

# Acoustic identification of female Steller sea lions (*Eumetopias jubatus*)<sup>a)</sup>

Gregory S. Campbell<sup>b)</sup>

Cetacean Behavior Laboratory, College of Sciences, San Diego State University, 5500 Campanile Drive,  
San Diego, California 92182-4611

Robert C. Gisiner

Code 342, Office of Naval Research, 800 North Quincy Street, Arlington, Virginia 22217

David A. Helweg

Code D3501E, Space and Naval Warfare Systems Center San Diego, 53560 Hull Street, San Diego,  
California 92152-5000

Linda L. Milette

Marine Mammal Research Unit, Fisheries Center, University of British Columbia, 2204 Main Mall,  
Vancouver, British Columbia V6T 1Z4, Canada

(Received 30 May 2001; revised 28 January 2002; accepted 27 February 2002)

Steller sea lion (*Eumetopias jubatus*) mothers and pups establish and maintain contact with individually distinctive vocalizations. Our objective was to develop a robust neural network to classify females based on their mother-pup contact calls. We catalogued 573 contact calls from 25 females in 1998 and 1323 calls from 46 females in 1999. From this database, a subset of 26 females with sufficient samples of calls was selected for further study. Each female was identified visually by marking patterns, which provided the verification for acoustic identification. Average logarithmic spectra were extracted for each call, and standardized training and generalization datasets created for the neural network classifier. A family of backpropagation networks was generated to assess relative contribution of spectral input bandwidth, frequency resolution, and network architectural variables to classification accuracy. The network with best overall generalization accuracy (71%) used an input representation of 0–3 kHz of bandwidth at 10.77 Hz/bin frequency resolution, and a 2:1 hidden:output layer neural ratio. The network was analyzed to reveal which portions of the call spectra were most influential for identification of each female. Acoustical identification of distinctive female acoustic signatures has several potentially important conservation applications for this endangered species, such as rapid survey of females present on a rookery. © 2002 Acoustical Society of America. [DOI: 10.1121/1.1474443]

PACS numbers: 43.80.Ka, 43.80.Lb, 43.64.Tk [WA]

## I. INTRODUCTION

Steller sea lions were selected for this study because biological theory predicts that colonially breeding individuals will possess refined means of individual identification in order to invest finite resources in a manner most likely to enhance their reproductive success (Trivers, 1972). Individual identification is potentially possible for any individual of any species; however, evolutionary theory predicts that the need to reduce confusion over reproductive investment will result in more obvious and consistent communication of individual identity in colonially breeding species (Beecher, 1982). The task of developing a mathematical means of discriminating amongst individuals is thus more likely to succeed with a colonially breeding species than with a species

where opportunities for mistaking identity and misplacing competitive, breeding, or parental resources are less frequent and therefore less consequential.

Sea lions and fur seals (Otariids) typically demonstrate a polygynous breeding system, with females gathering in dense congregations to give birth and to copulate (Riedman, 1990). Mothers alternate between nursing their pups on land and feeding at sea, sometimes leaving the pup unattended for several days (Riedman, 1990). Acoustic mother-pup call recognition appears to be the primary means of initiating contact and orienting searching behavior after a period of separation (Insley, 1989, 1992; Gisiner and Schusterman, 1991; Schusterman *et al.*, 1992a; Phillips and Stirling 2000, 2001). Playback experiments have provided additional evidence to support the existence of mother-pup acoustic recognition systems. Studies on both the subantarctic fur seal (*Arctocephalus tropicalis*) and Galapagos fur seal (*A. galapagoensis*) have suggested that mothers and pups react positively to recordings of each others' calls but not to those of strangers (Trillmich, 1981; Roux and Jouventin, 1987). While individual call recognition among Otariids appears to be common, few studies have used acoustical and statistical

<sup>a)</sup>Portions of this work were presented in "Acoustic identification of female Steller sea lions," Proceedings of the 140th meeting of the ASA/Noise Congress, Newport Beach, CA, November 2000, and "Neural network classification of individual female Steller sea lions (*Eumetopias jubatus*)," Proceedings of the 14th Biennial Conference on the Biology of Marine Mammals, Vancouver, BC, Canada, November 2001.

<sup>b)</sup>Electronic mail: gcampbel@sonofsun.sdsu.edu



analysis techniques to evaluate call stereotypy. Only studies of the northern fur seal *Callorhinus ursinus* (Insley, 1989, 1992), South American sea lion (Fernández-Juricic *et al.*, 1999), and South American fur seal *Arctocephalus australis* (Phillips and Sterling, 2000) have classified known females based on the acoustic structure of their calls.

The Steller sea lion ranges along the North Pacific Rim from Japan to Central California, with centers of abundance and distribution in the Aleutian Islands and the Gulf of Alaska (Loughlin *et al.*, 1984; Riedman, 1990). In response to population declines in Alaska since the 1970s, the Steller sea lion was listed in 1990 as *threatened* under the U.S. Endangered Species Act (NMFS, 1992). Genetic and distribution information subsequently led to the classification of two separate breeding stocks of Steller sea lions: a “western” stock, which included sea lions in the Gulf of Alaska, Aleutian Islands, and Bering Sea, and an “eastern” stock which included sea lions from Southeast Alaska to California (NMFS, 1995; Bickham *et al.*, 1996; Loughlin, 1997). In 1997, the western stock of Steller sea lions was relisted as *endangered* under the U.S. Endangered Species Act. This reclassification was in response to a dramatic population decline during the past 30 years, estimated at over 80% (Braham *et al.*, 1980; Merrick *et al.*, 1987; Loughlin *et al.*, 1992; Trites and Larkin, 1996; Sease *et al.*, 2001). During this same time period, Steller sea lion numbers have remained stable or increased in Southeast Alaska and throughout most of the range of the eastern stock (Calkins *et al.*, 1999; Sease *et al.*, 2001).

The causes of the decline in the western stock of Steller sea lions are not known, though poor nutrition is currently a leading hypothesis. Undernutrition resulting from changes in the abundance or appropriateness of Steller sea lion prey items may lead to reduced birth rates and increased mortality through starvation, disease, or predation (Merrick and Loughlin, 1997; Calkins *et al.*, 1999). Considerable research on Steller sea lion population dynamics and energy requirements is currently being conducted in an effort to better understand the factors influencing the population decline in the western stock and the stability of the eastern stock.

Between May and July, Steller sea lions gather on offshore rookeries for reproductive activities. Females give birth to a single precocial pup, and for one week following parturition, mothers fast and suckle their pups on land. No allomaternal behavior has been reported. Following this perinatal period, mothers begin a cycle of feeding at sea for approximately 24 hours and then returning to the rookery for approximately 24 hours to suckle their young and to rest (Higgins, 1984; Riedman, 1990; Milette *et al.*, 1999). When Steller sea lion mothers return from foraging trips, they must repeatedly find their pups on the large crowded rookery. Acoustic mother-pup recognition appears to be the primary means by which female Steller sea lions achieve this essential task (Higgins, 1984; Milette *et al.*, 1999).

Monitoring the arrival and departure activities of individually identified mothers within and across breeding seasons is one method that can provide important information on maternal care, energy transfer, site fidelity, patterns of rookery usage, and philopatry or dispersion of matriline.

Currently, individual identification is done visually by comparison of photographs or drawings of distinctive markings (Gisiner, 1985; Milette *et al.*, 1999). However, the accuracy of this identification method depends on the observer’s experience and can lead to misidentification. Markings may be small and inconspicuous and only visible from one side, animals in the water usually cannot be identified visually, and crowded conditions make it difficult to single out and maintain visual contact with an identified individual. As an alternative, call identification has the potential for reliable discrimination of individuals based on acoustical parameters. Since females call whenever they move and whenever they depart to and return from the sea on feeding trips, acoustic identification may be preferred in conditions where direct visual observations may be difficult.

Campbell *et al.* (2000) demonstrated that female Steller sea lions can be discriminated by call characteristics. They evaluated a set of classifiers including linear discriminant analysis and three artificial neural network architectures (backpropagation, learning vector quantization, and adaptive resonance) using standardized training and generalization datasets of calls from 14 females. The backpropagation architecture clearly provided the highest classification accuracy on the novel generalization data.

Here, we advance development of a robust neural network to classify females based on their mother-pup contact calls. Average logarithmic spectra were extracted for calls from 26 females, and standardized training and generalization datasets created. A family of backpropagation networks (BPNs) was generated to assess relative contribution of spectral and architectural variables to classification accuracy. The BPN with best overall generalization accuracy was analyzed to reveal which portions of the call spectra were most influential for identification of each female.

## II. MATERIALS AND METHODS

### A. Study site and data collection

Our study site was Sugarloaf Island (58°53′N; 152°02′W), one of numerous offshore locations in the Gulf of Alaska that serve as rookery/haul out sites for the western population of Steller sea lions (Fig. 1). Prior to the current study, 70 females were visually identified and photographed at Sugarloaf during the breeding seasons of 1994–1997 (Milette *et al.*, 1999).

During the breeding seasons of 1998 and 1999, acoustic recordings were conducted with a two-channel Sony TCD-D8 DAT recorder connected to a Sennheiser M-60 Directional Microphone. One channel was devoted to sea lion calls and the second channel was used for observer narration. A typical recording contained a repetitive series of female calls with the name of the vocalizing individual stated after each call. We catalogued 573 contact calls from 25 females in 1998 and 1323 calls from 46 females in 1999. Nine known females were recorded in both years, yielding a total of 62 identified individuals.



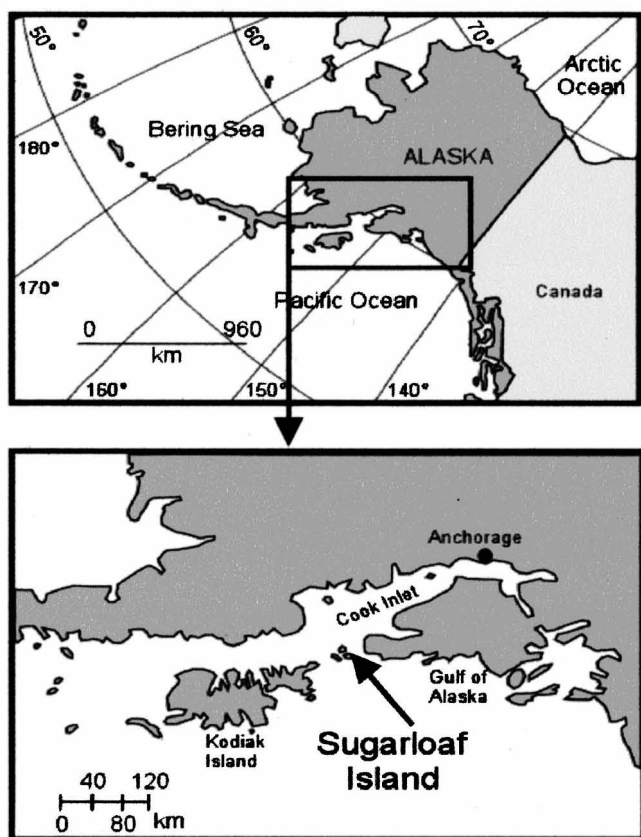


FIG. 1. The study site at Sugarloaf Island, Alaska. The top panel gives an overview of the western Gulf of Alaska region. The box and arrow indicate the location of study site, with a detail presented in the bottom panel.

## B. Acoustical analyses

Acoustical analyses were limited to nonagonistic calls directed by mothers toward their pups while on the rookery. The digital data were transferred from DAT to PC with a Zefiro Acoustics ZA-2 digital sound card over a Sony POC-DA12P fiber optic cable. Transfer was managed via SoundForge 4.5 multimedia editing software. Raw data preserved the 44.1-kHz sampling rate and stereo format in PCM WAV files. The stereo WAV files were scanned using the narration channel, and each call was segmented using a SoundForge "region." Each of the 1896 calls was labeled according to an individual female and one of four possible call quality ratings. Quality 1 (Q1) calls were vocalizations that had excellent signal to noise conditions with no little or no ambient noise. Quality 2 (Q2) calls were degraded by ambient noise (e.g., wind, rain or surf). Quality 3 (Q3) calls were degraded by other sea lion sounds and/or weather conditions, which made segmentation of the calls ambiguous. Quality 4 (Q4) calls were samples in which female identity was uncertain, and were not analyzed.

## C. Backpropagation architectures

After all recordings were segmented, a batch process low-pass filtered (4 kHz) and decimated each call to an 11.025-kHz sampling rate, and extracted the call to an individual WAV file. The calls were cross-tabulated by female and call quality. Q1 and Q2 calls were pooled, which yielded

a set of 26 females, each of which was represented by more than ten calls. Backpropagation is a supervised learning algorithm that requires a set of exemplars to be compiled into a learning set, and generalization ability can be tested using additional sets of novel exemplars. For this study, a standardized training set of calls was generated by drawing ten calls for each female at random ( $n=260$ ), using Q1 calls and supplementing with Q2 calls where necessary. All remaining Q1–Q3 calls ( $n=1281$ ) were placed into a generalization set. The data are summarized in Table I.

Each call was represented by an average logarithmic spectrum on the BPN input layer. We evaluated the effects of spectral and architectural variables on the generalization accuracy of the BPNs. Two spectral variables were manipulated. Frequency resolution was set at 5.38, 10.77, 21.53, or 43.07 Hz/bin, and bandwidth was set at 2 or 3 kHz. The BPN processing architecture was manipulated by modifying the ratio of hidden layer elements to output layer elements, using settings of 2:1, 4:1, and 6:1 hidden:output ratios. The output layer always had 26 elements, one for each female. Thus, for example, a BPN with a 3-kHz input bandwidth, a 10.77-Hz/bin frequency resolution, and a connectivity ratio of 4:1 would have 280 input elements (corresponding to the first 280 spectral bins, which span 0 to 3 kHz), 104 hidden layer elements, and 26 output layer elements.

The BPNs were fully interconnected and used standard cumulative-delta learning propagated across the bipolar tanh transfer function. Each BPN was trained to a rms error criterion of 0.05, which corresponds to approximately 95% correct across the training set. After training, the classification accuracy of each BPN for novel exemplars was tested using the generalization dataset. The overall generalization percent correct for each BPN was used to select the most accurate network. A detailed analysis of the relative importance of each frequency band in the input spectra was conducted. For every call in the generalization set that the BPN correctly classified, the value of each input bin was dithered by 20%, and the resulting percent change on the output neuron activation strength was measured. Dithering the values of input bins that had strong connection weights would result in large changes on the output neuron, thereby providing insight into the spectral characteristics that were most useful in classifying the females.

## III. RESULTS

### A. Female call characteristics

Female calls ranged in frequency from 30 to 3000 Hz with peak frequencies from 150 to 1000 Hz. Typical call duration was 1000 to 1500 ms. Call structure varied across females, illustrated in Fig. 2. Some individuals typically produced tonal calls, demonstrated by a call from Amatuli in the left column of Fig. 2, while others typically produced pulsed calls, demonstrated by a call from Ethel in the right column. Calls containing both mixed and pulsatile portions were also common in our data set. Note that the tonal call was characterized by clearly defined harmonic structure, which gives rise to the rippled spectra and harmonic lines in the spectrogram. In contrast, the pulsed call lacked strong harmonic structure and little spectral rippling.



TABLE I. The sample dataset. Data for each of the 26 females are broken out by quality factor within the training and generalization subsets. Each female was represented by ten calls in the training set, and by unequal  $n$ 's in the generalization set.

Female	Training		Generalization		
	Q1	Q2	Q1	Q2	Q3
763Girl	1	9	0	0	19
Amatuli	10	0	7	8	4
Bellow	6	4	0	9	6
Breeze	3	7	0	12	38
Bubbles	1	9	0	23	57
Chamille	10	0	7	45	82
Cher	10	0	1	82	96
Clover	1	9	0	1	3
Crunchie	2	8	0	25	40
Carcy	3	7	0	7	13
Demera	3	7	0	4	26
Dice	2	8	0	12	50
Ethel	4	6	0	9	53
Flo	9	1	0	12	28
Fungus Ear	0	10	0	0	65
GooGoo Doll	3	7	0	23	46
Jay	2	8	0	0	8
Jessica	7	3	0	9	5
Lucy	10	0	7	40	26
Moss	10	0	6	31	47
Pebbles	5	5	0	3	12
Pinky	4	6	0	1	5
Sasha	4	6	0	4	5
Scooby Doo	4	6	0	11	8
Snicker	9	1	0	16	25
Ziggy	0	10	0	33	66

### B. Neural network evaluation

All BPNs learned to classify the set of 26 females with 100% accuracy on the training dataset. Overall generalization accuracy ranged from 58.5% on the BPN with 2-kHz bandwidth, 43.07-Hz/bin frequency resolution, and 2:1 hidden:output layer connectivity to 70.7% on the BPN with 3-kHz bandwidth, 10.77-Hz/bin frequency resolution, and

TABLE II. Comparison of the generalization accuracy of the backpropagation networks. The first column is the frequency resolution of the spectral input representation in Hz. The second column is the connectivity ratio between the number of elements in the hidden layer to the output layer. The third column is the average percent correct (APC) (with standard deviation, s.d.) computed over all 26 females, using a 2-kHz bandwidth spectral representation. The fourth column is average percent correct using a 3-kHz bandwidth.

Frequency resolution (Hz/bin)	H:O Ratio	APC (s.d.) 2 kHz BW	APC (s.d.) 3 kHz BW
5.38	2:1	61.77 (22.26)	68.81(19.50)
10.77	2:1	64.69 (20.23)	70.69(17.15)
21.53	2:1	62.58 (18.54)	69.19(18.99)
43.06	2:1	58.50 (19.03)	69.19(20.39)
5.38	4:1	64.54 (21.20)	69.19(17.88)
10.77	4:1	63.38 (20.85)	69.81(17.72)
21.53	4:1	64.54 (17.70)	68.38(17.37)
43.06	4:1	60.00 (20.22)	58.96(20.61)
5.38	6:1	64.19 (20.51)	68.46(17.8)
10.77	6:1	66.31 (20.06)	70.62(17.36)
21.53	6:1	65.73 (18.43)	68.92(18.93)
43.06	6:1	60.65 (18.93)	59.73(19.45)

2:1 hidden:output layer connectivity. The results are summarized in Table II. The accuracy of BPNs with 2-kHz bandwidth was significantly lower than those with 3-kHz bandwidth [ $t(22)=2.74$ ,  $p=0.012$ ], with mean percent correct of 63.07 vs 66.99, respectively. Analysis of variance revealed that manipulating frequency resolution also resulted in significant differences in accuracy [ $F(3,8)=148.16$ ,  $p<0.01$ ]. Scheffe's *post hoc* comparison of means ( $\alpha=0.01$ ) demonstrated that accuracy of networks using a 43.05-Hz/bin resolution was significantly lower than that of the networks using 21.53-, 10.77-, and 5.38-Hz/bin. No significant effect of manipulating hidden:output layer connectivity ratios was observed.

On the basis of these evaluations, the BPN with 3 kHz of input bandwidth, 10.77-Hz/bin frequency resolution, and 2:1 hidden:output connectivity was selected for further study.

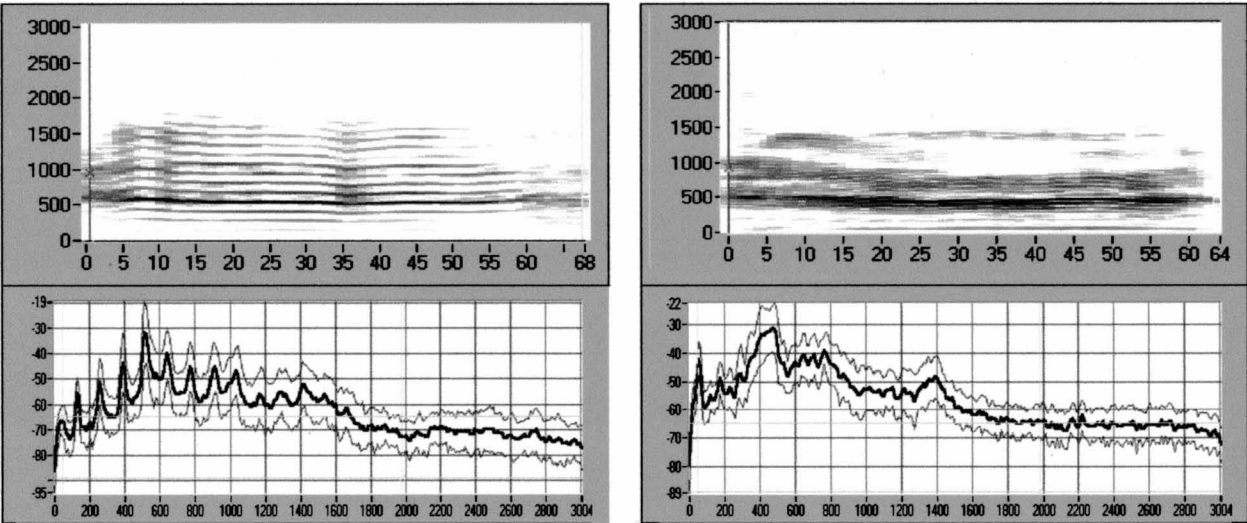


FIG. 2. Sample call types. A tonal call from Amatuli is presented on the left, and a pulsatile call from Ethel is presented on the right. The top panels are spectrograms and the bottom panels are average log power spectra. Duration of the calls were approximately X seconds. Frequency ranges from 0 to 3 kHz on both the spectrograms and power spectra; amplitude is in relative dB.



TABLE III. Confusion matrix for the optimal BPN on the generalization dataset. Actual female identity is in rows, and predicted identity is in columns. To facilitate comparison, values are percent classification, which normalizes across the unequal sampling. Percent correct is on the major diagonal.

	763GIRL	AMATULI	BELLOW	BREEZE	BUBBLES	CHAMILLE	CHER	CLOVER	CRUNCHIE	DARCY	DEMERA	DICE	ETHEL	FLO	FUNGUSEAR	GOOGOO	JAY	JESSICA	LUCY	MOSS	PEBBLES	PINKY	SASHA	SCOOPYDOO	SNICKER	ZIGGY
763GIRL	74	0	5	0	0	0	0	0	0	0	0	0	5	0	0	0	11	0	0	0	0	0	5	0	0	0
AMATULI	0	89	0	0	0	5	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
BELLOW	0	0	80	7	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BREEZE	0	0	2	74	4	0	0	0	6	0	2	6	0	0	0	0	0	0	2	0	0	0	0	0	4	0
BUBBLES	0	0	0	0	76	0	0	0	1	0	1	19	0	0	0	0	0	2	0	0	0	0	0	0	0	0
CHAMILLE	13	3	1	1	1	51	0	1	0	1	1	1	0	0	9	0	5	1	1	1	0	4	2	1	0	0
CHER	1	0	7	1	0	0	63	1	7	0	0	0	1	1	0	0	0	3	0	8	1	0	1	1	0	0
CLOVER	0	0	0	0	0	0	0	75	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0
CRUNCHIE	0	0	0	0	3	0	0	0	91	0	2	2	0	0	0	0	0	0	0	3	0	0	0	0	0	0
DARCY	0	0	0	0	5	0	0	5	0	65	0	5	5	0	0	0	0	5	0	0	5	0	0	0	0	0
DEMERA	0	0	0	0	0	0	0	0	0	0	90	7	0	0	0	0	0	3	0	0	0	0	0	0	0	0
DICE	2	0	0	3	16	2	0	0	0	0	3	68	2	0	0	0	0	3	0	0	0	0	2	0	0	0
ETHEL	0	0	0	2	6	0	6	0	0	0	0	2	53	0	0	0	0	0	0	16	0	8	2	2	0	0
FLO	0	0	0	2	0	2	5	2	8	0	2	5	40	0	0	0	0	0	0	0	0	0	0	0	30	2
FUNGUSEAR	2	0	0	0	0	2	0	0	2	0	0	0	0	0	51	2	2	3	0	18	0	17	0	2	2	0
GOOGOO	1	4	0	0	4	0	3	4	0	0	0	1	0	0	0	65	0	0	0	0	0	6	0	10	0	0
JAY	12	0	0	12	0	0	0	12	0	0	0	12	0	0	0	0	38	0	0	0	0	0	0	12	0	0
JESSICA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93	7	0	0	0	0	0	0	0
LUCY	0	1	0	1	8	7	0	0	1	0	0	8	0	0	0	0	3	5	56	0	0	3	3	3	0	0
MOSS	1	0	0	0	2	0	4	1	0	2	0	2	12	1	0	0	0	0	0	57	5	11	1	0	0	0
PEBBLES	0	0	0	0	0	0	0	13	0	13	0	0	0	0	0	0	0	0	0	7	60	0	0	7	0	0
PINKY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
SASHA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	78	0	0	22
SCOOPYDOO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
SNICKER	2	0	0	0	2	2	2	0	0	2	0	0	2	2	0	0	0	0	0	2	2	0	2	2	73	0
ZIGGY	0	0	0	0	0	0	10	2	0	1	0	2	0	2	0	0	0	1	0	0	0	2	0	2	0	78

No difference was observed between the 2:1 and 6:1 architectures, so the simpler 2:1 network was chosen. The confusion matrix for this network is presented in Table III. Calls were correctly classified 70.69% of the time ( $\pm 17.15$ ), with zero errors for some subjects. Interindividual variability in percent correct classification suggests that the BPN can distinguish some females more readily than others, though all individuals had scores that were well above what would be expected by chance (i.e., 26 females =  $1/26 = 3.8\%$ ). The most notable degradation in classification was observed for Flo (40%) and Jay (38%). The pattern of errors was not related to distribution of Q1 versus Q2 samples in the training set nor to the distribution of samples across quality levels in the generalization set. Also, the pattern of errors could not be attributed to the female call type (tonal versus pulsatile). Further analysis of errors would require a larger generalization set to justify more complex statistical treatment.

### C. Representation of call features within the BPN

A detailed analysis of the relative importance of each frequency band in the input spectra was conducted. For every call in the generalization set that the BPN correctly classified, the value of each input bin was dithered by 20%, and the resulting percent change on the output neuron activation strength was measured. Dithering the values of input bins that had strong connection weights would result in large changes on the output neuron, thereby providing insight into the spectral characteristics that were most useful in classifying the females.

The results of this analysis are summarized in Fig. 3, one female per panel. The data are percent change in activation of the output neuron assigned to that female, ranging between  $\pm 60\%$  change. The results can be interpreted as follows: The BPNs represented data on a  $-1$  to  $+1$  full-scale

range (bipolar). An output neuron with activation of  $+1$  indicates strong "confidence" of correct classification. An output neuron with activation of  $-1$  indicates no confidence of correct classification. Therefore, strong positive-going changes in the activation levels indicate that those frequency bands made important contributions to confidence in correctly classifying the female. Strong negative-going changes in the activation levels indicate that those frequency bands made important contributions to confidence in classifying the call as *not* coming from that female. Small changes indicate weak areas of frequency differentiation.

The overall picture suggests that most of the highest frequency weightings lie below about 2 kHz, which is supported by an approximate 5–15-dB rolloff in sound energy above those frequencies. However, recall that the generalization accuracy of BPNs with 2-kHz bandwidth was significantly poorer than those with 3-kHz bandwidth. Thus, although the frequency weightings were attenuated in the 2–3-kHz range, they added about 10% improvement in classification accuracy. Visual comparison of the weightings indicates a diverse set of positive- and negative-going patterns that defy straightforward summary. However, note that the percent changes for Flo and Jay are attenuated (especially those for Jay)—these are the two females for which generalization accuracy was poorest. Dithering the input representation resulted in very little percent change in the output neuron, suggesting that the network could neither identify features for positive attribution nor features for positive rejection of these females.

## IV. DISCUSSION

Acoustic signature recognition requires low intraindividual call variability and high interindividual call variability. Our analyses suggest that female Steller sea lions have



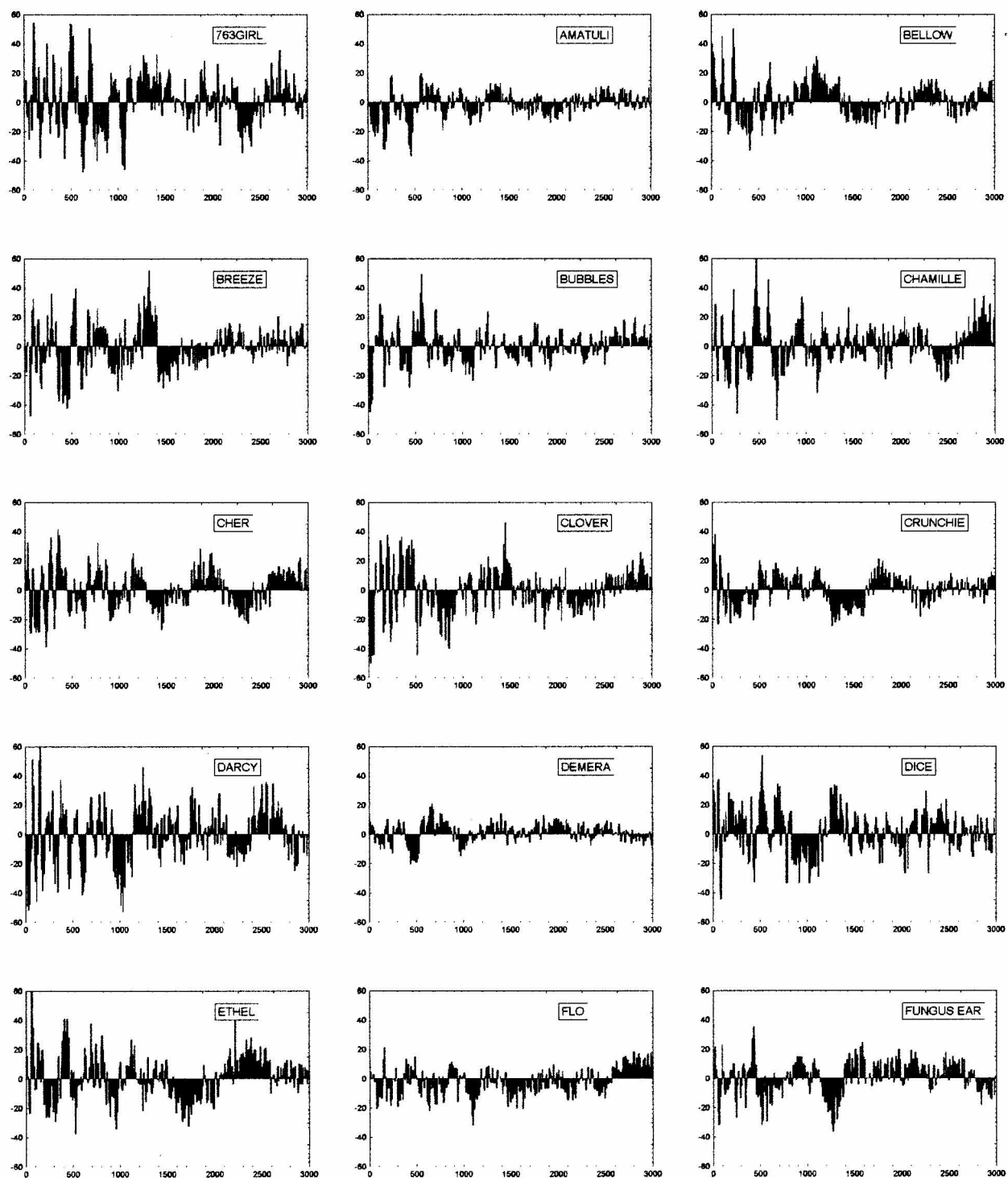


FIG. 3. Relative importance of spectral regions for neural network classification accuracy. Spectral frequency lies on the x-axis from 0 to 3000 Hz in 10.77 Hz bins. Percent change in output neuron activation lies on the y-axis from  $-60\%$  to  $+60\%$ . Large positive-going changes in the activation level indicate that those frequency bands made important contributions to confidence in classifying the calls as coming from that female. Large negative-going changes in the activation level indicate that those frequency bands made important contributions to confidence in rejecting the call as coming from that female. Small changes suggest that the network could identify features neither for confident attribution nor rejection of that female.

evolved sufficient vocal stereotypy to differentiate between individuals. The BPN was relatively accurate, and clearly able to consistently differentiate females based on their call parameters. The results of this research are consistent with other studies on colonially breeding pinniped species, which have suggested the existence of acoustic mother-pup recognition systems facilitated by individually distinctive vocal signatures (Bartolomew, 1959; Peterson and Bartholomew,

1969; Trillmich, 1981; Roux and Jouventin, 1987; Insley, 1989, 1992; Gisiner and Schusterman, 1991; Hanggi, 1992; Schusterman *et al.*, 1992b; Fernández-Jurísic *et al.*, 1999; Phillips and Stirling, 2000, 2001).

In this study, the best BPN yielded an overall percent correct classification of 71% on novel data (100% correct on the training data). In contrast, Phillips and Stirling (2000) reported that linear discriminant analysis resulted in 55%



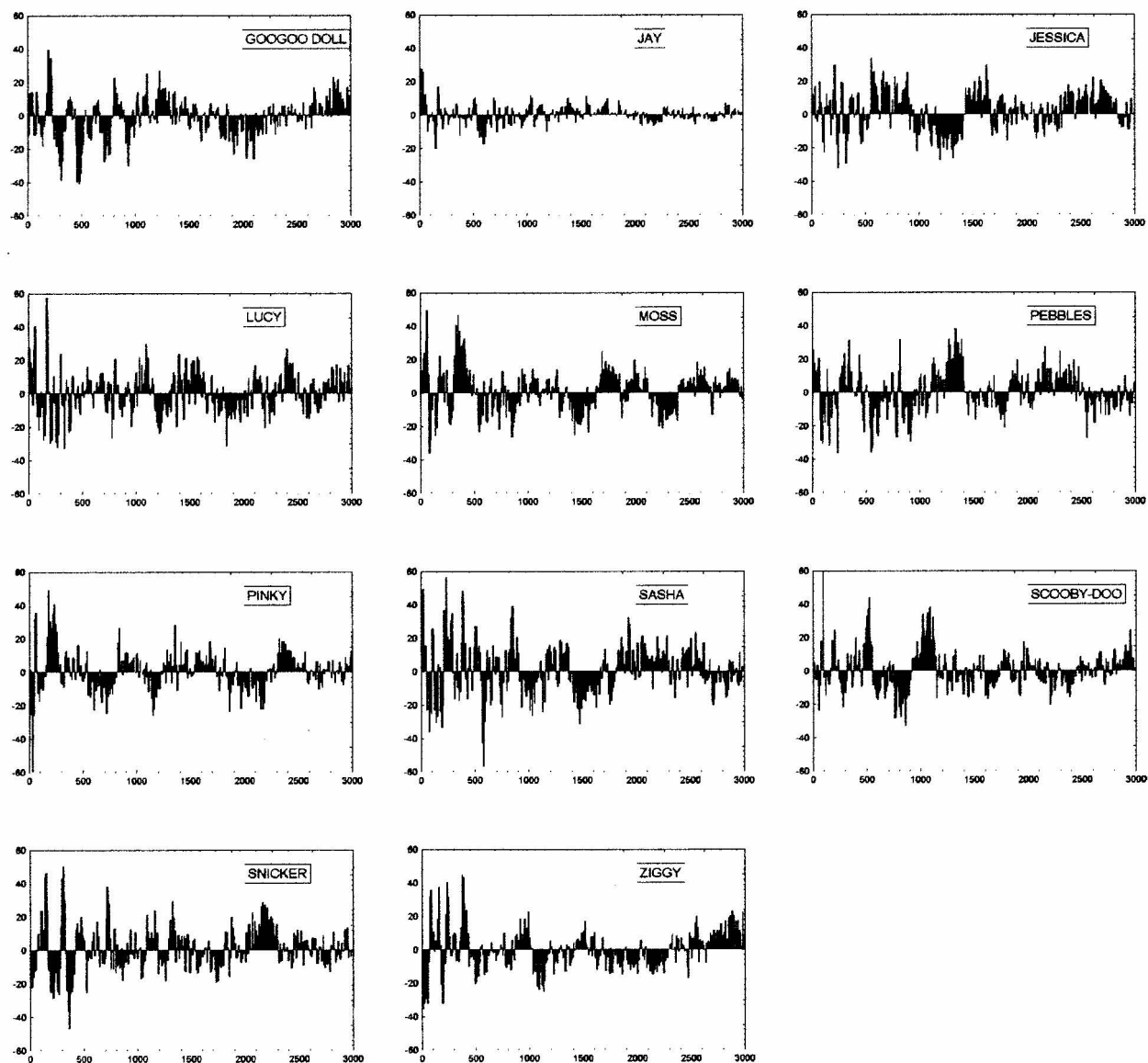


FIG. 3. (Continued.)

correct on novel data (70% on the training data) for a set of 15 female South American fur seals. Neural networks clearly outcompeted LDA on our Steller sea lion data as well (Campbell *et al.*, 2000). Though our BPN analyses of female Steller sea lion calls demonstrated the existence of individual acoustic signatures, the system was still prone to decision errors. The fact that classification rates were not higher overall may be explained by a number of factors. First, the specific acoustical parameters that pups use to identify mothers are not known and our inability to discern what variables are most important for identification may have limited the system's discriminatory power. In addition, recognition may occur in two ways, either by discrimination of only one individual (i.e., mother) but not of others, or discrimination among a number of related individuals (Beecher, 1982). Moreover, we do not understand the degree to which other factors may have contributed to variability in call structure. These factors include changes in motivational state within and between call bouts, effects of physiognomy (e.g., female size and condition), stability of female vocal tract character-

istics across years, and varying signal-to-noise conditions in the recordings.

Implementing an autonomous acoustic recognition system in the real world faces several challenges. For example, the deployment of the acoustic sensor(s) needs to optimize the desire for maximal signal-to-noise ratio for female calls and the desire to survey as many females as possible. In the current study, calls were recorded by a field technician using a directional microphone, which provides both for intelligent survey of females, and best possible recording quality. An affordable autonomous system is not likely to be mobile (i.e., robotic), but an array of fixed microphones could be deployed which enclose the female haul-out site within their aperture. Beamforming would then substitute for the ability of a field technician to isolate a specific female and maximize signal strength by manually orienting the handheld directional microphone.

The recognition system must be robust with respect to signal quality and the presence of distractors such as male calls, pup calls, other types of female calls, and other discrete



ambient biotic sounds (e.g., songbirds). Moreover, the signal processing system must be able to extract female calls from the real-time acoustic data stream. Finally, the recognition system must be able to form new output categories for new females, ideally without supervision. We speculate that the signal processing issues might be solved as follows: The system would consist of several cascaded stages. In the first stage, the real-time data would be parsed into successive time slices. Each time slice would be presented to a filter in which mother-pup contact calls would be discriminated from other coarse classes of calls (e.g., male calls, pup calls, other female calls, and other biotic sounds). Once categorized as a female pup-contact call, the data segment would be passed to the second-stage classifier in which individual female recognition would occur. The second-stage classifier could utilize unsupervised learning algorithms, such as adaptive resonance. Theoretically, this type of architecture would permit formation of new categories for new females by detection of "sufficiently novel" calls. Testing the utility of this concept will require several years of field data from individually identified females.

These points illustrate the fact that development of an autonomous, real-time, adaptive recognition system is not trivial. Tradeoffs will exist inherently between the desire for real-time processing speed and the desire for accuracy. If the system is to be remotely deployed and the data are not required in real time, we suggest that real-time recognition be abandoned and that slower, more accurate signal processing methods be used. A haul-out site could be sampled for a segment of time, and the archived data processed between sampling epochs. Post-processing introduces the ability to apply tools such as adaptive "whitening" filters that increase signal-to-noise ratio by reduction of local ambient noise. In addition, post-processing could be used to scan multiple beams (from an array) to collect calls from multiple females, and to potentially allow the recognition system to reorganize as a result of novel incoming calls. Although an autonomous adaptive recognition system is technically exciting, it is important not to lose sight of the importance of verifying acoustic classification with visual recognition by an expert, such as reported in the current study.

Individually distinctive acoustic signatures in female Steller sea lions have several potentially important scientific applications. Statistically based voice identification may be a particularly important tool for studying the decline of this endangered species. First, because females usually call when they enter and leave the rookery but are not always visible, the classification system described here could be used in the field to document the precise female arrival and departure times at the rookery. This information could be used to document patterns of maternal care/energy transfer during the critical development stage of pups. Time devoted to foraging and feeding trips by different individuals in different age classes could be examined. Comparisons of these parameters between the western (endangered) and eastern (threatened) populations of Steller sea lions may be of particular interest due to the different conservation status of the two populations. In addition, data collected over multiple years could

allow us to determine site fidelity of females, interannual patterns of rookery usage, and philopatry or dispersion of matriline.

## ACKNOWLEDGMENTS

We are grateful to R. H. Defran for providing advice and assistance with many aspects of this research. The Alaska Sea Life Center deserves recognition for its support of the field effort on Sugarloaf Island from 1994 to 1999. The SDSU Foundation provided essential administrative assistance. We thank two anonymous reviewers for their helpful comments. Funding was provided by an ONR Research Opportunities for Program Officers (ROPO) award to Dr. R. H. Defran at San Diego State University (ONR N00014-00-1-0114). This paper is dedicated to the loving memory of Olga and Woody Moore.

- Bartholomew, G. A. (1959). "Mother-young relations and the maturation of pup behavior in the Alaska fur seal," *Behaviour* 7, 187–208.
- Beecher, M. D. (1982). "Signature systems and kin recognition," *Am. Zool.* 22, 477–490.
- Bickham, J. W., Patton, J. C., and Loughlin, T. R. (1996). "High variability for control-region sequences in a marine mammal: Implications for conservation and biogeography of Steller sea lions (*Eumetopias jubatus*)," *J. Mammal.* 77, 95–108.
- Braham, H. W., Everitt, R. D., and Rugh, D. J. (1980). "Northern sea lion decline in the eastern Aleutian Islands," *J. Wildl. Manage.* 44, 25–33.
- Calkins, D. G., McAllister, D. C., Pitcher, K. W., and Pendleton, G. W. (1999). "Steller sea lions status and trend in Southeast Alaska: 1979–1997," *Marine Mammal Sci.* 15, 462–477.
- Campbell, G. S., Gisiner, R., and Helweg, D. A. (2000). "Acoustic identification of female Steller sea lions," Abstracts of the joint 140th meeting of the ASA/Noise Congress, Newport Beach, CA, November, p. 2541.
- Fernández-Juricic, E., Campagna, C., Enriquez, V., and Ortiz, C. L. (1999). "Vocal communication and individual variation in breeding South American sea lions," *Behaviour* 136, 495–517.
- Gisiner, R. (1985). "Male territoriality and reproductive behavior in the Steller sea lion, *Eumetopias jubatus*," Ph.D. dissertation, Univ. of Calif. at Santa Cruz.
- Gisiner, R., and Schusterman, R. J. (1991). "California sea lion pups play an active role in reunions with their mothers," *Anim. Behav.* 41, 364–366.
- Hanggi, E. B. (1992). "The importance of vocal cues in mother-pup recognition in a California sea lion," *Marine Mammal Sci.* 8, 430–432.
- Higgins, L. V. (1984). "Maternal behavior and attendance patterns of the Steller sea lion in California," M.S. thesis, Univ. of Calif. at Santa Cruz.
- Insley, S. J. (1989). "Vocal recognition between mothers and pups in the northern elephant seal (*Mirounga angustirostris*) and northern fur seal *Callorhinus ursinus*," M.S. thesis, Univ. of Victoria.
- Insley, S. J. (1992). "Mother offspring separation and acoustic stereotypy: a comparison of call morphology in two species of pinnipeds," *Behaviour* 120, 103–121.
- Loughlin, T. R. (1997). "Using the phylogeographic method to identify Steller sea lion stocks," in *Molecular Genetics of Marine Mammals*, edited by A. E. Dizon, S. J. Chivers, and W. F. Perrin (Soc. Marr. Mammal.), Spec. Publ. No. 3, pp. 329–341.
- Loughlin, T. R., Perlov, A. S., and Vladimirov, V. A. (1992). "Range-wide survey and estimation of total abundance of Steller sea lions in 1989," *Marine Mammal Sci.* 8, 220–239.
- Loughlin, T. R., Rugh, D. J., and Fiscus, C. H. (1984). "Northern sea lion distribution and abundance: 1956–80," *J. Wildl. Manage.* 48, 729–740.
- Merrick, R. L., and Loughlin, T. R. (1997). "Foraging behavior of adult female and young-of-the-year Steller sea lions in Alaskan waters," *Can. J. Zool.* 75, 776–786.
- Merrick, R. L., Loughlin, T. R., and Calkins, D. G. (1987). "Decline in abundance of the northern sea lion, *Eumetopias jubatus*, in Alaska, 1956–1986," *Fish. Bull.* 85, 351–365.
- Milette, L. L., Trites, A. W., and Calkins, D. (1999). "Behavior of lactating Steller sea lions: are they nutritionally stressed?," Abstracts of the 13th Biennial Conference on the Biology of Marine Mammals, Maui, HI, December, p. 125.



- NMFS (National Marine Fisheries Service) (1992). "Recovery plan for the Steller sea lion (*Eumetopias jubatus*)" (NMFS, Silver Spring, MD).
- NMFS (National Marine Fisheries Service) (1995). "Status Review of the United States Steller sea lion (*Eumetopias jubatus*) population" (NMFS, National Marine Mammal Laboratory, Seattle, WA).
- Peterson, R. S., and Bartholomew, G. A. (1969). "Airborne vocal communication in the California sea lion, *Zalophus californianus*," *Anim. Behav.* **17**, 17–24.
- Phillips, A. V., and Stirling, I. (2000). "Vocal individuality in mother and pup South American fur seals, *Arctocephalus australis*," *Marine Mammal Sci.* **16**, 592–616.
- Phillips, A. V., and Stirling, I. (2001). "Vocal repertoire of South American fur seals, *Arctocephalus australis*: structure, function, and context," *Can. J. Zool.* **79**, 420–437.
- Riedman, M. (1990). *The Pinnipeds: Seals and Sea Lions* (Univ. of California, Berkeley).
- Roux, J. P., and Jouventin, P. (1987). "Behavioral cues to individual recognition in the Subantarctic fur seal, *Arctocephalus tropicalis*," NOAA Tech. Rep. NMFS No. 51.
- Schusterman, R. J., Gisner, R., and Hanggi, E. B. (1992a). "Imprinting and other aspects of pinniped/human interactions," in *The Inevitable Bond*, edited by H. Davis and D. Balfour (Cambridge U. P., New York), pp. 334–356.
- Schusterman, R. J., Hanggi, E. B., and Gisner, R. (1992b). "Acoustic signalling in mother-pup reunions, interspecies bonding, and affiliation by kinship in California sea lions *Zalophus californianus*," in *Marine Mammal Sensory Systems*, edited by J. A. Thomas, R. A. Kastelein, and A. E. Supin (Plenum, New York), pp. 533–551.
- Sease, J. L., Taylor, W. P., Loughlin, T. R., and Pitcher, K. W. (2001). "Aerial and land-based surveys of Steller sea lions (*Eumetopias jubatus*) in Alaska, June and July 1999 and 2000," U.S. Dept. Commer., NOAA Tech. Memo. NMFS-AFSC-122.
- Trillmich, F. (1981). "Mutual mother-pup recognition in Galapagos fur seals and sea lions: Cues used and functional significance," *Behaviour* **78**, 21–42.
- Trites, A. W., and Larkin, P. A. (1996). "Changes in the abundance of Steller sea lions (*Eumetopias jubatus*) in Alaska from 1956 to 1992: how many were there?" *Aquatic Mamm.* **22**, 153–166.
- Trivers, R. L. (1972). "Parental investment and sexual selection," in *Sexual Selection and the Descent of Man*, edited by B. Campbell (Aldine, Chicago), pp. 136–179.