Chapter 34
Artificial Neural Networks: A New Tool for Studying Lemur Vocal Communication

Luca Pozzi, Marco Gamba, and Cristina Giacoma

Abstract Previous studies have applied Artificial Neural Networks (ANNs) successfully to bioacoustic problems at different levels of analysis (individual and species identification, vocal repertoire categorization, and analysis of sound structure) but not to nonhuman primates. Here, we report the results of applying this tool to two important problems in primate vocal communication. First, we apply a supervised ANN to classify 222 long grunt vocalizations emitted by five species of the genus *Eulemur.* Second, we use an unsupervised self-organizing network to identify discrete categories within the vocal repertoire of black lemurs (*Eulemur macaco*). Calls were characterized by both spectral (fundamental frequency and formants) and temporal features. The result show not only that ANNs are effective for studying primate vocalizations but also that this tool can increase the efficiency, objectivity, and biological significance of vocal classification greatly. The advantages of ANNs over more commonly used statistical techniques and different applications for supervised and unsupervised ANNs are discussed.

Resume Des études antérieures ont appliqué avec succès les Réseaux Neuronaux Artificiels (RNA) aux questions bioacoustiques (reconnaissance individuelle et inter-spécifique, catégorisation des répertoires vocaux, et analyse des structures sonores), mais pas sur les primates non-humains. Ici nous appliquons cet outil à

L. Pozzi (✉)
Department of Anthropology, Center for the Study of Human Origins, New York University, 25 Waverly Place, New York 10003, NY, USA
e-mail: luca.pozzi@nyu.edu

M. Gamba • C. Giacoma
Dipartimento di Biologia Animale e dell'Uomo, Università degli Studi di Torino, Via Accademia Albertina, 13, 10124 Torino, Italy
e-mail: marco.gamba@unito.it; cristina.giacoma@unito.it

J. Masters et al. (eds.), *Leaping Ahead: Advances in Prosimian Biology*, Developments in Primatology: Progress and Prospects, DOI 10.1007/978-1-4614-4511-1_34,

deux problèmes concernant la communication vocale des Primates. Premièrement, nous utilisons un modèle de RNA « supervisé » à la classification de 222 « grognements longs » émis par 5 espèces du genre *Eulemur*. Deuxièmement, nous utilisons un modèle auto-organisé « non supervisé » de RNA pour identifier des catégories discrètes dans le répertoire vocal du lémur Macaco (*Eulemur macaco*). Les vocalisations sont caractérisées par leurs propriétés spectrales (fréquence fondamentale et formant) et temporelles. Les résultats montrent que les RNA sont des outils efficaces pour l'étude des vocalisations des primates, mais aussi que cette méthode qui accroit l'efficacité, l'objectivité, et la signification biologique des classifications vocales. Les avantages des RNA sur d'autres méthodes communément utilisées, ainsi que différentes applications des RNA supervisées et non supervisées sont discutés.

Introduction

Identifying discrete categories is one of the most challenging problems in bioacoustics. In studies of animal communication, the classification of vocal signals into discrete categories is a preliminary step to comparing acoustic variability at different levels of analysis (definition of vocal repertoires, identification of species or individuals, analysis of call structure, etc.). Unfortunately, current statistical methods do not always give satisfactory results, particularly when the data are nonlinearly distributed (Demuth and Beale 1993). Moreover, most statistical techniques involve a number of a priori assumptions that may affect the objectivity of the classification and, therefore, the comparison of results between studies.

In recent years, many attempts have been made to develop mathematical tools for automatic and "objective" classification. Artificial Neural Networks (ANNs) have been widely used in the field of pattern recognition and have been applied successfully within the biological sciences to model complex functions and solve problems of classification, regression, and prevision (Ghirlanda and Enquist 1998).

The Theory of Artificial Neural Networks: Structure and Function

ANNs are computer simulations of biological nervous systems and mimic their fault-tolerance and learning capacity by modeling the low-level structure of the brain. Although this technique cannot reach the complexity of most biological nervous systems, it provides a powerful classificatory tool on account of its ability to learn a specific classification scheme and deal with incomplete or noisy data.

ANNs can be classified into two main groups: s*upervised* and *unsupervised* neural networks. The operation of a supervised neural network can be divided into two phases (1) *learning*, in which the network is trained to recognize different output categories (*targets*) and (2) *generalization*, where the network autonomously classifies a new set of previously unseen data according to the classification scheme acquired during the previous phase. Unlike supervised ANNs, unsupervised networks do not require a priori definition of output categories, and the process of classification is based on local information only. The network is thus able to self-organize data, to detect patterns in inputs autonomously, and to classify them into categories without requiring the number or types of categories to be predefined (Kohonen 1988).

Application of ANNs to Bioacoustic Problems

Because of their ability to model complex functions, ANNs have been used to address several issues in animal bioacoustics. The main applications of ANNs in this field have been to human speech, but they have also been applied to a variety of animal vocalizations. ANNs have been successfully applied to categorization of species' vocal repertoires (whales: Mercado and Kuh 1998; Murray et al. 1998), individual (sea lions: Campbell et al. 2002; deer: Reby et al. 1997; birds: Adi et al. 2010) and species identification (insects: Chesmore 2004; Chesmore and Ohya 2004; birds: Derégnaucourt et al. 2001; Lopes et al. 2011; bats: Parsons 2001; Jennings et al. 2008), and studies of the acoustic structure of vocalizations (birds: Dawson et al. 2006; Nickerson et al. 2006; whales: Green et al. 2011).

To date, the application of this computational technique in the study of nonhuman primate utterances has been extremely limited. Zimmermann (1995) suggested neural network modeling as a tool for the analysis and interpretation of primate sounds but did not perform any experimental studies using the technique. Recently, we applied ANNs to the study of the vocal repertoire of black lemurs, *Eulemur macaco* (Pozzi et al. 2010), and demonstrated the ability of supervised ANNs to classify lemur vocalizations into discrete categories with 94% correct prediction overall. In this chapter, we explore this approach further for classifying primate vocalizations, specifically (1) to differentiate the calls of different *Eulemur* species and (2) to categorize the vocal repertoire of black lemurs using unsupervised ANNS.

Materials and Methods

Application of Supervised ANNs to Differentiate Eulemur Species Vocalizations

A sample of 222 long grunts emitted by *Eulemur macaco* (115), *E. mongoz* (33), *E. coronatus* (23), *E. rubriventer* (16), and *E. fulvus* (35) was randomly selected from a larger data set recorded between 1995 and 2005 [see Gamba and Giacoma (2005) for recording techniques]. This vocal type was chosen because (1) it is present in the vocal repertoire of all lemurid species (Macedonia and Stanger 1994; Gamba and Giacoma 2005); (2) it is emitted by all species at a high rate, providing a good sample for the analyses; and (3) previous studies have demonstrated that long grunts are species specific (Gamba and Giacoma 2005). The different rates of long grunt emission across the five species and the need for good sound quality biased the number of recordings included in the sample. However, ANNs are not influenced by unbalanced numbers in different groups, only by the number of inputs during the training phase (Pozzi et al. 2010).

Each vocalization was split into a series of very short duration (0.01 s) nonoverlapping time windows, and 30 windows were sampled at regular intervals. The average duration of each vocalization was 0.521 s (SD 0.145 s), meaning that most vocalizations had approximately 50 windows. Vocalizations shorter than 0.30 s were excluded from the analysis. The inputs for the ANNs were two-dimensional characterizations of long grunts, in which each signal was described by a combination of pitch (F0) and the first four formants (i.e., spectral peaks), F1–F4. Each vector was constructed by concatenating the five 30-element vectors into a single 150-element vector and served as inputs into a supervised artificial neural network using a back-propagation algorithm during the training phase. Neural network analyses were performed using Statistica Neural Networks 7.1 (StatSoft, Trajan Software Ltd 1996–2000).

Identification of Distinct Calls Within the Eulemur macaco Vocal Repertoire Using an Unsupervised ANN

A total of 311 vocalizations including seven different vocal types (alarm calls, grunted hoots, hoots, grunts, long grunts, long grunt clear calls, and tonal calls) described in previous studies (Gosset et al. 2002; Macedonia and Stanger 1994) was used in the study. Each vocalization was characterized by a 10-element vector: three elements each corresponding to measurements of fundamental frequency (F0) and the first two formants (F1 and F2) at three different points (at the beginning, middle and end of the signal), and one element corresponding to duration. These vectors were the inputs for an unsupervised ANN (Self-Organizing Neural Network or SONN). Analyses were performed using both Matlab 7.0 (Matlab Neural Network Toolbox, Demuth and Beale 1993) and Statistica Neural Networks 7.1 (StatSoft, Trajan Software Ltd, 1996–2000).

Because unsupervised ANNs do not require target categories to be defined a priori, the architecture of the network has been characterized as a trial-and-error procedure (protocol described in Murray et al. 1998). We tested 12 networks using different combinations of learning rate (0.01 and 0.03), numbers of iterations (5,000 and 10,000), and numbers of neurons (40, 20, and 10). The number of units in a

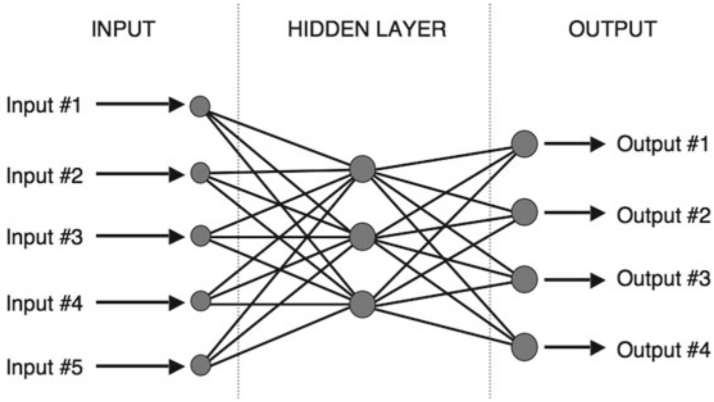

Fig. 34.1 Structure of the artificial neural network used in this study

SONN can be taken as the maximum number of potential categories that the network can identify during a training session. For each run, we analyzed the number of neurons activated that correspond to the number of categories recognized by the system of classification. The output neurons activated during the analysis represent the number of categories recognized by the network. To match these categories to the vocal types defined a priori for the black lemur vocal repertoire, we analyzed the weight vectors of each activated unit: the pattern of the spectral values (F0–F2) and the call durations were compared with the vocal repertoire description to assess concordance between the network results and the categories previously recognized.

Results

Application of Supervised ANNs to Differentiate Eulemur Species Vocalizations

We applied a one-hidden layer network with 70 units and 1,000 epochs (Fig. 34.1); this topology was the best network in the preliminary analyses, with 98% correct classifications in the training phase and 81% in the test.

The data set was randomly divided into a training (67% signals) and a test set (33%), and the analyses were run independently 15 times. The percentage correct prediction during the test phase varied from 67% in *E. coronatus* to 99% in *E. macaco*. The means and standard deviations for all vocal types analyzed are reported in Table 34.1: four out of five species were recognized with >80% correctly classified signals. The overall average reliability of the network was 89%.

Table 34.1 Number of signals, correct classifications (mean performance) and standard deviation for each of the five species within the genus *Eulemur* resulted in the application of the supervised neural network with 70 hidden units

	E. macaco	*E. fulvus*	*E. mongoz*	*E. coronatus*	*E. rubriventer*
N	115	35	33	23	16
Mean performance	99.42	84.44	80.00	66.67	84.00
Standard deviation	1.18	14.17	11.34	14.95	17.24

Identification of Distinct Calls Within the Eulemur macaco Vocal Repertoire Using an Unsupervised ANN

The weight vectors resulting from each activated unit in the analysis are shown in Fig. 34.2. The *x*-axis represents the ten elements making up the vectors, while the normalized *z*-scores are on the *y*-axis. Zero represents the mean, and values above and below zero are deviations from the mean. All networks but one identified five different categories. We thus compared the resulting five weight vectors with the sound structure of the different vocal types. The association between the weight vectors and the vocal categories recognized by the network is represented in Fig. 34.2. The five categories recognized by the unsupervised neural networks are:

1. *Alarm call*, represented by the weight vector W1, shows long duration and high frequency values in spectral parameters (above the mean).
2. *Long grunt,* represented by the vector W2, is characterized by low fundamental frequency and formant values and by long duration.
3. *Tonal* or *clear calls* (weight vector W3) have relatively high spectral values and average duration (around zero).
4. The composite vocalization *long grunt clear call* (W4) is made up of two parts: the first similar to the *long grunt* category (low spectral parameters) and the second with very high fundamental frequency and formant values.
5. A combined category including *hoot*, *grunt,* and *grunted hoot* (W5) has a spectral structure similar to the category 2 *long grunt* (low F0 and formant values) but is of much shorter duration.

Discussion

ANNs are widely used in bioacoustics to address numerous problems, but ours are the first applications of ANNs to nonhuman primate vocalizations. We applied different network typologies to address two common issues in the study of primate communication (1) the identification of closely related species and (2) the categorization of a species' vocal repertoire. The supervised network in the first case study distinguished five *Eulemur* species with an overall performance of almost 90%.

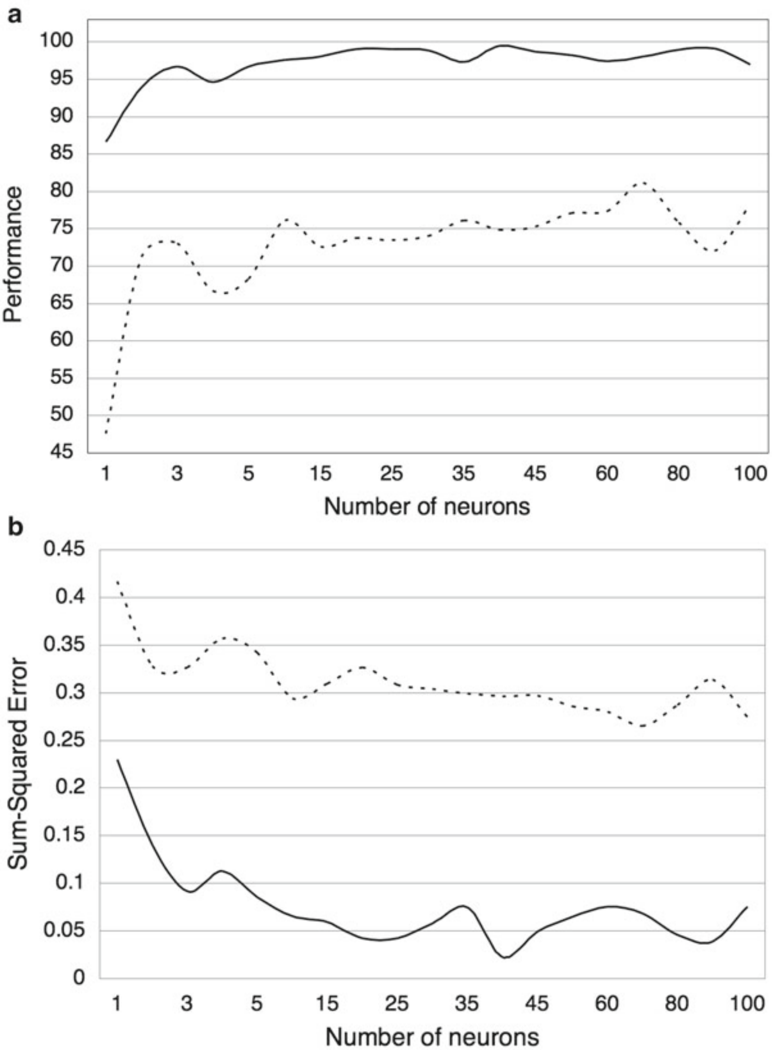

Fig. 34.2 Results of the SONN: the 10-element weight vectors of each of the five units resulting from the supervised network analysis. Zero (y-axis) represents the grand mean for each dimension (normalized z scores). Values above and below zero represent deviations from the mean. Each weight vector is associated with one or more of the call typologies recognized in this study (see the text for details). *LG* long grunt, *LGCC* long grunt clear call, *G* grunt, *GH* grunted hoot, *H* hoot

The unsupervised ANN used to categorize the black lemur vocal repertoire recognized five out of the seven vocal types defined a priori. The remaining two types (*grunted hoot* and *hoot*), with very similar structure, were merged under a larger category with *grunts*. This result is probably due to the relatively low number of appropriate signals in our sample.

Our study shows that ANNs provide an effective tool for acoustic categorization and have several advantages over other systems of classification. First, our approach requires fewer a priori assumptions than other statistical methods, reducing the degree of subjectivity in the classification procedure. Second, ANNs can deal with incomplete or noisy data, which is especially important in biological classification tasks, where unambiguous, discrete categories are often difficult to define. Finally, ANNs allow researchers to develop automatic or semiautomatic acoustic screening of animal vocalizations.

Our use of both supervised and unsupervised networks allows us also to draw some conclusions regarding their advantages and constraints for primate call classification. The application of supervised learning systems is obviously more efficient in recognizing discrete categories but requires the definition of a priori categories during the learning phase and is therefore more prone to subjective choices. Since ANNs are able to detect a classification rule autonomously during the learning phase and to reduce the noise in the signals, supervised ANNs are extremely powerful at classifying vocalizations into previously defined categories, even with low quality recordings (Placer and Slobodchikoff 2000). The main advantage of supervised ANNs is that these networks can be trained on a well-defined set of signals and then used to classify previously unseen records (Pozzi et al. 2010). Such networks are thus better suited to classification tasks where some aspect of pattern is already known. On the other hand, unsupervised ANNs (SONNs) are capable of detecting regularities and classifying inputs into discrete categories without numbers and types of outputs being defined a priori. Unsupervised neural networks reduce the role of the external operator in the classification process, increasing the objectivity of the classification system. SONNs are thus excellent for classification where no a priori knowledge of vocalizations is available (Murray et al. 1998).

The great potential of this approach indicates that more effort should be directed towards the development of neural networks to classify complex signals emitted by nonhuman primates. Such applications may be used to clarify various aspects of primate vocal behavior when discrete categories are hard to define using more conventional techniques. The integration of behavioral and statistical approaches with ANNs can increase greatly the efficiency, objectivity, and biological significance of vocal classification, and so advance our efforts to understand the evolution of vocal behavior in nonhuman primates.

Acknowledgments We thank Judith Masters for providing the opportunity to contribute to this volume. We also thank Cesare Avesani Zaborra and Parco Natura Viva for their continued support.

References

Adi K, Johnson MT, Osiejuk TS (2010) Acoustic censusing using automatic vocalization classification and identity recognition. J Acoust Soc Am 127:874–883

Campbell GS, Gisiner RC, Helweg DA, Milette LL (2002) Acoustic identification of female Steller sea lions (*Eumetopias jubatus*). J Acoust Soc Am 111:2920–2928

Chesmore D (2004) Automated bioacoustic identification of species. An Acad Bras Cienc 76:435–440

Chesmore ED, Ohya E (2004) Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. Bull Entomol Res 94:319–330

Dawson MR, Charrier I, Sturdy CB (2006) Using an artificial neural network to classify black-capped chickadee (*Poecile atricapillus*) call note types. J Acoust Soc Am 119:3161–3172

Demuth H, Beale M (1993) Neural network toolbox. The MathWorks, Inc., Natick, MA

Derégnaucourt S, Guyomarc'h J-C, Richard V (2001) Classification of hybrid crows in quail using artificial neural networks. Behav Proc 56:103–112

Gamba M, Giacoma C (2005) Key issues in the study of primate acoustic signals. J Anthropol Sci 83:61–87

Ghirlanda S, Enquist M (1998) Artificial neural networks as models of stimulus control. Anim Behav 56:1383–1389

Gosset D, Fornasieri I, Roeder J-J (2002) Acoustic structure and contexts of emission of vocal signals by black lemurs. Evol Commun 4:225–251

Green SR, Mercado E, Pack AA, Herman LM (2011) Recurring patterns in the songs of humpback whales (*Megaptera novaeangliae*). Behav Proc 86:284–294

Jennings N, Parsons S, Pocock MJO (2008) Human vs. machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. Can J Zool 86:371–377

Kohonen T (1988) The 'neural' phonetic typewriter. Computer 21:11–22

Lopes MT, Gioppo LL, Higushi TT, Kaestner CAA, Silla Jr, CN, Koerich AL (2011) Automatic bird species identification for large number of species. IEEE international symposium on multimedia, pp. 117–122

Macedonia JM, Stanger KF (1994) Phylogeny of the Lemuridae revisited: evidence from communication signals. Folia Primatol 63:1–43

Mercado E, Kuh A (1998) Classification of humpback whale vocalizations using a self-organizing neural network. Proceedings of the 1998 IEEE joint conference on neural networks, pp 1584–1589. doi: 10.1109/IJCNN.1998.686014

Murray SO, Mercado E, Roitblat HL (1998) The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. J Acoust Soc Am 104:3626–3633

Nickerson CM, Bloomfield LL, Dawson MR, Sturdy CB (2006) Artificial neural network discrimination of black-capped chickadee (*Poecile atricapillus*) call notes. J Acoust Soc Am 120:1111–1117

Parsons S (2001) Identification of New Zealand bats (*Chalinolobus tuberculatus* and *Mystacina tuberculata*) in flight from analysis of echolocation calls by artificial neural networks. J Zool 253:447–456

Placer J, Slobodchikoff CN (2000) A fuzzy-neural system for identification of species-specific alarm calls of Gunnison's prairie dogs. Behav Proc 52:1–9

Pozzi L, Gamba M, Giacoma C (2010) The use of artificial neural networks to classify primate vocalizations: a pilot study on black lemurs. Am J Primatol 72:337–348

Reby D, Lek S, Dimopoulos I, Joachim J, Lauga J, Aulagnier S (1997) Artificial neural networks as a classification method in the behavioural sciences. Behav Proc 40:35–43

Zimmermann A (1995) Artificial neural networks for analysis and recognition of primate vocal communication. In: Zimmermann E, Newman JO, Jürgens UWE (eds) Current topics in primate vocal communication. Plenum, New York, pp 29–46