2, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, 25 pp., 4 appendices.

Watkins, William A., Mary Ann Daher, and Nancy J. Haley 1990. Documentation for the CETACEA database of marine mammal literature references. Technical Report WHOI-90-19, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, 111 pp.

SPECIES LIST -- SOUND Databases

SPECIES LIST -- SOUND Databases

Organization of the Species List --

The list of species for the SOUND Databases is the same as is used for CETACEA, the literature reference database (Watkins, Bird, Moore, and Tyack 1988; Watkins, Daher, and Haley 1990). The species are arranged alphabetically within families. Alphanumeric codes have been assigned to each species by (1) order/suborder, (2) family, (3) genus, and (4) species. These were assigned initially in ascending (alphanumeric) order, with spaces for taxonomic revision when needed.

The first place of the species code is a letter representing the order or suborder. The letter "A" denotes the suborder Mysticeti and includes 11 species. The letter "B" denotes the suborder Odontoceti and includes 68 species. The letter "C" denotes the order Carnivora, including 37 species of pinnipeds, sea otters, and polar bear. The letter "D" denotes Sirenia and includes 5 species. Other orders and suborders are included in general categories and are coded by letter (E-Z).

The second place of the marine mammal species code is a letter representing the family. For example, the Species List indicates the Balaenopteridae as code "AC" -- or suborder Mysticeti "A" and family Balaenopteridae "C". The code for Ziphiidae is "BC" -- or suborder Odontoceti "B", family Ziphiidae "C".

The third place of the species code is a number of one or two digits representing the genus. For example, the code for Mesoplodon is "BC5" -- or suborder Odontoceti "B", family Ziphiidae "C", genus Mesoplodon "5". The code for Kogia is "BA1"-- or suborder Odontoceti "B", family Physeteridae "A", genus Kogia "1". The code for the genus Stenella is "BD15" (suborder Odontoceti "B", family Delphinidae "D", and genus Stenella "15").

The last place of the species code is a letter representing species. For example, the code for Kogia breviceps is "BA1A" (suborder Odontoceti "B", family Physeteridae "A", genus Kogia "1", species breviceps "A"). The code for the related species, Kogia simus is "BA1B". The code for Cephalorhynchus hectori is "BD1D" (suborder Odontoceti "B", family Delphinidae "D", genus Cephalorhynchus "1", species hectori "D"). Both the scientific names and the genus/species codes are given in the GENSP field -- for example, a reference about finback whales will have "Balaenoptera physalus AC1F" in the genus/species field.

SPECIES LIST -- SOUND Databases

Mysticeti -------------- A

Balaenidae -------- AA

Balaena mysticetus Linnaeus 1758	AA1A
Eubalaena glacialis (Borowski) 1781	AA3A
Eubalaena australis (Desmoulins) 1822	AA3B

Eschrichtiidae ---- AB

Eschrichtius robustus (Lilljeborg) 1861	AB1A

Balaenopteridae --- AC

Balaenoptera acutorostrata Lacépède 1804	AC1A
Balaenoptera bonaerensis Burmeister 1867	AC1D
Balaenoptera borealis Lesson 1828	AC1B
Balaenoptera edeni Anderson 1878	AC1C
Balaenoptera musculus (Linnaeus) 1758	AC1E
Balaenoptera physalus (Linnaeus) 1758	AC1F
Megaptera novaeangliae (Borowski) 1781	AC2A

Neobalaenidae ----- AD

Caperea marginata (Gray) 1846	AD1A

Odontoceti ------------- B

Physeteridae ------ BA

Kogia breviceps (Blainville) 1838	BA1A
Kogia simus (Owen) 1866	BA1B
Physeter catodon Linnaeus 1758	BA2A

Monodontidae ------ BB

Delphinapterus leucas (Pallas) 1776	BB1A
Monodon monoceros Linnaeus 1758	BB2A

Ziphiidae --------- BC

Berardius arnuxii Duvernoy 1851	BC1A
Berardius bairdii (Stejneger) 1883	BC1B
Hyperoodon ampullatus (Forster) 1770	BC2A
Hyperoodon planifrons Flower 1882	BC2B
Indopacetus pacificus (Longman) 1926	BC3D
Mesoplodon bidens (Sowerby) 1804	BC5A
Mesoplodon bowdoini Andrews 1908	BC5B
Mesoplodon carlhubbsi Moore 1963	BC5C
Mesoplodon densirostris (Blainville) 1817	BC5D
Mesoplodon europaeus Gervais 1855	BC5E
Mesoplodon ginkgodens Nishiwaki and Kamiya 1958	BC5H
Mesoplodon grayi von Haast 1876	BC5J
Mesoplodon hectori (Gray) 1871	BC5K
Mesoplodon layardii (Gray) 1865	BC5L
Mesoplodon mirus True 1913	BC5M
Mesoplodon peruvianus Reyes, Mead, Waerebeek 1991	BC5P
Mesoplodon stejnegeri True 1885	BC5S
Tasmacetus shepherdi Oliver 1937	BC7A
Ziphius cavirostris G. Cuvier 1823	BC9A

SPECIES LIST -- SOUND Databases

Delphinidae --------- BD	
Cephalorhynchus commersonii Lacépède 1804	BD1A
Cephalorhynchus eutropia (Gray) 1846(9?)	BD1B
Cephalorhynchus heavisidii (Gray) 1828	BD1C
Cephalorhynchus hectori van Beneden 1881	BD1D
Delphinus bairdii Dall 1873	BD3A
Delphinus delphis Linnaeus 1758	BD3B
Delphinus tropicalis van Bree 1971	BD3C
Grampus griseus (Cuvier) 1812	BD4A
Lagenodelphis hosei Fraser 1957	BD5A
Lagenorhynchus acutus (Gray) 1828	BD6A
Lagenorhynchus albirostris Gray 1846	BD6B
Lagenorhynchus australis (Peale) 1848	BD6C
Lagenorhynchus cruciger (Quoy and Gaimard) 1824	BD6E
Lagenorhynchus obliquidens Gill 1865	BD6G
Lagenorhynchus obscurus (Gray) 1828	BD6H
Lissodelphis borealis (Peale) 1848	BD8A
Lissodelphis peronii (Lacépède) 1804	BD8B
Peponocephala electra (Gray) 1846	BD10A
Sotalia borneensis Lydekker 1901	BD12A
Sotalia brasiliensis Van Beneden 1875	BD12C
Sotalia fluviatilis (Gervais) 1855	BD12B
Sotalia guianensis Van Beneden 1864	BD12D
Sousa chinensis (Osbeck) 1765	BD13A
Sousa plumbea (Cuvier) 1829	BD13B
Sousa tëuszii (Kukënthal) 1892	BD13C
Stenella attenuata (Gray) 1846	BD15A
Stenella clymene Gray 1850	BD15B
Stenella coeruleoalba (Meyen) 1833	BD15C
Stenella frontalis (G. Cuvier) 1829	BD15F
Stenella longirostris (Gray) 1828	BD15L
Steno bredanensis (Cuvier) 1828	BD17A
Tursiops aduncus (Ehrenberg) 1832	BD19A
Tursiops catalania (Gray) 1868	BD19B
Tursiops gillii Dall 1873	BD19C
Tursiops truncatus (Montagu) 1821	BD19D
Globicephalidae - BE	
Feresa attenuata Gray 1874	BE1A
Globicephala edwardii Smith 1934	BE3A
Globicephala macrorhynchus (Gray) 1846	BE3B
Globicephala melaena (Traill) 1809	BE3C
Globicephala scammoni Cope 1869	BE3D
Orcaella brevirostris (Owen 1866)	BE5A
Orcaella fluminalis Anderson 1871	BE5B
Orcinus orca (Linnaeus) 1758	BE7A
Pseudorca crassidens (Owen) 1846	BE9A
Phocoenidae ---- BF	
Australophocaena dioptrica Lahille 1912	BF1A
Phocoena phocoena (Linnaeus) 1758	BF2A
Phocoena spinipinnis Burmeister 1865	BF2B
Phocoena sinus Norris and McFarland 1958	BF2C
Phocoenoides dalli (True) 1885	BF4A
Neophocaena phocaenoides (G.Cuvier) 1829	BF6A

Susuidae ---------- BG	
Susu gangetica Lebeck 1801	BG1A
Susu indii Blyth 1859	BG1B
Inia geoffrensis Blainville 1817	BG2A
Lipotes vexillifer Miller 1918	BG3A
Pontoporia blainvillei (Gervais) 1844	BG4A
Carnivora ------------- C	
Otariidae --------- CA	
Arctocephalus australis (Zimmerman) 1783	CA1A
Arctocephalus forsteri Lesson 1828	CA1F
Arctocephalus galapagoensis Heller 1904	CA1G
Arctocephalus gazella Peters 1875	CA1H
Arctocephalus philippii Peters 1866	CA1P
Arctocephalus pusillus (Schreber) 1776	CA1R
Arctocephalus townsendi Merriam 1897	CA1T
Arctocephalus tropicalis (Gray) 1872	CA1W
Callorhinus ursinus (Linnaeus) 1758	CA2A
Eumetopias jubatus (Schreber) 1776	CA3B
Neophoca cinerea (Peron) 1816	CA4A
Otaria flavescens (Shaw) 1800	CA6A
Phocarctos hookeri (Gray) 1844	CA8A
Zalophus californianus (Lesson) 1828	CA9A
Odobenidae -------- CB	
Odobenus rosmarus (Linnaeus) 1758	CB1A
Phocidae ---------- CC	
Cystophora cristata (Erxleben) 1777	CC1A
Erignathus barbatus (Erxleben) 1777	CC2A
Halichoerus grypus (Fabricius) 1791	CC3A
Hydrurga leptonyx (Blainville) 1820	CC4A
Leptonychotes weddellii (Lesson) 1826	CC5A
Lobodon carcinophagus (Hombron & Jacquinot) 1842	CC6A
Monachus monachus (Hermann) 1779	CC8A
Monachus schauinslandi Matschie 1905	CC8B
Monachus tropicalis Gray 1850	CC8C
Mirounga angustirostris Gill 1866	CC10A
Mirounga leonina (Linnaeus) 1758	CC10B
Phoca caspica Gmelin 1788	CC12C
Phoca fasciata Zimmermann 1783	CC12F
Phoca groenlandica Erxleben 1777	CC12G
Phoca hispida Schreber 1775	CC12H
Phoca largha Pallas 1811	CC12L
Phoca sibirica Gmelin 1788	CC12S
Phoca vitulina Linnaeus 1758	CC12V
Ommatophoca rossi Gray 1844	CC14A
Mustelidae -------- CD	
Enhydra lutris (Linnaeus) 1758	CD1A
Lutra felina Molina 1782	CD2B

SPECIES LIST -- SOUND Databases

Ursidae ------------ CE
- *Ursus maritimus* Phipps 1774 — CE1A

Sirenia ---------------- D

Dugongidae --------- DA
- *Dugong dugon* Muller 1776 — DA1A
- *Hydrodamalis gigas* Zimmermann 1780 — DA2B

Trichechidae ------- DB
- *Trichechus inunguis* (Natterer) 1883 — DB1A
- *Trichechus manatus* Linnaeus 1758 — DB1B
- *Trichechus senegalensis* Link 1795 — DB1C

OTHER MAMMALS
- Primates --------------------- E
- Chiroptera ------------------- F
- Ungulates, *sensu lato* -------- G
- Other mammals ---------------- H

VERTEBRATES
- Aves ------------------------- I
- Reptilia --------------------- J
- Amphibia --------------------- K
- Fish, *sensu lato* ------------- L
- Other vertebrates ------------ S

INVERTEBRATES
- Molluscs
  - Cephalopoda ------------- M
  - Other molluscs ---------- N
- Arthropods
  - Crustacea --------------- O
  - Insecta ----------------- P
  - Other arthropods -------- Q
- Other invertebrates ---------- R

GENERAL
- Fossils ------------------------------ T
- Uncertain (sea serpents and other indeterminate animals) ----- U
- General pinniped --------------------- V
- General cetacean --------------------- W
- Ambient noise (ship, geologic, ice) --- X
- General mammal ----------------------- Y
- Animals in general ------------------- Z

List of Common Names for Marine Animals --

The "common" names given to different species of marine animals vary considerably, with only a few of these names consistently used by English speakers. Other language areas, of course, have their own (variable) set of common names. For accuracy, therefore, the databases always use the scientific nomenclature. However, for cross-reference, we have included the following list of some of the most commonly used names for these animals. The common names are listed in the same order, directly with the scientific names and the species codes used by the databases. Many species are not well known by any but their scientific names -- it is always accectable, and often much easier, to refer to these marine animals by their scientific names. The scientific nomenclature is recognized worldwide, regardless of language.

COMMON NAMES FOR MARINE ANIMALS

Baleen whales (Mysticeti) -- A	
Family Balaenidae ----------- AA	
Bowhead whale -- *Balaena mysticetus*	AA1A
Northern right whale -- *Eubalaena glacialis*	AA3A
Southern right whale -- *Eubalaena australis*	AA3B
Family Eschrichtiidae ------- AB	
Gray whale -- *Eschrichtius robustus*	AB1A
Family Balaenopteridae ------ AC	
Minke whale -- *Balaenoptera acutorostrata*	AC1A
Southern minke whale -- *Balaenoptera bonaerensis*	AC1D
Sei whale -- *Balaenoptera borealis*	AC1B
Bryde's whale -- *Balaenoptera edeni*	AC1C
Blue whale -- *Balaenoptera musculus*	AC1E
Fin, Finback whale -- *Balaenoptera physalus*	AC1F
Humpback whale -- *Megaptera novaeangliae*	AC2A
Family Neobalaenidae -------- AD	
Pigmy right whale -- *Caperea marginata*	AD1A
Toothed whales (Odontoceti) -- B	
Family Physeteridae --------- BA	
Pigmy sperm whale -- *Kogia breviceps*	BA1A
Dwarf sperm whale -- *Kogia simus*	BA1B
Sperm whale -- *Physeter catodon*	BA2A
Family Monodontidae --------- BB	
Beluga, white whale -- *Delphinapterus leucas*	BB1A
Narwhal -- *Monodon monoceros*	BB2A
Family Ziphiidae ------------ BC	
Arnoux bottlenose whale -- *Berardius arnuxii*	BC1A
Baird's bottlenose -- *Berardius bairdii*	BC1B
Northern bottlenose -- *Hyperoodon ampullatus*	BC2A
Southern bottlenose -- *Hyperoodon planifrons*	BC2B
-- *Indopacetus pacificus*	BC3D
Sowerby's beaked whale -- *Mesoplodon bidens*	BC5A
-- *Mesoplodon bowdoini*	BC5B
Hubb's beaked whale -- *Mesoplodon carlhubbsi*	BC5C
Blainville's beaked -- *Mesoplodon densirostris*	BC5D
-- *Mesoplodon europaeus*	BC5E
-- *Mesoplodon ginkgodens*	BC5H
-- *Mesoplodon grayi*	BC5J
Hector's beaked whale -- *Mesoplodon hectori*	BC5K
-- *Mesoplodon layardii*	BC5L
True's beaked whale -- *Mesoplodon mirus*	BC5M
-- *Mesoplodon peruvianus*	BC5P
-- *Mesoplodon stejnegeri*	BC5S
Tasman beaked whale -- *Tasmacetus shepherdi*	BC7A
Cuvier's, goose-beaked -- *Ziphius cavirostris*	BC9A

COMMON NAMES FOR MARINE ANIMALS

Family Delphinidae (dolphins) ------------- BD

Commerson's dolphin Cephalorhynchus commersonii	BD1A
-- Cephalorhynchus eutropia	BD1B
-- Cephalorhynchus heavisidii	BD1C
Hector's dolphin -- Cephalorhynchus hectori	BD1D
Pacific common dolphin -- Delphinus bairdii	BD3A
Saddleback, common dolphin -- Delphinus delphis	BD3B
-- Delphinus tropicalis	BD3C
Grampus, Risseau's dolphin -- Grampus griseus	BD4A
Fraser's dolphin -- Lagenodelphis hosei	BD5A
White-sided dolphin -- Lagenorhynchus acutus	BD6A
White-beaked dolphin Lagenorhynchus albirostris	BD6B
Peale's dolphin -- Lagenorhynchus australis	BD6C
-- Lagenorhynchus cruciger	BD6E
Pacific white-sided Lagenorhynchus obliquidens	BD6G
Dusky dolphin -- Lagenorhynchus obscurus	BD6H
North. rt. whale dolphin - Lissodelphis borealis	BD8A
South. rt. whale dolphin -- Lissodelphis peronii	BD8B
-- Peponocephala electra	BD10A
Borneo dolphin -- Sotalia borneensis	BD12A
Amazon dolphin -- Sotalia brasiliensis	BD12C
-- Sotalia fluviatilis	BD12B
Guiana dolphin -- Sotalia guianensis	BD12D
Chinese river dolphin -- Sousa chinensis	BD13A
-- Sousa plumbea	BD13B
-- Sousa tëuszii	BD13C
Spotted dolphin -- Stenella attenuata	BD15A
-- Stenella clymene	BD15B
Striped dolphin -- Stenella coeruleoalba	BD15C
-- Stenella frontalis	BD15F
Spinner dolphin -- Stenella longirostris	BD15L
Rough-toothed dolphin -- Steno bredanensis	BD17A
Southern bottlenose dolphin -- Tursiops aduncus	BD19A
-- Tursiops catalania	BD19B
Pacific bottlenose dolphin -- Tursiops gillii	BD19C
(Atlantic) bottlenose -- Tursiops truncatus	BD19D

Family Globicephalidae (small toothed whales) -- BE

-- Feresa attenuata	BE1A
Southern pilot whale -- Globicephala edwardii	BE3A
Short-finned pilot -- Globicephala macrorhynchus	BE3B
Long-finned pilot -- Globicephala melaena	BE3C
Pacific pilot whale -- Globicephala scammoni	BE3D
Irrawaddy dolphin -- Orcaella brevirostris	BE5A
-- Orcaella fluminalis	BE5B
Killer whale -- Orcinus orca	BE7A
False killer -- Pseudorca crassidens	BE9A

Family Phocoenidae (porpoises) -- BF

-- Australophocaena dioptrica	BF1A
Harbor porpoise -- Phocoena phocoena	BF2A
Burmeister's porpoise -- Phocoena spinipinnis	BF2B
-- Phocoena sinus	BF2C
Dall's porpoise -- Phocoenoides dalli	BF4A
Finless porpoise -- Neophocaena phocaenoides	BF6A

Family Susuidae (dolphins -- River & South) -- BG	
Ganges river dolphin -- Susu gangetica	BG1A
Indus river dolphin -- Susu indii	BG1B
Boutu (Amazon river) -- Inia geoffrensis	BG2A
Yangtze river dolphin -- Lipotes vexillifer	BG3A
Franciscana -- Pontoporia blainvillei	BG4A
Seals, Sea otters, etc. (Carnivora) -- C	
Family Otariidae (fur seals, sea lions) ------ CA	
Southern fur seal -- Arctocephalus australis	CA1A
N. Z. fur seal -- Arctocephalus forsteri	CA1F
Galapagos fur -- Arctocephalus galapagoensis	CA1G
Kerguelen fur -- Arctocephalus gazella	CA1H
Juan Fernandez fur -- Arctocephalus philippii	CA1P
South African fur -- Arctocephalus pusillus	CA1R
-- Arctocephalus townsendi	CA1T
-- Arctocephalus tropicalis	CA1W
Northern fur seal -- Callorhinus ursinus	CA2A
Steller sea lion -- Eumetopias jubatus	CA3B
Australian sea lion -- Neophoca cinerea	CA4A
Southern sea lion -- Otaria flavescens	CA6A
N. Z. sea lion -- Phocarctos hookeri	CA8A
California sea lion -- Zalophus californianus	CA9A
Family Odobenidae (walrus) ------------------ CB	
Walrus -- Odobenus rosmarus	CB1A
Family Phocidae (seals) --------------------- CC	
Hood seal -- Cystophora cristata	CC1A
Bearded seal -- Erignathus barbatus	CC2A
Gray seal -- Halichoerus grypus	CC3A
Leopard seal -- Hydrurga leptonyx	CC4A
Weddell seal -- Leptonychotes weddellii	CC5A
Crabeater seal -- Lobodon carcinophagus	CC6A
Med. monk seal -- Monachus monachus	CC8A
Hawaiian monk seal -- Monachus schauinslandi	CC8B
Caribbean monk seal -- Monachus tropicalis	CC8C
North elephant seal -- Mirounga angustirostris	CC10A
South elephant seal -- Mirounga leonina	CC10B
Caspian seal -- Phoca caspica	CC12C
Ribbon seal -- Phoca fasciata	CC12F
Harp seal -- Phoca groenlandica	CC12G
Ringed seal -- Phoca hispida	CC12H
Spotted seal -- Phoca largha	CC12L
Baikal seal -- Phoca sibirica	CC12S
Harbor seal -- Phoca vitulina	CC12V
Ross seal -- Ommatophoca rossi	CC14A
Family Mustelidae (sea otter) --------------- CD	
Sea otter -- Enhydra lutris	CD1A
-- Lutra felina	CD2B

COMMON NAMES FOR MARINE ANIMALS

Family Ursidae (polar bear) ------------------ CE
Polar bear -- _Ursus maritimus_ CE1A

Manatees (Sirenia) ---------------- D

Family Dugongidae (dugong) ------------------- DA
Dugong -- _Dugong dugon_ DA1A
-- _Hydrodamalis gigas_ DA2B

Family Trichechidae (manatee) ---------------- DB
Amazonian manatee -- _Trichechus inunguis_ DB1A
West Indies manatee -- _Trichechus manatus_ DB1B
West African manatee -- _Trichechus senegalensis_ DB1C

OTHER MAMMALS
Primates --------------------- E
Chiroptera ------------------- F
Ungulates, _sensu lato_ -------- G
Other mammals ---------------- H

VERTEBRATES
Aves ------------------------- I
Reptilia --------------------- J
Amphibia --------------------- K
Fish, _sensu lato_ ------------- L
Other vertebrates ------------ S

INVERTEBRATES
Molluscs
Cephalopoda ------------- M
Other molluscs ---------- N
Arthropods
Crustacea --------------- O
Insecta ----------------- P
Other arthropods -------- Q
Other invertebrates ---------- R

GENERAL
Fossils ------------------------------- T
Uncertain (sea serpents and
other indeterminate animals) ----- U
General pinniped ---------------------- V
General cetacean ---------------------- W
Ambient noise (ship, geologic, ice) --- X
General mammal ------------------------ Y
Animals in general -------------------- Z

DOCUMENT LIBRARY

March 11, 1991

Distribution List for Technical Report Exchange

Attn: Stella Sanchez-Wade
Documents Section
Scripps Institution of Oceanography
Library, Mail Code C-075C
La Jolla, CA 92093

Hancock Library of Biology &
Oceanography
Alan Hancock Laboratory
University of Southern California
University Park
Los Angeles, CA 90089-0371

Gifts & Exchanges
Library
Bedford Institute of Oceanography
P.O. Box 1006
Dartmouth, NS, B2Y 4A2, CANADA

Office of the International
Ice Patrol
c/o Coast Guard R & D Center
Avery Point
Groton, CT 06340

NOAA/EDIS Miami Library Center
4301 Rickenbacker Causeway
Miami, FL 33149

Library
Skidaway Institute of Oceanography
P.O. Box 13687
Savannah, GA 31416

Institute of Geophysics
University of Hawaii
Library Room 252
2525 Correa Road
Honolulu, HI 96822

Marine Resources Information Center
Building E38-320
MIT
Cambridge, MA 02139

Library
Lamont-Doherty Geological
Observatory
Colombia University
Palisades, NY 10964

Library
Serials Department
Oregon State University
Corvallis, OR 97331-5503

Pell Marine Science Library
University of Rhode Island
Narragansett Bay Campus
Narragansett, RI 02882

Working Collection
Texas A&M University
Dept. of Oceanography
College Station, TX 77843

Library
Virginia Institute of Marine Science
Gloucester Point, VA 23062

Fisheries-Oceanography Library
151 Oceanography Teaching Bldg.
University of Washington
Seattle, WA 98195

Library
R.S.M.A.S.
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Maury Oceanographic Library
Naval Oceanographic Office
Stennis Space Center
NSTL, MS 39522-5001

Marine Sciences Collection
Mayaguez Campus Library
University of Puerto Rico
Mayagues, Puerto Rico 00708

Library
Institute of Oceanographic Sciences
Deacon Laboratory
Wormley, Godalming
Surrey GU8 5UB
UNITED KINGDOM

The Librarian
CSIRO Marine Laboratories
G.P.O. Box 1538
Hobart, Tasmania
AUSTRALIA 7001

Library
Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead
Merseyside L43 7 RA
UNITED KINGDOM

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO. WHOI-91-21	2.	3. Recipient's Accession No.
4. Title and Subtitle Marine Animal SOUND Database			5. Report Date August 1991
			6.
7. Author(s) William A. Watkins, Kurt Fristrup, and Mary Ann Daher			8. Performing Organization Rept. No. WHOI-91-21
9. Performing Organization Name and Address Woods Hole Oceanographic Institution Woods Hole, Massachusetts 02543			10. Project/Task/Work Unit No.
			11. Contract(C) or Grant(G) No. (C) N00014-88-K-0273 (G) N00014-91-J-1445
12. Sponsoring Organization Name and Address Office of Naval Research, NOARL			13. Type of Report & Period Covered Technical Report
			14.

15. Supplementary Notes

This report should be cited as: Woods Hole Oceanog. Inst. Tech. Rept., WHOI-91-21.

16. Abstract (Limit: 200 words)

The Marine Animal SOUND Database system encompasses (1) descriptive text databases cataloging the WHOI collection of underwater sound recordings from marine animals, (2) sets of files of digital sound sequences, (3) text databases organizing the digital sound sequences, and (4) software for analysis, display, playback, and export of selected sound files. The text databases index and sort the information on the sounds. The digital sound files are accessed directly from the text record, analyzed on screen, listened to, and compared or exported as desired. These databases provide comprehensive means for quantitative analyses and statistical comparisons of marine animal vocalizations.

The objective has been to develop basic tools for the study of marine animal sounds. The text database for cataloging the collection of recordings provides convenient sorting and selection of sounds of interest. Then, as specific sequences are digitized from these recordings, they become part of a second database system that manages these sound data. Once a digital sound is part of the database, several tools are available for interactive spectrogram display, sound playback, statistical feature extraction, and export to other application programs.

17. Document Analysis a. Descriptors

sound database
marine animals
underwater sounds
animal vocalizations

b. Identifiers/Open-Ended Terms

c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Approved for public release; distribution unlimited.		56
	20. Security Class (This Page)	22. Price

(See ANSI-Z39.18) *See Instructions on Reverse* OPTIONAL FORM 272 (4-77) (Formerly NTIS-35) Department of Commerce