Animal Welfare 2004, 13: 105-110
ISSN 0962-7286

Automated recording of stress vocalisations as a tool to document impaired welfare in pigs

PC Schön, B Puppe and G Manteuffel*

Research Institute for the Biology of Farm Animals (FBN), Behavioural Physiology Unit, Wilhelm-Stahl-Allee 2, D-18196 Dummerstorf, Germany
* Contact for correspondence and requests for reprints: manteuff@fbn-dummerstorf.de

Abstract

The vocalisations of animals are results of particular emotional states. For example, the stress screams of pigs may be indicators of disturbed welfare. Our objective was to develop a system to monitor and record levels of stress calls in pigs, which could be employed in environments of breeding, transportation and slaughter. Using a combination of sound analysis by linear prediction coding and artificial neural networks, it was possible to detect the stress vocalisations of pigs in noisy pig units with few recognition errors (<5%). The system (STREMODO: stress monitor and documentation unit) running on PCs is insensitive to environmental noise, human speech and pig vocalisations other than screams. As a stand-alone device it can be routinely used for the objective, non-invasive measurement of acute stress in various farming environments. The system delivers reliable, reproducible registrations of stress vocalisations. Its detection quality in commercial systems was found to correlate well with that of human experts. STREMODO is particularly well-suited for comparisons of housing and management regimes. Since the system can be trained to recognise various animal vocalisations, its use with other species is also well within its scope.

Keywords: *animal welfare, neural networks, pig, stress, vocalisation, welfare assessment*

Introduction

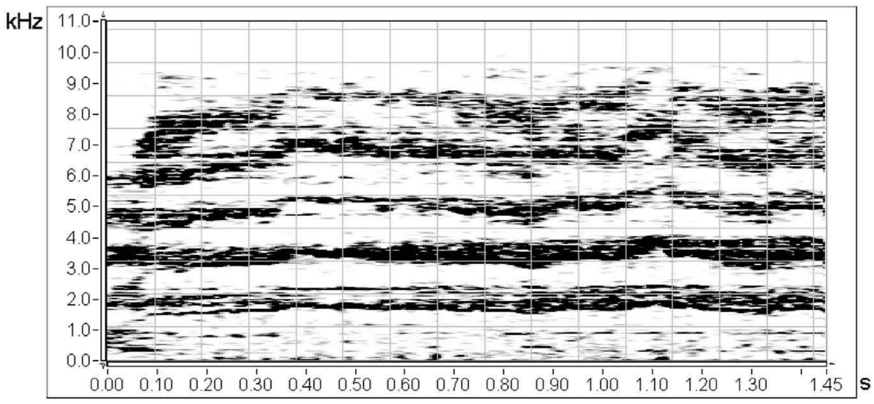
Animal vocalisations are actively produced sounds that are able to indicate specific emotional states occurring spontaneously or induced by external events and contexts (Jürgens 1979; Weary & Fraser 1995; Schrader & Todt 1998). In contrast to human language, where meaning is attributed in a discursive process, animal vocalisations are usually produced according to fixed programs developed during phylogeny and achieved in ontogeny. Hence, animal vocalisations are specifically attributed to particular inner states, maybe with the exception of apes, where a more creative use might be possible under certain circumstances.

Changing emotional states can be the cause or the effect of physiological and/or behavioural reactions. The latter can be measured and taken as a reference for the species-specific and individual meaning of a vocalisation (Schrader & Rohn 1997). The dependence of animal vocalisation on psycho-physiological conditions makes sound analysis a well-suited tool for non-invasive judgments of welfare and stress. The procedure consists of two elements: a syntactic, which describes the formal features, and a semantic, which derives the meaning from coincident behavioural and physiological reactions.

In farm animals, housing conditions may produce stress and may impair welfare if biological needs remain transiently or permanently unfulfilled. Pigs in particular can be easily stressed, which leads to an activation of sensory and limbic brain centres and eventually to reactions of the sympathico-adrenomedullary and the hypothalamo–pituitary–adrenal hormonal axes. In parallel, motor centres that control behavioural components of the stress reaction may be triggered. One of these is vocalisation, which in acutely stressed pigs is displayed as a rather sustained scream containing high frequency components (Schön *et al* 2001) (Figure 1).

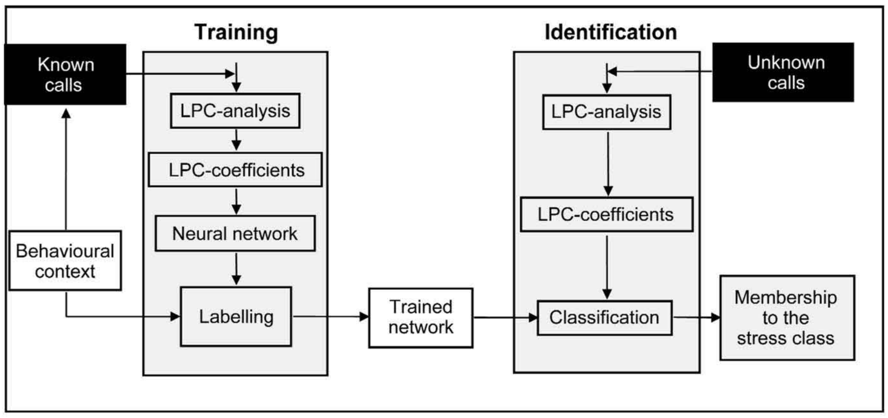
Up to now, methods of analysis that allowed the recognition of arbitrary sounds did not exist because general solutions for extraction and classification were virtually absent. False or too few features may result in insufficient classification and too many features can overload most computers, especially if real-time capability is demanded. Hence, differentiated, well-adapted procedures are required for nearly every classification task (Schön *et al* 1999). Here, a system termed STREMODO (STREss MOnitor and DOcumentation unit) is presented that allows the recognition of the stress screams of domestic pigs. It combines Linear Prediction Coding (LPC) with an artificial neural network, as has been similarly proposed for the detection of pigs' cough (Moshou *et al* 2001). The proposed system is able to calculate and record in real-time while disregarding other vocalisations and sounds (patent pending). The results are documented in a time-related list.

Figure 1



Sonogram (from Fast Fourier Transformation) of a growing pig's stress vocalisation. The grey scale represents the intensities of the respective frequency bands.

Figure 2



Training and application phases of STREMODO.

Methods

LPC is used to obtain formal parameters of pigs' screams (Pham & Le Breton 1991). In a first step, the continuous analogue time signal from a microphone is digitised and sampled using time windows of 100 ms duration. From the series of samples, the LPC coefficients are calculated, as described in detail elsewhere (Schön *et al* 2001). The LPC procedure is in formal equivalence to the source-filter model of the vocal tract (Fant 1970). The LPC coefficients correspond to the filter coefficients of the vocal tract. Hence the model is sensitive to variation of the resonance frequencies of the vocal tract and indirectly to the motor efforts that are required to obtain them. Polynomial development allows the calculation of the LPC spectrum from the LPC coefficients. The resulting LPC sonograms resemble sonograms based on the Fast Fourier Transformations (FFT), but are based on distinctly fewer parameters. For the detection of pigs' stress screams, 12 LPC coefficients, equivalent to the first six resonance frequencies, have been found to be sufficient.

For real-time classification of the LPC spectra, a neural network is applied that has been trained with the LPC coefficients of pigs' stress screams against non-scream vocalisations and against various other types of noise. After training, the system should be able to detect the stress vocalisations of pigs in arbitrary contexts and environments (Figure 2).

We used a four-layer Perceptron type network (Rosenblatt 1962), taking the LPC coefficients of single time windows (100 ms) as 12-dimensional input vectors. The Perceptron consisted of an input layer (12 neurons), two hidden layers (120 and 60 neurons) and an output layer (2 neurons). The activation functions were sigmoid and training was executed with an error back-propagation algorithm.

At the output of the neural network it is indicated whether or not the vector has been classified as a stress scream. The percentage of stress vectors is stored in time bins, each containing 100 vectors (= 10 s), and can be displayed immediately or later as a list or as a graphic (Figure 3).

The system was trained with the stress calls of pigs (German Landrace) of various ages that were bred and housed in the pig plant of the Research Institute for the Biology of Farm Animals, Dummerstorf, Germany. The vocalisations were recorded in a noise-reduction chamber with a Sennheiser MKE 46 microphone and were stored on a Sony DCT-790 DAT-Recorder (Schön *et al* 1998). The stressors applied were: 1) immobilisation of piglets by holding them upright at the thorax and keeping them above the floor, 2) immobilisation of growing pigs by forcing them on their backs, and 3) immobilisation of sows, which were held with a nose snare. These procedures not only reliably elicit stress vocalisations as indicators of high levels of

https://doi.org/10.1017/S096272860002683X Published online by Cambridge University Press

Table 1 Classification results of the STREMODO prototype.

Animals (age)	n	Calls/noise	Number of LPC vectors	Type	Misclassification (%)
Piglets (2nd week)	10	Cries	1904	Stress	0.58
Piglets (5th week)	7	Cries	2476	Stress	0.85
Piglets (2nd week)	3	Grunts	171	No stress	2.34
Piglets (5th week)	3	Grunts	245	No stress	2.04
Sows (1st lactation)	5	Suckling grunts	60	No stress	1.67
Noise	–	Farm noise	1706	No stress	1.23

excitement, but also elicit hormonal and neurophysiological reactions that indicate stress, ie increased cortisol, adrenaline and noradrenaline levels (Kanitz *et al* 1998, 1999; Otten *et al* 2001; Tuchscherer *et al* 2002).

The standardised stress situations provided prototypical, generalised stress vocalisations by the pooling of the examples in the trained neural network. To date there is no convincing evidence from the literature that the stress screams of pigs differ depending on the type of stress being experienced. If such data become available in the future, the system will be trained with respective examples and then its capability to discriminate between different types of stress scream will be assessed. In addition to the recording of stress screams, pig grunts and different types of technical noise occurring in pig units were used as training examples of non-stress sounds. The result is a classification neural network that discriminates between stress and non-stress acoustic input. Only inputs that belong to the class 'stress' are forwarded to the output and result in a registration on the display and a data record in which the occurrence of stress screams is related to time.

Subsequently, STREMODO was tested for its applicability and performance in a more realistic situation: a commercial pig farm in Mecklenburg-Vorpommern, Germany. We compared the occurrence of stress screams in two different feeding regimes. In one case, the animal-to-feeding-place ratio was 6:1 with 24 fattening pigs (average weight 50 kg) in a 4.85×3.90 m^2 concrete/slatted floor pen (0.8 m^2/animal; sensor-aided system where food was added at a fixed time if the 1.5 m trough was emptied). In the second system, enough food was supplied once daily at an animal-to-feeding-place ratio of 1:1 (4 m trough) in a 4.0×1.8 m^2 slatted floor pen with 11 animals (average weight 80 kg; 0.7 m^2/animal). For measurement, the microphone of STREMODO was placed 2 m above floor level in the centre of the pen and the system was allowed to record for several hours. In parallel, a video system recorded the animals' behaviour and vocalisations. Since the only aim of this experiment was to demonstrate the technical suitability of the system in practical farming, the different numbers of animals in the two situations was not considered.

For comparisons of judgements of STREMODO with those of six human experts, a 10 min recording from a commercial pig unit was presented. Humans were individually instructed to press a pre-defined key of a computer keyboard for as long as they believed that they could hear stress-like screaming. Their responses were stored in the same 10 s bins (n = 60) as those used by STREMODO, and their responses were correlated with the system's result, which was recorded afterwards.

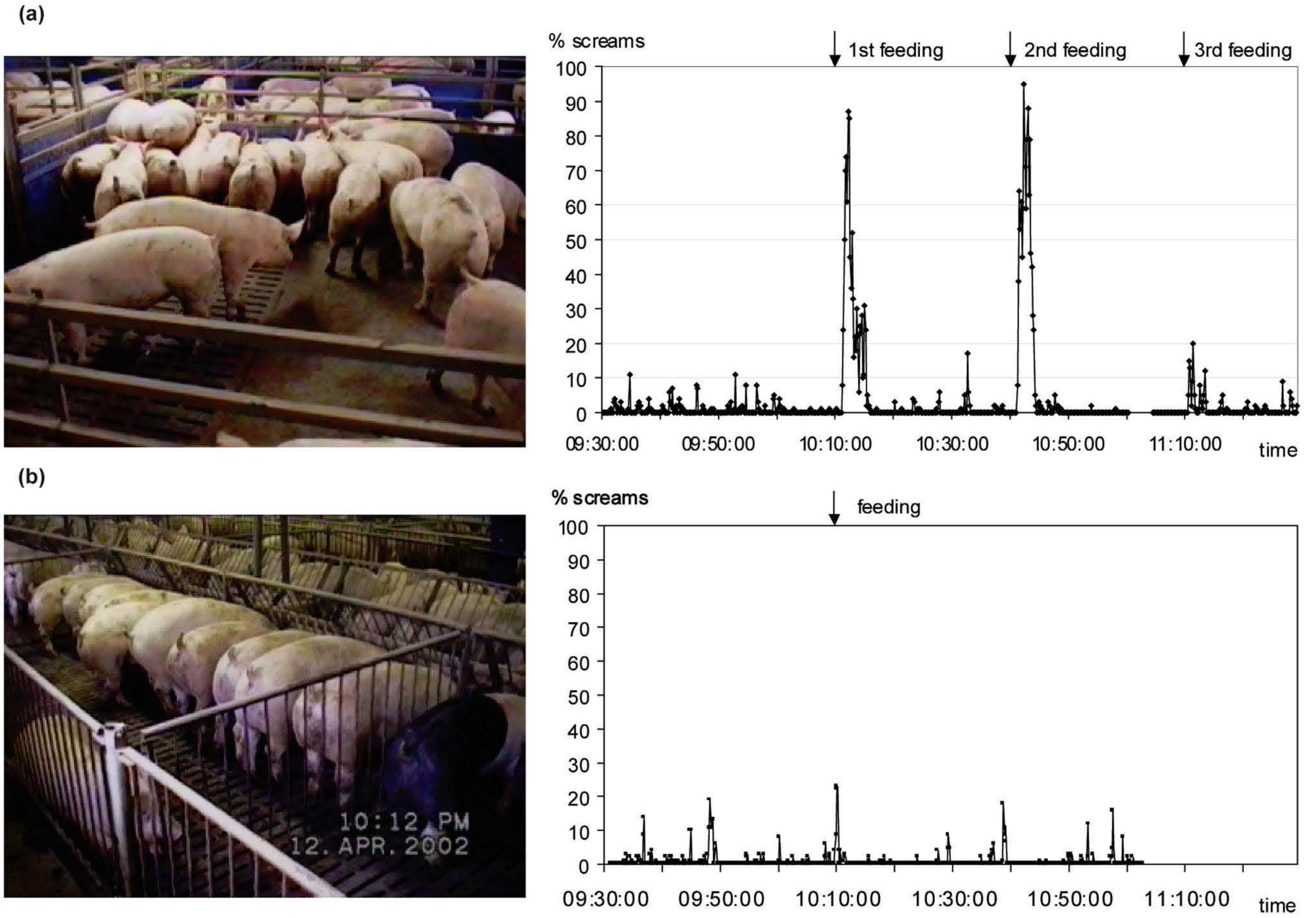
Results and performance

The system was tested in a first run using vocalisations and noise samples that were recorded together with the training sets but were not included within them (Schön & Manteuffel 2001; Schön *et al* 2001). The vocalisations and noise were played back via loudspeaker and classified in real-time. From the calls of the animals used, the LPC coefficients were calculated for each time window (100 ms). Since in this test situation the origin of the input data (ie from stressed animals, other animals or noise) was known, the relation between the network's classification in the right and wrong class determined the misclassification rate in percent. The result is shown in Table 1, demonstrating that very few misclassifications of stress calls (<1%) occurred, and that few acoustical non-stress events were incorrectly classified as stress calls (<5%).

In the practical farming application, STREMODO recorded increased screaming during the first two feeding sessions at the 6:1 animal-to-feeding-place ratio (Figure 3a). A qualitative evaluation of the respective behavioural recordings displayed a high level of competition between the animals for the food in these situations, in which virtually all of the animals at the trough were temporarily involved. In the third feeding episode, the animals displayed less screaming and less competition, probably because they were less hungry. In contrast, in the stall with the 1:1 animal-to-feeding-place ratio, the animals showed no detectable increase in screaming at feeding time compared to other times when occasional stress-like screams during aggressive interactions were detected (Figure 3b). These results demonstrate that STREMODO is able to reveal the level of stress vocalisations in noisy farm environments, and can be used for the continuous monitoring of porcine stress calls. When compared to six expert humans who judged the same stress screaming of pigs from a 10 min recording in a commercial unit (see Figure 4), the average correlation between STREMODO and the humans' decisions was 0.84 (Spearman correlation coefficient; $P < 0.001$; range of

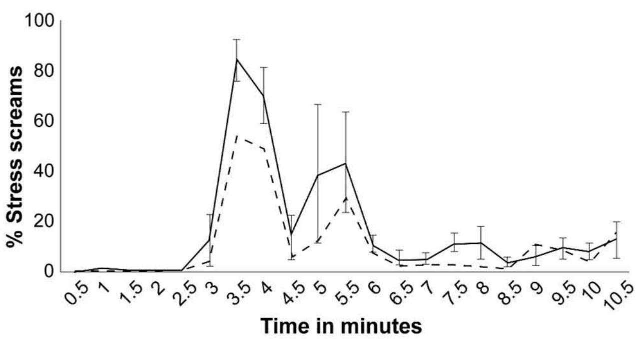
https://doi.org/10.1017/S096272860002683X Published online by Cambridge University Press

Figure 3



The temporal percentage of the duration of stress vocalisations as recorded by STREMODO on two different days (ordinate: % of sound classified as stress screaming within each time bin, also see Methods): (a) Restrictive feeding using a 6:1 animal-to-feeding-place ratio (a further measurement was added for the third additional feeding episode on the same day), and (b) Trough feeding using a 1:1 animal-to-feeding-place ratio. Vertical arrows indicate feeding times.

Figure 4



Rate of stress screaming in a 10 min recording from a commercial pig unit as judged by six human experts (solid line, mean and standard deviation) and by STREMODO (broken line).

0.78–0.87). The expert humans always had difficulty in exactly determining the beginning and end of individual screams and in concentrating even for the relatively short duration of the recording.

Discussion

The STREMODO system for the recognition of stress as expressed by vocalisations, consists of an initial step of extraction of the elementary characteristic syntactic elements using LPC analysis and a subsequent classification by a trained artificial neural network. Its performance is well within the range of benchmarks given in the respective literature (cf Dreyfus 1992; Gramß & Strube 1992; Moshou *et al* 2001). Hence, the system is suited to detect the stress calls of pigs with a high level of accuracy, as was also shown by its practical application in a farming environment.

STREMODO indicates the duration and intensity (as a percentage of stress scream vectors found in each 10 s) of stress screaming, but not the number of individual screams. This measure was considered to be a better indication of effective distress than the number of separate screaming episodes, assuming that prolonged stress is more relevant in terms of welfare than are occasional short screams (Blokhuis *et al* 1998; Manteuffel & Puppe 2000; Manteuffel 2002).

Two specific characteristics of STREMODO are particularly advantageous for its tasks. First, the system is considerably insensitive to noise and to the non-stress vocalisations that constantly occur in pig units. Second, being computer-based it is always ready for monitoring because of its real-time performance, and it delivers a list of objective, reproducible results for later evaluation, even at times when no human surveillance is provided. At the present development stage, STREMODO runs on a laptop or desktop computer. For most practical purposes, however, a special low-cost stand-alone unit that is robust against the dust and harmful gas that is frequently found in pig units would be more favourable. By choosing an appropriate directional characteristic of the microphone, it is also possible to restrict recording to particular pen areas within a unit. A wide characteristic allows the recording of all stress screams from an area that is only limited by a sufficiently high sound pressure level. All stress screams that originate from that area are registered independent of their (above threshold) volume since the sound analysis procedure of STREMODO does not depend on volume. Alternatively, a directionally restricted microphone characteristic allows the recording of stress screams from a specific area of interest within housing.

In part, the efficiency of STREMODO is due to the relatively simple structure of the stress screams of pigs, which are quite sustained and not very modulated in frequency. If modulations occur at all, their temporal order is largely irrelevant. Since vocalisations that may indicate emotional states in other mammalian farm animals are likewise comparatively simply structured, rapid adaptation of the system to other species is possible (eg for the isolation calls and calls during oestrus of goats, sheep and cows).

Animal welfare implications

The ability to automatically record the time and level of stress calls makes STREMODO suitable for objective welfare monitoring in large commercial pig units where it will be able to tell the farmer when excessive stress is occurring. This should eventually lead farmers to look for possible causes of stress and to eliminate them in order to improve product quality and animal welfare. Furthermore, objectively documented low stress levels in pig production can be demonstrated to interested consumers.

Further applications include the monitoring of stress during transport and slaughter, and comparisons of stress levels in different housing systems and between different strains of pig to assess their susceptibility to stress. For large-scale commercial application, tolerable stress levels, as measured by STREMODO, must be determined on the basis of scientific research that correlates these levels with relevant physiological and welfare parameters, because of course stress screaming alone is not sufficient to indicate the impaired welfare of pigs. Stress screaming has to be interpreted in the contexts in which it occurs, because any single parameter is to some degree ambiguous with respect to animal well-being (Broom 1991; Clark *et al* 1997; Manteuffel & Puppe 2000). However, if the level of screaming in particular husbandry situations is excessive, this might be taken as a serious indicator of prolonged and/or repetitive stress, which is likely to be psychologically and physically detrimental.

Acknowledgement

The authors gratefully acknowledge Olaf Tober's (Landesforschung für Landwirtschaft, Mecklenburg-Vorpommern) cooperation in recording vocalisations on a commercial farm.

References

Blokhuis H J, Hopster H, Geverink N A, Korte S M and van Reenen C G 1998 Studies of stress in farm animals. *Comparative Haematology International 8*: 94-101

Broom D M 1991 Animal welfare: concepts and measurement. *Journal of Animal Science 69*: 4167-4175

Clark J D, Rager D R and Calpin J D 1997 Animal well-being: I. General considerations. *Laboratory Animal Science 47*: 564-570

Dreyfus G 1992 Neural networks for the automatic recognition of handwritten digits. In: Schuster H G (ed) *Applications of Neural Networks* pp 35-60. Verlag Chemie: Weinheim, Germany

Fant G 1970 *Acoustic Theory of Speech Production.* Mouton: The Hague, The Netherlands

Gramß T and Strube H W 1992 Word recognition with a fast learning neural net. In: Schuster H G (ed) *Applications of Neural Networks* pp 223-237. Verlag Chemie: Weinheim, Germany

Jürgens U 1979 Vocalisation as an emotional indicator. A neuroethological study in the squirrel monkey. *Behaviour 69*: 88-117

Kanitz E, Manteuffel G and Otten W 1998 Effects of weaning and restraint stress on glucocorticoid receptor binding capacity in limbic areas of domestic pigs. *Brain Research 804*: 311-315

Kanitz E, Otten W, Nürnberg G and Brüssow K P 1999 Effects of age and maternal reactivity on the stress response of the pituitary-adrenocortical axis and the sympathetic nervous system in neonatal pigs. *Animal Science 68*: 519-526

Manteuffel G 2002 Central nervous regulation of the hypothalamic-pituitary-adrenal axis and its impact on fertility, immunity, metabolism and animal welfare — a review. *Archiv für Tierzucht Dummerstorf 45*: 575-595

Manteuffel G and Puppe B 2000 Animal well-being in husbandry and coping with stress. *Archiv für Tierzucht Dummerstorf 43*: 140-143 (Special issue)

Moshou D, Chedad A, Van Hirtum A, De Baerdemaeker J, Berkmans D and Ramon H 2001 Neural recognition system for swine cough. *Mathematics and Computers in Simulation 56*: 475-487

Otten W, Kanitz E, Tuchscherer M and Nürnberg G 2001 Effects of prenatal restraint stress on hypothalamic-pituitary-adrenocortical and sympatho-adrenomedullary axis in neonatal pigs. *Animal Science 73*: 279-287

Pham D T and Le Breton A 1991 Levinson-Durbin type algorithm for continuous-time autoregressive models and applications. *Mathematics Control Signals Systems 4*: 69-79

Rosenblatt F 1962 *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* Spartan Books: Washington, DC, USA

Schön P C and Manteuffel G 2001 Nichtinvasive Beurteilung emotionaler Stressbelastung von Nutztieren mittels Vokalisationserkennung durch ein künstliches neuronales Netzwerk. In: Schäfer D and von Borell E (eds) *Proceedings of the 15th IGN Congress: Tierschutz und Nutztierhaltung* (*Animal Welfare and Farm Animal Housing*) pp 104-109. Internationale Gesellschaft für Nutztierhaltung: Halle, Germany [Title translation: Non-invasive judgement of emotional stress of farm animals by vocalisation recognition using an artificial neuronal network]

Schön P C, Puppe B and Manteuffel G 1998 A sound analysis system based on LabVIEW® applied to the analysis of suckling grunts of domestic pigs (*Sus scrofa*). *Bioacoustics 9*: 119-133

Schön P C, Puppe B and Manteuffel G 1999 Common features and individual differences in nurse grunting of domestic pigs (*Sus scrofa*): a multi-parametric analysis. *Behaviour 136*: 49-66

Schön P C, Puppe B and Manteuffel G 2001 Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (*Sus scrofa*). *Journal of the Acoustical Society of America 110*: 1425-1431

Schrader L and Rohn C 1997 Lautäußerungen von Hausschweinen als Indikator für Stressreaktionen. *Landbauforschung Völkenrode 47*: 89-95 [Title translation: The quality of pig vocalisations as an indicator of stress response]

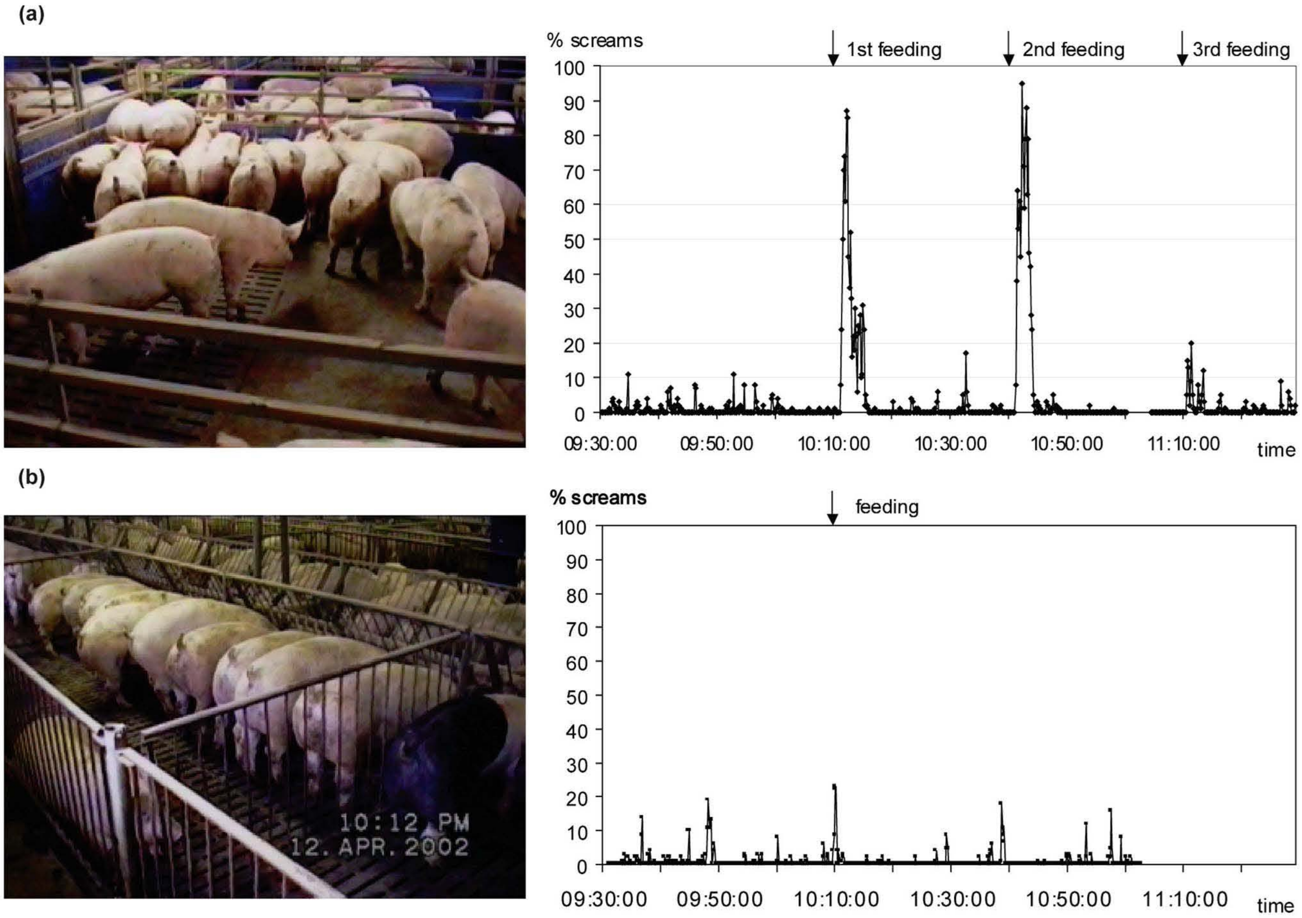
Schrader L and Todt D 1998 Vocal quality is correlated with levels of stress hormones in domestic pigs. *Ethology 104*: 859-876

Tuchscherer M, Kanitz E, Otten W and Tuchscherer A 2002 Effects of prenatal stress on cellular and humoral immune responses in neonatal pigs. *Veterinary Immunology and Immunopathology 86*: 175-203

Weary D M and Fraser D 1995 Calling by domestic piglets: reliable signals of need? *Animal Behaviour 50*: 1047-1055

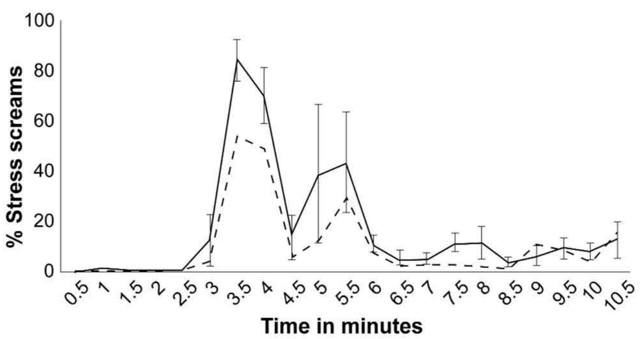
https://doi.org/10.1017/S096272860002683X Published online by Cambridge University Press

Figure 3



The temporal percentage of the duration of stress vocalisations as recorded by STREMODO on two different days (ordinate: % of sound classified as stress screaming within each time bin, also see Methods): (a) Restrictive feeding using a 6:1 animal-to-feeding-place ratio (a further measurement was added for the third additional feeding episode on the same day), and (b) Trough feeding using a 1:1 animal-to-feeding-place ratio. Vertical arrows indicate feeding times.

Figure 4



Rate of stress screaming in a 10 min recording from a commercial pig unit as judged by six human experts (solid line, mean and standard deviation) and by STREMODO (broken line).

0.78–0.87). The expert humans always had difficulty in exactly determining the beginning and end of individual screams and in concentrating even for the relatively short duration of the recording.

Discussion

The STREMODO system for the recognition of stress as expressed by vocalisations, consists of an initial step of extraction of the elementary characteristic syntactic elements using LPC analysis and a subsequent classification by a trained artificial neural network. Its performance is well within the range of benchmarks given in the respective literature (cf Dreyfus 1992; Gramß & Strube 1992; Moshou *et al* 2001). Hence, the system is suited to detect the stress calls of pigs with a high level of accuracy, as was also shown by its practical application in a farming environment.

STREMODO indicates the duration and intensity (as a percentage of stress scream vectors found in each 10 s) of stress screaming, but not the number of individual screams. This measure was considered to be a better indication of effective distress than the number of separate screaming episodes, assuming that prolonged stress is more relevant in terms of welfare than are occasional short screams (Blokhuis *et al* 1998; Manteuffel & Puppe 2000; Manteuffel 2002).

Two specific characteristics of STREMODO are particularly advantageous for its tasks. First, the system is considerably

https://doi.org/10.1017/S096272860002683X Published online by Cambridge University Press