Classification of dolphin echolocation clicks by energy and frequency distributions

D. S. Houser
Department of Biology, E&MS A316, University of California, Santa Cruz, California 95064

D. A. Helweg and P. W. Moore
SPAWAR Systems Center San Diego, Code 351, 49620 Beluga Road, San Diego, California 92152-6506

(Received 7 May 1998; accepted for publication 30 April 1999)

Dolphins demonstrate an adaptive control over echolocation click production, but little is known of the manner or degree with which control is exercised. Echolocation clicks (N=30,000) were collected from an Atlantic bottlenose dolphin (Tursiops truncatus) performing object discrimination tasks in order to investigate differential click production. Seven categories of clicks were identified using the spectral conformation and relative position of -3 and -10 dB peaks. A counterpropagation network utilizing 16 inputs, 50 hidden units, and 8 output units was trained to classify clicks using the same spectral variables. The network classified novel clicks with 92% success. Additional echolocation clicks (N=24,000) from two other dolphins were submitted to the network for classification. Classified echolocation clicks were analyzed for animal specific differences, changes in predominant click type within click trains, and task-related specificity. Differences in animal and task performance may influence click type and click train length. © 1999 Acoustical Society of America. [S0001-4966(99)05109-7]

PACS numbers: 43.80.Ka [FD]

INTRODUCTION

Dolphins can inspect objects by emitting trains or sequences of impulsive sounds, termed clicks, with the inter-click interval changing proportionally to target range (Penn, 1988). Bottlenose dolphins (Tursiops truncatus) display adaptive control over the emission of echolocation clicks both with respect to amplitude and frequency modulation, although the two are not fully independent of one another (Moore and Pawloski, 1990). Changes in the manner that click trains are utilized have been noted with respect to environmental noise, task specificity, and learning by both bottlenose dolphins and the false killer whale (Pseudorca crassiden) (Au et al., 1995b; Brill et al., 1992; Sigurdson, 1995). Furthermore, recent analysis of click train production through the use of chaos mathematics suggests underlying patterns may exist within click trains that are not detected by conventional mathematical techniques (Kremlivsky et al., 1998). Evidence regarding the dynamic sound production system of these small odontocetes and its voluntary control prompts further investigation into click train structure as well as the structure of individual clicks.

Au et al. (1995b) classified clicks of the false killer whale into four categories based upon the frequency spectrum of collected echolocation signals. The distribution of categories suggested that the false killer whale utilized relatively broadband signals with peak frequencies between 46 and 100 kHz and that an association between source levels and frequency modulation indicated a physiological constraint on the sound production mechanism. Our goal was to investigate adaptive control over click structure in bottlenose dolphins. We expanded upon this classification technique by establishing criteria for more strictly defined categories, and applied it to a much larger set of data. To fully assess the utility of such a classification scheme and its applicability to the study of dolphin echolocation, comparisons of click utilization were made between bottlenose dolphins performing object detection tasks and between the different intervals of a three-alternative match-to-sample task performed by a single dolphin. The general stability of the scheme was further evaluated by submitting the same data set to an artificial neural network and comparing its classification to the user-defined system.

I. MATERIALS AND METHODS

A. Subjects

Two Atlantic bottlenose dolphins (Tursiops truncatus: Tt751F, Tt018M) were trained to perform detection tasks in San Diego Bay, and another (Tt598M) performed a two-interval three-alternative match-to-sample (3A MTS) task in Kaneohe Bay, Hawaii. The object detection task was based upon a standard go/no-go paradigm in which the ‘‘go’’ response was emitted with the presence of the stimulus object. The two-interval paradigm required the animal to echo-inspect a sample target in the first interval, and echo-inspect a set of comparison targets in the second interval. Following the second interval the dolphin attempted to match the appropriate comparison target to the sample target (for details see Helweg et al., 1996).

B. Three alternative match-to-sample (3A MTS)

Data were collected during match-to-sample tasks. Sample targets were placed 4.65 m in front of the subject and comparison targets were placed 3.65 m in front of the subject and 1.6 m to the left and right of center. The subject was required to echo-inspect objects and signal
TABLE I. Categories of click types, click type description, and a representative spectrum for each. The horizontal dotted line signifies $-3$ dB region and the vertical dotted line signifies peak frequency.

<table>
<thead>
<tr>
<th>Click type</th>
<th>Description</th>
<th>Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>unimodal, low frequency ($&lt; 70$ kHz)</td>
<td>![Graph A]</td>
</tr>
<tr>
<td>B</td>
<td>unimodal, low frequency ($&lt; 70$ kHz); secondary high-frequency peak ($&gt; 70$ kHz) between $-3$ dB and $-10$ dB down</td>
<td>![Graph B]</td>
</tr>
<tr>
<td>C</td>
<td>bimodal</td>
<td>![Graph C]</td>
</tr>
<tr>
<td>D</td>
<td>unimodal, high frequency ($&gt; 70$ kHz); secondary low-frequency peak ($&lt; 70$ kHz) between the $-3$ dB and $-10$ dB down</td>
<td>![Graph D]</td>
</tr>
<tr>
<td>E</td>
<td>unimodal, high frequency ($&gt; 70$ kHz)</td>
<td>![Graph E]</td>
</tr>
<tr>
<td>W</td>
<td>wideband (single continuously bounded region within the $-3$ dB bandwidth with frequency bandwidth $&gt; 85$ kHz)</td>
<td>![Graph W]</td>
</tr>
<tr>
<td>M</td>
<td>3 or more distinctly bounded regions within the $-3$ dB</td>
<td>![Graph M]</td>
</tr>
</tbody>
</table>

“comparison” choices by responding to foam rubber balls attached to flexible PVC rods located above and to the sides of the station (Helweg et al., 1996).

A Bruel & Kjaer 8103 hydrophone (B&K) mounted 2 m from the subject and 1 m underwater was used to detect echolocation clicks. The B&K had a flat frequency response ($\pm 3$ dB) up to 150 kHz, with a sensitivity of $-211$ dB at 100 kHz. Detection of a click triggered the computer to store the click after an appropriate delay. The trigger threshold was set at 150 dB which resulted in capture of all clicks in each click
train and clicks were amplified 20 dB (Hewlett-Packard 465A). The clicks were digitized at 500 kHz with 12-bit resolution using an RC Electronics ICS-16 ComputerScope A/D board and 256 points per waveform were stored in a PC.

C. Object detection

During the object detection task, targets were placed 10 m in front of the subject. All targets were attached to monofilament line and lowered during target “present” trials. Targets were lowered to a depth of 1 m for presentation to the subject. A circular aperture was placed with the subject 1 m underwater. The subject was required to echo-inspect objects through this “window.” An aluminum sheet was used to block attempts to echolocate on targets prior to and between trials. It was placed between the subject and targets and suspended from pulleys for lowering and raising from the shelter. Responses by the subject were made to a foam rubber ball attached to a flexible PVC rod located above and to the side of the station.

A Bruel & Kjaer 8103 hydrophone (B&K) mounted 1 m from the subject and 1 m underwater was used to detect echolocation clicks. The B&K had a flat frequency response (±3 dB) up to 150 kHz, with a sensitivity of −211 dB at 100 kHz. Detection of a click triggered the computer to store the click after an appropriate delay. The trigger threshold was set at 150 dB which resulted in capture of all clicks in each click train and clicks were amplified 60 dB (Stanford Research Systems model SR560). The clicks were digitized at 500 kHz with 12-bit resolution using an RC Electronics ICS-16 ComputerScope A/D board and 256 points per waveform were stored in a PC.

D. Click classification

The frequency spectrum of >30,000 clicks collected from Tt598M were visually inspected and seven categories of click types were developed from these observations (Table I). Each category was based upon Boolean characters that described the form of the spectrum. Clicks were classified according to: (1) peak frequency; (2) the number of distinctly bounded regions existing within the 3-dB bandwidth; (3) the secondary peak frequency of a region, if one exists within the 3-dB bandwidth; (4) the frequency bandwidth of distinctly bounded regions existing in the 3-dB bandwidth; (5) the 10-dB bandwidth; (6) the number and peak frequency of modal regions existing within −3 dB and −10 dB bandwidths; (7) the drop in power of distinctly bounded regions existing between the −3 dB and −10 dB boundaries. A complete list of the rules utilized in the classification process is available upon request.

Type A clicks were defined by unimodal low-frequency (<70 kHz) spectral distributions (see Table I). Type B clicks were defined as unimodal low-frequency clicks with a secondary peak existing at a higher frequency between the −3 dB and −10 dB regions. Type E and type D clicks were, respectively, spectral mirror images of the previously described click types with the primary peak being high frequency (>70 kHz) and the secondary peak occurring at lower frequencies. Type C clicks contained a distinctly bounded bimodal distribution within the −3 dB bandwidth. Type W clicks were defined as wide-band clicks and contained a single bounded region of the spectrum within the −3 dB bandwidth with a frequency bandwidth of >85 kHz. All clicks that had three or more distinctly bounded regions within −3 dB of the peak frequency constituted type M, or multimodal, clicks.

The classification process was automated by creating a computer program that used the same Boolean decisions employed by a human expert using frequency spectrum. The implementation of the Boolean rules eliminated the potential for error by human classifiers and, because of the dichotomous nature of the scheme, necessarily classified clicks into one of the defined categories. A threshold peak SPL of 150 dB was established for inclusion of clicks through the Boolean classification program (BCP) and the analysis of the frequency spectrum was restricted from 27 to 150 kHz. A total of 54,283 clicks collected from the three dolphins were classified with the BCP (Table II). Subsets of clicks were randomly chosen and visually compared to the automated scheme to verify the automated process. In all cases clicks were correctly classified. For each trial clicks were categorically summed and each trial was considered an observation for statistical analysis. A Mann–Whitney U-test on rank sums was used to test for differences in click type usage between dolphins that performed object discrimination and between the sample and comparison tasks completed by Tt598M. All tests were performed with p < 0.05.

In order to test the intuitiveness of the classification scheme, i.e., without the implementation of Boolean rules, echolocation clicks were submitted to an artificial neural network (ANN) for classification. Because of its modular construction relevant to biological models and its potential and historical success at pattern recognition (Dayhoff, 1990; Roitblat et al., 1989; Rojas, 1996), a counterpropagation network was created that utilized 16 inputs, 50 hidden units, and 8 output units. The learning sequence consisted of submitting spectra of a given category, as defined by the Boolean scheme, as input. Training sets of 25 ideal (easily categorized) click spectra representing each click type were submitted to the network during the learning sequence. Additional sets of 125 novel click spectra, termed the “generalization sets,” were selected from the remaining and most ideal spectral forms and submitted to the neural network for the initial testing of its categorization ability. After initial testing, the entire click data set was submitted to the ANN and the output was compared to the automated classification program to determine the percentage of overall agreement between the two methods.

<table>
<thead>
<tr>
<th>Dolphin ID</th>
<th>Gender</th>
<th># of clicks</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tt51F</td>
<td>Female</td>
<td>13,679</td>
<td>Detection</td>
</tr>
<tr>
<td>Tt01B</td>
<td>Male</td>
<td>11,043</td>
<td>Detection</td>
</tr>
<tr>
<td>Tt598M</td>
<td>Male</td>
<td>29,561</td>
<td>3A MTS</td>
</tr>
</tbody>
</table>
II. RESULTS

A. Neural network

The counterpropagation network achieved a 92% success rate classifying the "generalization sets" of selected click types when compared to the automated Boolean classification scheme. The categorization of all click data by the ANN is presented in Fig. 1. Agreement between the network and the automated BCP for the entire click data set was variable; the percentage of agreement ranged from 45.5% to 82.0%. ANN classification of click types was most variable for Tt018M. In contrast, the ANN predominantly classified clicks produced by Tt751F as type E, or unimodal high-frequency clicks, while predominantly classifying clicks produced by Tt598M as type A clicks. The ANN classified 12,989 clicks as type D clicks, even though type D clicks, defined by Boolean category rules, were highly underrepresented within the data set (~30 of 54,000 clicks).

B. Click usage

1. Object detection task

Significant differences between Tt751F and Tt018M in click usage were observed for all categories (Table III). Except for unimodal high-frequency clicks (type E), Tt018M produced a greater number of clicks of each given category. Tt751F produced the lowest mean number of clicks per click train (mean751 = 12.1 ± 11.0). Most of her clicks were type E, with spectra that had only one peak existing above 70 kHz and within the −3-dB band (Fig. 2). Tt018M also produced relatively few clicks per train (mean018 = 15.8 ± 8.8) and displayed a broader use of click types (Fig. 2). Tt018M emitted type W clicks during early portions of click trains but did not persist as the click train progressed (Fig. 2). In both cases, the contribution of type B and type D clicks to click train composition was minimal.

Observations of position specific click type proportions indicated a change in click type production for Tt751F as the click train lengthened (Fig. 2). As the mean click train length was exceeded the production of type E clicks switched to that of type A, unimodal low-frequency clicks, and type M clicks which had multiple peak regions within −3 dB of peak amplitude and across the frequency range. In contrast, Tt018M demonstrated no changes in click type across click trains, but produced stable proportions of type A, type E, and

<table>
<thead>
<tr>
<th>Animal ID</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>M</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tt018M</td>
<td>1582</td>
<td>1582</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tt598M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1582</td>
</tr>
<tr>
<td>Tt751F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1582</td>
</tr>
</tbody>
</table>

FIG. 1. Distribution of clicks as categorized by the counterpropagation neural network. Clicks are distributed against the Boolean classification scheme and the overall agreement between the neural network and the classification program is given for each animal.

<table>
<thead>
<tr>
<th>Object detection task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tt751F vs Tt018M</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Matching to sample task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample vs comparison</td>
</tr>
</tbody>
</table>

TABLE III. Comparison of click types used by animals performing an object detection task (Tt751F and Tt018M) and a comparison of click types used by Tt598M in both sample and comparison segments of a two-interval match-to-sample task. Asterisks designate significant differences between animals or task interval.
type M when click train lengths of less than 60 clicks were considered (Fig. 2). In the few click trains exceeding 60 clicks in length, the clicks appeared erratic and without uniformity.

2. Matching-to-sample task

Statistical analysis of sample and comparison intervals performed by Tt598M indicated significant differences between all click categories except type D clicks (Table III). During the sample interval a greater number of clicks of all types except A and D were produced. In contrast, the production of type A clicks during the comparison interval was overwhelming (see Fig. 3). Proportional click usage throughout the click train for the other categories appeared to be stable but minimal during this interval. Unfortunately, a true mean click train length could not be determined for Tt598M because emitted click trains were often of a length that exceeded the capacity of the recording system (=99 clicks).

In both sample and comparison intervals, Tt598M produced type A clicks more often than any other click type. During sample intervals, type W and type M clicks were produced at a near-constant proportion across the duration of the train and, early in the click train, even comprised a greater portion of the clicks than type A clicks. In contrast, type A clicks always comprised ~70%–80% of the clicks for the comparison interval, regardless of position within the click train.

III. DISCUSSION

The click classification scheme described here demonstrated qualitative differences between click production by dolphins in similar echolocation tasks, as well as between the intervals of a task performed by the same dolphin (sample inspection versus comparison). Future comparisons should continue to focus on dolphins of different ages and sexes performing identical tasks with identical targets, as well as on the same dolphin performing multiple tasks. If the classification scheme continues to demonstrate qualitative differences in dolphin-dependent and task-dependent echolocation strategies, it should further our understanding of the adaptive control of echolocation as well as the ecological and physiological influences over it.
Consideration of the relevance of type B and type D
categories to the overall classification system may need to be
taken into consideration. Both click types contributed rela-
tively little to the overall click production of any animal and
type D clicks occurred only 30 times out of >54,000 clicks
produced. The correlation between frequency and amplitude
of emitted clicks observed in this species and other odonto-
cetes supports the existence of mechanistic constraints on the
sound production system (Au et al., 1995b; Moore and
Pawloski, 1990; Thomas and Turl, 1990; Thomas et al.,
1988). Lack of production of type D clicks, and possibly
Type B clicks, may relate to these constraints. Alternatively,
given that type B and D clicks are similar in spectrum shape
to type A and E clicks, respectively, they may represent tran-
sitional states away from or toward one of the latter click
types. In either case, the classification scheme appears to be
emphasizing characteristics that are of little use in determin-
ing differences in click production. Further comparisons be-
tween dolphins and tasks will determine whether type B
clicks should be merged with type A clicks and whether type
clicks should be merged with type E clicks.

Artificial neural networks have been utilized in dolphin
bioacoustics but have primarily focused upon aspects of tar-
get discrimination (Au et al., 1995a; Au, 1994; Helweg and
Moore, 1997; Moore et al., 1991; Roitblat et al., 1992,
1989). These studies utilized various ANN schemes to ad-
dress the importance of echo features to the discrimination
task as well as the biological relevance of the neural pro-
cesses involved. This study took a different approach by us-
ing a counterpropagation network to assess spectral prop-
erties of emitted clicks, specifically addressing characteristics
that may be used to distinguish between variable types of
click production. The ANN performed well when given ideal
frequency spectra for a given click type but performance declined when submitted with the entire data set. This sugg-
est that the ANN was capable of learning patterns that
were distinctive and of ideal spectral shape for a category,
but that performance deteriorated as spectral distributions
shifted from the ideal shape. The ANN classification of a
large number of clicks as type D is worth particular note. The
type D category appeared to act as an attractor in this system,
trapping energy states which did not readily settle into other
categorical states. The underlying cause of this was likely the
under-representation of type D categories within the data set
such that a more defined energy state for that category was
not learned. Still, the ANN provided general agreement
about overall click distributions as compared to the echol-
ocation click classification program; i.e., gross differences be-
tween dolphins and intervals of a task were still notable.

The three dolphins demonstrated different degrees of
production of specific click types with regard to a given task.
The variables which influenced these preferences are poten-
tially numerous and may have included such things as envi-
ronmental noise, the physiological condition of the animal
and the demands of the task. For example, Tt598M per-
formed his echolocation tasks in Kaneohe Bay, a noisy en-
vironment when compared to the relatively quiet waters of
San Diego Bay (Au et al., 1985). This may have impacted
decisions as to which frequencies and amplitudes the animal
utilized, presumably adapting sound production to optimize
useful echo returns (Au et al., 1985; Moore and Pawloski,
1990). Evidence presented here suggests Tt598M had a pref-
ference for low-frequency clicks, which contrasts reports
given for other animals performing similar tasks within Ka-
neohe Bay (Au et al., 1985).

Senescence of the auditory system may result in the al-
teration of click production in order to accommodate the loss
of sensitivity at certain frequencies. For example, Tt018M is
33 years of age and recent audiograms indicate a bilateral
decrease in sensitivity above 50 kHz, possibly as a result of
age-related retrograde neural loss (Ketten et al., 1997; Brill
et al., submitted). Alterations of click train structure in re-
sponse to the attenuation of returning echoes has been pre-
viously demonstrated by Brill and Harder (1991). Senes-
cence of the sound reception mechanism may create an
analogous scenario with regard to the sensitivity of both the
frequency and amplitude of returning echoes. This may ex-
plain the increased production of lower-frequency clicks by
Tt018M relative to Tt751F, since clicks of this type should
attenuate less rapidly in water than those of higher frequency
and would better match his hearing profile (Brill et al., sub-
mitted). The strategy used by Tt751F, a 14-year-old female,
contrasts that of the older male dolphin. Tt751F’s produc-
tion of high-frequency unimodal clicks may indicate a greater
sensitivity to higher-frequency echo returns and a sound re-
ception mechanisms as of yet unaffected by age- or sex-
related senescence.

Tt598M produced low-frequency clicks in a similar
manner to Tt018M. No audiograms currently exist for
Tt598M, but at 17 years of age, he was approaching the age
range for which decreases in high-frequency sensitivity have
been noted for males of this species (Ridgway and Carder,
1993, 1997). Unfortunately, it is impossible to differentiate
between the environmental influences, the demands of the
task, and possible physiological influences that may impact
click production without knowing the hearing sensitivity of
the animal or without comparing to other dolphins perform-
ing the same task.

Differences in click production by interval suggest that
variable strategies may be employed to optimize success at a
task. In the matching-to-sample task performed by Tt598M
there was a strong shift to type A click usage when perform-
ing the comparison interval. Although it is beyond the capac-
ity of this study to address the specifics underlying the
change in echolocation strategy, we can speculate that use of
low-frequency clicks may have related to the evocation of
salient echo features utilized by Tt598M in the decision mak-
ing process.

IV. CONCLUSIONS

The click classification scheme presented here demon-
strated differences between the types of clicks produced by
individual dolphins performing similar tasks and by a single
dolphin within a task. Further comparisons need to be made
between dolphins performing the same task and across a va-
riety of tasks in order to fully evaluate the utility of this
classification scheme. If it continues to prove useful, the
scheme may provide another tool with which to study dol-


