

Evidence for a perception of prosodic cues in bat communication: contact call classification by *Megaderma lyra*

Simone Janßen · Sabine Schmidt

Received: 13 October 2008 / Revised: 5 April 2009 / Accepted: 5 April 2009 / Published online: 23 April 2009
© Springer-Verlag 2009

Abstract The perception of prosodic cues in human speech may be rooted in mechanisms common to mammals. The present study explores to what extent bats use rhythm and frequency, typically carrying prosodic information in human speech, for the classification of communication call series. Using a two-alternative, forced choice procedure, we trained *Megaderma lyra* to discriminate between synthetic contact call series differing in frequency, rhythm on level of calls and rhythm on level of call series, and measured the classification performance for stimuli differing in only one, or two, of the above parameters. A comparison with predictions from models based on one, combinations of two, or all, parameters revealed that the bats based their decision predominantly on frequency and in addition on rhythm on the level of call series, whereas rhythm on level of calls was not taken into account in this paradigm. Moreover, frequency and rhythm on the level of call series were evaluated independently. Our results show that parameters corresponding to prosodic cues in human languages are perceived and evaluated by bats. Thus, these necessary prerequisites for a communication via prosodic structures in mammals have evolved far before human speech.

Keywords Acoustic communication · Affect cues · Bat · Evolution of prosody · Identity cues

Introduction

Prosody is an essential trait of human languages and is expressed by variations of frequency, rhythm and amplitude parameters in the speech signal (Lakshminarayanan et al. 2003; Rickheit et al. 2003). Prosody is not only used to accent the semantic structure of the signal, i.e. to serve linguistic functions, but also to transmit identity and the emotional state of the speaker (Scherer 1989). Whereas the linguistic functions of prosody are characteristic for human languages, the paralinguistic cues (cf. Belin et al. 2004; Bußmann 1990) transmitting identity and emotional state are to some extent based on the morphological structure of the vocal tract and on physiological condition (e.g. Fitch and Hauser 2002; Rendall 2003), which makes it likely that a communication of identity and affect via these structures evolved far before human language (Scherer 1989). Research aiming to explore the origin of prosodic cue perception revealed language discrimination abilities in cotton-top tamarins and rats that were similar to those found in human infants (Ramus et al. 2000; Tincoff et al. 2005; Toro et al. 2003). While these studies showed that the capacity to discriminate languages from different linguistic rhythmic classes depends on perceptual abilities that evolved at least as far back as rodents, it remains unclear how mammals make use of call parameters corresponding to paralinguistic prosodic cues in intra-specific communication. A necessary prerequisite for a communication of identity and affect is the ability to perceive and evaluate differences in the respective cues. To address this question, we studied the classification of a complex communication call that was varied systematically in rhythm and frequency parameters in a bat.

Socially living bat species are an ideal model to investigate the functional elements and evolutionary roots of

S. Janßen · S. Schmidt (✉)
Institute of Zoology, University of Veterinary Medicine
Hanover, Bünteweg 17, 30559 Hanover, Germany
e-mail: sabine.schmidt@tiho-hannover.de

acoustic communication, due to the exceptional role of acoustic information for these nocturnal mammals not only in spatial orientation (e.g. Neuweiler 2000) but also for social organization (Kulzer 2005), and their early separation from other groups in evolution (Eizirik et al. 2001; Koopman 1994; Kumar and Hedges 1998; Murphy et al. 2001). In addition, psychoacoustic time and frequency constants determined in passive listening paradigms for bats without an acoustic fovea (cf. Neuweiler 2000) are similar to those in other mammals (e.g. Fay 1988; Schmidt 2000). In a given behavioural situation, many bat species use specific vocalizations (e.g. Barclay and Thomas 1979; Barclay et al. 1979; Behr and von Helversen 2004; Clement et al. 2006; Davidson and Wilkinson 2004; Leippert 1994; Pfalzer and Kusch 2003), which have been shown to carry group- (Boughman 1997; Scherrer and Wilkinson 1993; Wilkinson and Boughman 1998) and individual-specific signatures (e.g. Brown 1976; Esser and Schmidt 1989; Gelfand and McCracken 1986; Leippert et al. 2000), and to indicate the affect of the sender (Bastian and Schmidt 2008). Likewise, a perception of identity (Balcombe 1990; Bohn et al. 2007), group affiliation (Boughman and Wilkinson 1998) and affect (Russ et al. 2005) has been reported for bats.

However, the interplay of parameters governing social call classification is still widely unknown. The aim of the present study is to investigate this interplay in the Indian False Vampire bat, *Megaderma lyra*, using contact call perception as a model. This species emits series of contact calls (Goymann et al. 1999) in situations of social isolation to attract conspecifics. These situations include reunion at the day roost in the morning (Goymann et al. 1999), as well as reunions of mother and pup, and of individuals of both sexes at the night roosts where stable long-term groups of different day roosts meet to interact (Schmidt 2005), e.g. for courtship and mating (S. Schmidt, personal observations). In all these situations, the identity of the sender, coded in individual-specific signatures (Doerrie et al. 2001), and the urgency of the situation in which contact calls are emitted are of vital interest for the communication partners. Frequency, and rhythm, within the multi-syllabic contact call (Fig. 1a), and rhythm across calls in a series, which are parameters corresponding to paralinguistic prosodic cues in human speech, may vary considerably within and between individuals and may be used to express identity and affect.

In a two-alternative, forced choice procedure we trained *M. lyra* to discriminate between two synthetic contact call series differing in frequency, rhythm on level of calls and rhythm on level of call series. We assumed that the bats evaluated these three parameters for classification and that all parameters contributed to the classification. To test these hypotheses, we presented bats with stimuli differing

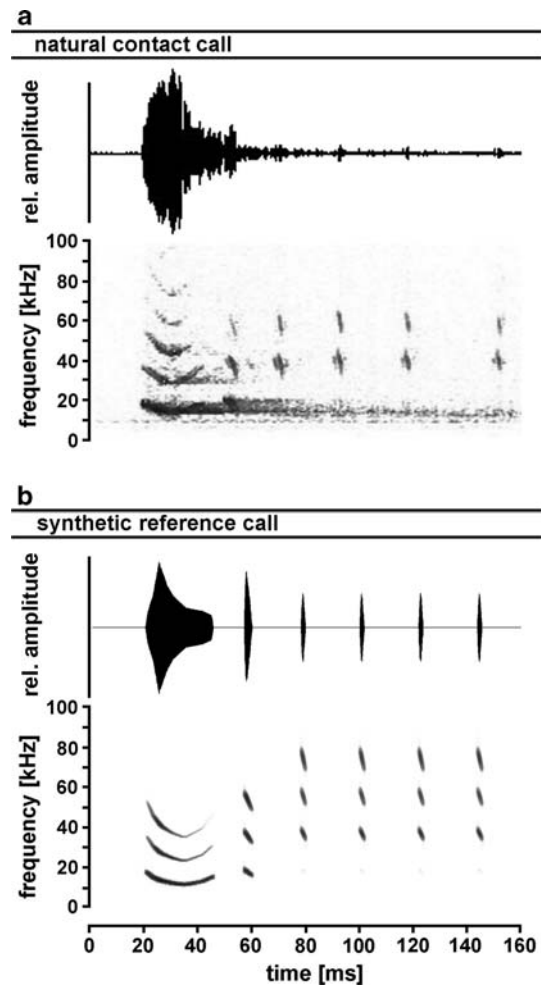


Fig. 1 Oscillogram and sonogram of a natural contact call of *M. lyra* (a), and of the synthetic reference call (b). The natural contact call was recorded from an isolated individual temporarily kept in an outdoor flight cage in its foraging habitat. The synthetic reference call was based on the median characteristics of 446 natural contact calls from eight Sri Lankan *M. lyra*

in one, or two, of the above parameters and compared the performance of the bats with the predictions from models assuming that one, a combination of two or all three parameters were used for classification. The results are discussed with respect to the physiological coupling of the parameters in call production and a potential formation of perceptual units by the receiver, and to the perception of parameters corresponding to paralinguistic prosodic cues across different mammalian orders.

Methods

Animals

Four adult *M. lyra*, two males and two females originating from Sri Lanka, took part in the experiments. One male

died before learning the task. The other male did not perform constantly in the experiment even after 9 months of training. Thus, experiments were successfully completed with two female bats. The animals were kept in a weakly illuminated room ($2.1 \times 2.9 \times 2.2$ m) with free access to water supplemented with vitamins and minerals [Polybion N (Merck), e-Mulsin (Mucos), Basica (Protina GmbH)]. Animals were trained once per day for 12 days followed by a 2-day training break. Throughout training periods, they were fed with mealworms (*Tenebrio molitor*) as a reward in the experimental sessions. During training breaks, they were fed with juvenile mice and grasshoppers (*Locusta*, *Schistocerca*).

Stimuli

Synthetic stimuli were used to achieve maximum stimulus control. As a basis for stimulus specification, we referred to 446 calls emitted in an isolation context by eight adult Sri Lankan bats of both sexes. The natural contact call typically consisted of three different multiharmonic syllable types. More than 80% of the natural contact calls of Sri Lankan bats started with a U-shaped syllable with a median peak-frequency of 13.9 kHz (ranging from 12.0 to 18.9 kHz) and a median duration of 25.5 ms. After a median interval of 11.8 ms, a downward frequency modulated syllable followed, with a median peak-frequency of 18.3 kHz (ranging from 14.0 to 22.65 kHz) and a median duration of 3.9 ms. After a median interval of 16.3 ms, 1–14 downward frequency modulated syllables of a third type with a median duration of 2.0 ms were emitted, characterized by a suppressed first harmonic and a median peak-frequency of 36.7 kHz (ranging from 28.3 to 46.9 kHz). The median inter-syllable interval between these type 3 syllables amounted to 21.7 ms and varied widely between 4.4 and 229.9 ms. Contact calls were regularly emitted in call series. In a call series, the intervals between the onsets of calls ranged between 220 and 1070 ms.

The syllables of the presented stimuli were synthesized with AviSoft-SASLab Pro, and combined to calls and call series using BatSound Pro 3.31. The synthetic reference call (Fig. 1b), reflecting median parameters and the shape of natural contact calls, consisted of three different syllable types: the first syllable (type 1) was U-shaped with three harmonics (frequency specifications of the first harmonic: peak = 13.8 kHz, start = 18.5 kHz, end = 16.3 kHz, minimum = 11.0 kHz) and had a duration of 25.3 ms; the second syllable (type 2) and syllables 3–6 (type 3) were downward frequency modulated. The second syllable consisted of harmonics 1–3 (first harmonic: peak = 17.6 kHz, start = 20.0 kHz, end = 15.4 kHz) and had a duration of 3.0 ms. The inter pulse-interval amounted to 11.4 ms between the first and second syllable and to

18.0 ms between the second and third syllable. Type 3 syllables consisted of harmonics 2–4 (second harmonic: peak = 35.9 kHz, start = 40.0 kHz, end = 33.7 kHz) and had a duration of 1.7 ms; they were repeated with an inter pulse-interval of 20.0 ms.

Based on this reference call, we generated eight different stimuli differing in frequency (F) and/or rhythm on level of call series (C) and/or of single calls (S). Each stimulus started with a reference call followed by a sequence of identical test calls (test sequence). As test sequence, we presented either three calls identical to the reference call or we varied the number of test calls, the number of type 3 syllables, and/or the frequency of the calls to create stimuli differing in one, two or three structural parameters (Table 1). Changes in rhythm of call series were achieved by omitting the second of the three test calls (missing call C), rhythmical changes within a call by removing syllables four and five in all calls of the test sequence (missing syllables S). Frequency changes comprised all harmonics of a test call and amounted to an upwards fundamental peak-frequency shift of 3 kHz (frequency shift F) in the respective test sequences. Differences between the stimuli were thus in the range of the naturally occurring variability. The durations of each call and of the total call sequence were kept constant for all stimuli. The interval between the onsets of calls was 530 ms between the reference call and the test sequence, and 371, or 742 ms, between test calls.

Experimental design

Stimulus presentation, controlled by a PC, was triggered by the experimenter. A custom made supervising software (J. Pillat) was used to transfer the stimuli, generated with a 10-fold time expansion, via a soundcard (Creative Sound Blaster AudioPCI 64 V) to a portable ultrasound signal processor (PUSP, Ultra Sound Advice), which compressed the signal by a factor of ten. The signal was monitored by an oscilloscope, amplified (harman/kardon, HK620) and fed into a speaker (Technics EAS 10 TH 800C). Stimuli were regularly checked using a ¼-inch microphone (Brüel and Kjør, Type 4135, and pre-amplifier Type 2670) connected to a measuring amplifier (Brüel and Kjør, Type 2610). Maximum stimulus intensity at the bats' head position was 75 ± 2 dB SPL (peak-to-peak).

Experimental procedure

Experiments took place in a weakly illuminated, echo-attenuated chamber ($3.5 \times 2.2 \times 2.3$ m). Animals were trained in a two-alternative, forced choice procedure to wait for stimulus presentation at a starting perch fixed on one side of the chamber, 220 cm in front of an ultrasonic

Table 1 List of stimuli

Stimulus	Changes in			Decision index (x_r/x_u) for model								
	F	C	S	FCS	FC	FS	CS	F	C	S	2FC	F2C
0				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
s			+	0.33	0.00	0.50	0.50	0.00	0.00	1.00	0.00	0.00
c		+		0.33	0.50	0.00	0.50	0.00	1.00	0.00	0.33	0.67
cs		+	+	0.67	0.50	0.50	1.00	0.00	1.00	1.00	0.33	0.67
f	+			0.33	0.50	0.50	0.00	1.00	0.00	0.00	0.67	0.33
fs	+		+	0.67	0.50	1.00	0.50	1.00	0.00	1.00	0.67	0.33
fc	+	+		0.67	1.00	0.50	0.50	1.00	1.00	0.00	1.00	1.00
fcs	+	+	+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Stimuli differed from stimulus 0 by a frequency shift (f), missing calls (c) and/or missing syllables (s) in the test sequence. Changes are indicated by +. The decision index (x_r/x_u), with x_r = number of parameters directing to the right and x_u = total number of parameters used by the respective model, is given for each stimulus and all models assuming the use of a frequency shift (F), missing calls (C) or missing syllables (S), or different parameter combinations, for classification. Please note that while all models predict the same decision behaviour for the training stimuli, they differ in their predictions for the test stimuli

loudspeaker directed towards the bat's head. The perch was positioned 190 cm, the loudspeaker 70 cm above the ground. Two feeding dishes were mounted symmetrically, 15 cm above the ground, at a distance of 75 cm on either side of the loudspeaker. The experimenter was seated behind the loudspeaker and feeders to control the experimental procedure and data storage.

To indicate its decision after stimulus presentation, the bat had to fly directly to one of the feeders. Between stimulus presentation and the bat's decision, the bat made a pronounced body rotation while still hanging on the starting perch, turning from the loudspeaker (Fig. 2a) towards one of the feeders (Fig. 2b); this initial orientation was kept towards this, or rejected by a rotation towards the other, feeder when the animal flew off (Fig. 2c). Bats were trained with the two stimuli differing in three structural parameters (Table 1), with training stimulus 0 rewarded at the left feeder and training stimulus fcs at the right feeder, i.e. a given parameter constellation had to be associated with each feeder. Training stimuli were presented in a pseudo-random order, with no more than four consecutive trials to one side. A session consisted of a maximum of 40 trials. When animals met a criterion of over 75% correct decisions for at least 10 consecutive sessions with the training stimuli, test sessions started in which the test stimuli were interspersed at random. 10–20% of all stimuli presented within a session were test stimuli. To obtain a spontaneous classification and to avoid training effects, responses to test stimuli were rewarded at both feeders.

Data analysis

Responses to training stimuli were summed up for a particular test session. Test sessions in which a bat performed

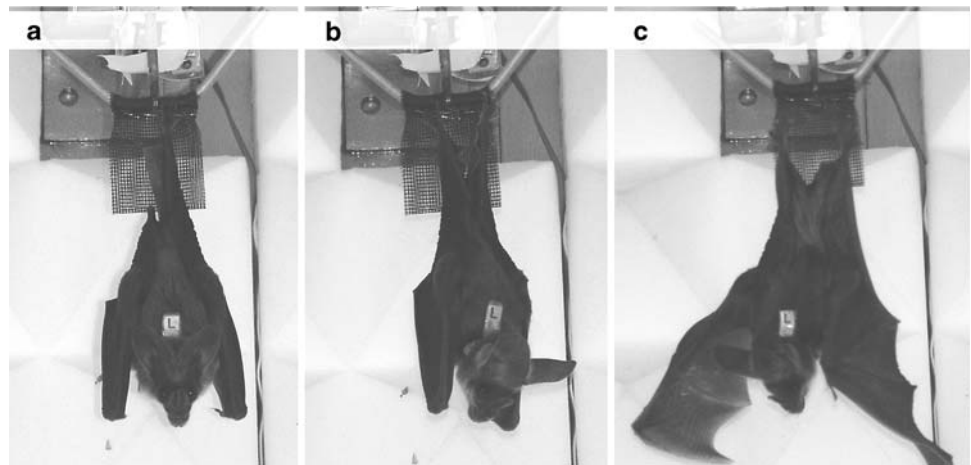
below 75% correct to training stimuli were discarded. Across all valid test sessions, we analyzed the decisions of a bat, as indicated by a flight to one of the feeders, as well as a change of decision, as indicated by an alteration between the initial orientation of the bat towards a feeder after stimulus presentation and its final decision. The parameter change of decision was visually scored by the experimenter and only analyzed for trials in which the bat's body and head were pointing directly towards the loudspeaker in the moment of stimulus presentation (Fig. 2a). For each test stimulus, at least 30 trials were performed. In a two-alternative, forced choice experiment, 75% choices to one side are significantly different from random choice for $n \geq 30$ ($P < 0.01$; Koller 1967). χ^2 tests were used to test whether choices for stimuli differed among each other, and whether the changes of decision for test stimuli differed from that for training stimulus 0 (Statistica, version 6.1, StatSoft, Inc. 2004).

Models

The models aim to reveal the relevant parameters governing the classification of test stimuli by the bats. As the stimuli differed in up to three structural parameters, the bats may have based their classification on different decision criteria: on frequency (F) only, on missing calls (C) only, on missing syllables (S) only or on a combination of two or all three structural parameters. A combination would result in ambiguous information for some stimuli, i.e. at least one cue guided the bat to the left while another one guided it to the right.

We calculated the expected decisions to the right for each stimulus (\exp_{stimulus}) for nine different decision models as

Fig. 2 A bat at the starting perch orienting towards the loudspeaker (a), towards the left feeder (b) and when flying off to the right feeder (c), as seen from the perspective of the experimenter



$$\text{exp}_{\text{stimulus}} = \frac{x_r}{x_u} \times n,$$

with x_r number of parameters directing to the right (indicated by + in Table 1), x_u total number of parameters used by the respective model, n number of trials for the respective stimulus in a bat. Three models assumed that a single parameter was used (models F, C and S), five that a combination of two parameters was used (models FC, FS, CS, 2FC and F2C). Among these, two models predicted a double weighting of one parameter (2FC, F2C). One model assumed a use of all three parameters (FCS). The decision index (x_r/x_u) is shown in Table 1. The decisions predicted by the models were then compared to the bats' decisions. The model quality was calculated as standardized Euclidean distance:

$$\sqrt{\sum_{i_{\text{stim}}=1}^{n_{\text{stim}}} \left(\frac{\text{obs}_{\text{stimulus}} - \text{exp}_{\text{stimulus}}}{n} \right)^2 \div n_{\text{stim}}}$$

with n_{stim} number of stimuli, $\text{obs}_{\text{stimulus}}$ observed number of flights to the right for a given stimulus, $\text{exp}_{\text{stimulus}}$ number of flights to the right as predicted by the model, n number of trials for the respective stimulus in a bat.

Results

Performance for the training stimuli amounted to about 90% correct decisions in both bats. The interspersed test stimuli were spontaneously classified by the bats. Thus, bats were able to transfer the experimental procedure to new stimuli.

Figure 3 shows the classification performance of the two bats. Bat wL classified stimuli s, c and cs significantly to the left, stimulus fc significantly to the right. Decisions for stimuli fs and f were not different from chance performance. Bat wH classified stimulus s significantly to the

left, and stimuli fc, fs and f significantly to the right. Decisions for stimuli cs and c were not different from chance performance. For both individuals, there were no statistical differences in decisions to stimulus pairs fcs–fc, fs–f, cs–c, and s–0 ($\chi^2 P > 0.05$). Thus, all stimuli that differed only in number of syllables were classified equally, i.e. rhythm on level of calls was not used for contact call classification.

Stimuli without frequency shift were predominantly classified to the left, stimuli with frequency shift to the right. Stimuli c, cs, f and fs were not homogeneously classified, whereas wH did not classify c and cs, wL did not classify f and fs differently from chance, hence, we can monitor individual strategies.

The results revealed a change of decision to training stimuli 0 and fcs of 3.5 and 26.6% in bat wL, and of 12.6 and 15.9% in bat wH, respectively. For stimuli classified by chance, changes of decision increased dramatically to nearly 60% for stimuli f and fs in bat wL and to about 40% for stimuli c and cs for bat wH. In fact, changes of decision increased significantly for test stimuli c, cs, f and fs when compared with training stimulus 0 in both bats ($\chi^2 P < 0.05$). This change between the initial orientation reaction during stimulus presentation and the bats' final decision may reflect the conflicting parameter constellation in these four test stimuli compared to the combinations learned during training.

Euclidean distances between bats' decisions and models are given in Fig. 4. Models predicting that only missing calls or missing syllables were relevant for classification (C, S) differed most from animals' decisions, whereas the model assuming a classification by frequency (F) ranked third best. In general, models assuming frequency as a relevant parameter predicted the performance of the bats better than those assuming that frequency was irrelevant for classification. Thus, frequency was the most relevant parameter for contact call classification. Rhythm on level

Fig. 3 Percentage of decisions to the left and to the right for bats wL (a) and wH (c). The 100%-bars indicate how the decisions were distributed, and values denoting the percentage of choices to the right are given. Percentage of change of decision for bats wL (b) and wH (d) for the different stimuli. *n* gives the number of trials taken into account per stimulus and animal

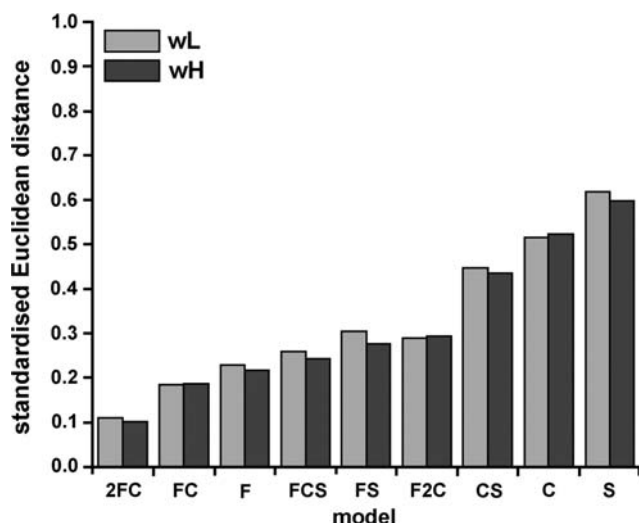
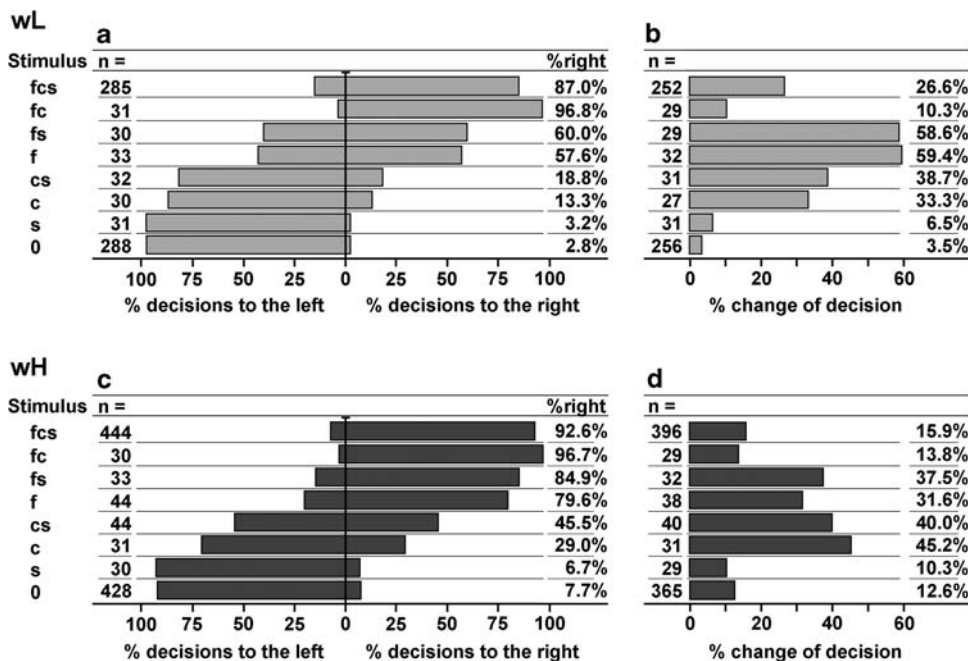


Fig. 4 Standardized Euclidean distances between model predictions and the performance of bats wL and wH, presented in order of minimal distance. Models assumed a classification based on frequency (F), missing calls (C), missing syllables (S), or combinations with an equal or double weighting. A distance of one indicates that the bats decide contrary to the predictions for all parameters, whereas a perfectly fitting model would result in a distance of zero

of call series (missing call C) was a second relevant parameter. Models assuming a combined usage of these two parameters corresponded best to bats' classifications. Minimal Euclidean distance, and therefore the best fit between animals' performance and models, was obtained for the 2FC-model, combining frequency and missing call with a double weighting of frequency.

Discussion

The variation introduced in the synthetic stimuli of the present study reflected the naturally occurring range of contact calls and can in this respect be considered as biologically relevant. Taking into account psychoacoustic data, the presented stimuli differed sufficiently in their parameters to be safely discriminated by *M. lyra* (see e.g. Schmidt 2000; Wiegrebe and Schmidt 1996). Yet the results of the present study supported our hypotheses only partially: the bats classified synthetic contact calls based on a combination of frequency and rhythm on level of call series, with a stronger weighting of frequency. Rhythm on level of calls, however, was not taken into account by the animals in the present paradigm. First, we shall consider how the bats solved the experimental task and to which extent they perceived and evaluated the above parameters taking the observed changes of decision as an indicator, and discuss why the bats may have failed to use rhythm on level of calls in this experiment. Then, we address the question whether the two parameters used in combination for classification were evaluated independently or as a perceptual unit. Finally, we shall discuss the results with respect to the perception of paralinguistic prosodic cues for the communication of identity and affect across different mammalian orders.

Decision behaviour and parameter use

It may appear surprising that the decision behaviour of both bats is described best by the same model, although we

monitored individual strategies. However, we would like to point out here that the models address exclusively the effects of the stimulus parameters on the expected decisions whereas the percentage of decisions to the left, or right, feeder, may reflect both, stimulus specific, i.e. parameter-based, as well as paradigm-dependent factors. In the present paradigm, the bats were trained to associate a given training stimulus of fixed parameter constellation with a specific feeder. This training task can be solved either by acquiring an internal representation of both training stimuli and associating it with the respective feeder, or by referring to one of the training stimuli, only. In the latter case, the learned training stimulus will be associated with one feeder while all stimuli that can be discriminated from this training stimulus by the animal, i.e. are not mapped on the same internal representation, will be associated with the other feeder. The graded decision performance of both bats for the test stimuli renders it likely that the bats have indeed acquired internal representations of both training stimuli, and used them to assess the test stimuli. A different weighting of the two training stimulus representations in the decision process may account for the individual differences in performance. Indeed, an integrative use and an individual-specific weighting of multiple internal representations of training stimuli for the classification of new stimuli has been shown for this bat species, before (Krumbholz and Schmidt 1999).

For all test stimuli, then, the presented parameter constellation was ambiguous compared to the internal representations of both training stimuli, as at least one parameter guided the animal to the left while the others pointed to the right, or vice versa. A change of decision can be interpreted as an indicator of a conflict. If the respective parameter constellation was perceived by the bat, we expected an increase of changes of decision. Substantially, increased change of decision rates were measured for all test stimuli combining frequency and rhythm on level of call series in an ambiguous way, which indicates that the bats perceived these two parameters and evaluated them as reflected in the classification performance. In contrast, for test stimuli differing from the training stimuli in rhythm on level of calls only, the change of decision rate was similarly low as for the training stimuli. This suggests that the rhythm on level of calls was either not perceived or it was evaluated as irrelevant by the bats in the present experiment. Consequently, all stimuli differing only in rhythm on level of calls were classified similarly by the bats.

From a psychoacoustic perspective, this result is remarkable. An experiment with *M. lyra* using series of tone pips presented with different repetition rates revealed a time window of about 152 ms in which the pips interfered (Wiegrebe and Schmidt 1996). All syllables of the synthetic contact calls with their overall duration of about

126 ms occurred within this time window, and the different number of type 3 syllables resulting in a change in rhythm on level of calls should therefore be detectable. In addition, it is likely that the bats were able to detect the time gap differences of about 40 ms in the stimuli differing in rhythm on level of calls, as these differences exceed typical gap detection thresholds for mammals by an order of magnitude. The gap detection threshold for broadband signals determined for a bat, *Tadarida brasiliensis*, amounted to about 2 ms (Nitsche 1993), which is comparable to that in humans (Zwicker and Fastl 1990) and other mammals (cf. Fay 1988), as is typical for a number of psychoacoustic time constants in bats (Schmidt 2000). Yet, the exact number and timing of type 3 syllables within the contact calls had no behavioural relevance in the present experiment. This may have been due to the fact that the individual contact call, rather than its single syllables, was perceived as an acoustical object, a notion supported by a spatial echo-suppression experiment with contact calls in *M. lyra* (Schuchmann 2006).

While the bats did not classify the present stimuli—based on median contact calls in duration and composition—by rhythm on level of calls, this does not rule out the possibility that the bats may be able to classify calls by using this parameter. Rhythm on level of calls may be relevant for a classification of calls composed of a much higher number of syllables, and consequently with an increased call duration, which may occur in communication calls emitted during high arousal situations (cf. Bastian and Schmidt 2008).

Are parameters evaluated independently or as a perceptual unit?

Our results suggest that the remaining two parameters, frequency and rhythm on level of call series, were used in combination for contact call classification. This may reflect a perceptual adaptation to the fact that arousal changes are expressed by a physiologically determined combination of parameters. Among other aspects of voice quality, and amplitude cues, rhythm and frequency cues in vocalizations are affected by arousal-based changes in respiration and in overall muscle tonus (Scherer 1986). An increase in arousal is regularly accompanied by a simultaneous increase of rhythm and frequency parameters in human speech (e.g. Banse and Scherer 1996; Barrett and Paus 2002; Johnstone and Scherer 2000; Ververidis and Kotropoulos 2006; Williams and Stevens 1972). A similar correlation of rhythm and frequency parameters with arousal has been reported across mammals (e.g. Manser 2001; Monticelli et al. 2004; Rendall 2003; Schehka et al. 2007; Weary and Fraser 1995) including *M. lyra* (Bastian and Schmidt 2008). Correspondingly, it may be expected

that a coupled increase, or decrease, of rhythm and frequency constitutes a perceptual unit necessary to evoke a behavioural response.

On the other hand, a perceptual separation of frequency and rhythm parameters may be essential for the recognition of individuals. Although temporal structures may contribute to encoding identity in mammals (e.g. Gelfand and McCracken 1986; Searby and Jouventin 2003) and a particular voice is identified by a specific combination of frequency and temporal cues in humans (e.g. Anward 2002; Lavner et al. 2000; Scherer 1974), caller identity is predominantly determined by frequency parameters resulting from filtering effects in the vocal tract (e.g. Fant 1960; Fitch 2000; Owren et al. 1997). In *M. lyra*, the frequency parameters of the response call, a social call consisting of multiharmonic components with distinct frequency contours such as the contact call, contributed substantially to individual distinctiveness (Bastian and Schmidt 2008). In contact calls, frequency parameters accounted for most of the inter-individual variability (Doerrie 2001; Doerrie et al. 2001).

By using a combination of high frequency with slow rhythm and low frequency with fast rhythm, respectively, as training stimuli, which is contrary to physiological arousal coding as discussed above, we tried to enforce an evaluation based on all structural differences of the stimuli rather than on arousal. In fact, a change of either frequency or rhythm on level of call series was sufficient to cause a different classification by the bats in the present study. This shows that these parameters were perceived independently and not as a perceptual unit, although they were used in combination for classification. The stronger weighting of frequency by the bats in this experiment, however, may reflect the particular relevance of frequency for encoding both, identity and affect in contact calls.

Comparative aspects of prosodic cue perception

Within a given call type, non-human mammals have been shown to discriminate between (e.g. Blumstein and Daniel 2004; Charrier et al. 2002, 2003; Esser and Lud 1997; Hare 1998; Reby et al. 2001), and recognize, individuals (e.g. Balcombe 1990; Kaplan et al. 1978; Searby and Jouventin 2003; Shizawa et al. 2005; Sousa-Lima et al. 2002; Sproul et al. 2006) and group membership (e.g. Boughman and Wilkinson 1998; Frommolt et al. 2003). Moreover, evidence has been provided that they are able to adapt their behaviour in accordance to the affect of the sender (e.g. Blumstein and Armitage 1997; Geiss and Schrader 1996; Manser et al. 2001). However, it is difficult to isolate the cues relevant for the perception of identity and/or affect in natural calls due to the variability within a given social call type which may comprise simultaneous changes in several parameters (Fitch and Kelley 2000). Experiments using

synthetic or natural stimuli with well-defined changes have revealed that the receivers changed their behaviour correlated to frequency (e.g. Fichtel and Hammerschmidt 2002, 2003), time (e.g. Blumstein and Armitage 1997; Randall and Rogovin 2002), or a combination of frequency and time parameters (e.g. Fischer et al. 2001; Geiss and Schrader 1996; Weary et al. 1996), which indicates a communication via paralinguistic prosodic cues in different species.

In bats, a first study with synthetic distress calls which varied in parameters typically associated with differences in affect revealed that an upward frequency shift in combination with an increase of syllable number resulted in the highest attractiveness for conspecific listeners (Russ et al. 2005). However, the paradigm used by Russ et al. measuring overall bat activity by sonar calls is unsuitable to assess the evaluation of stimuli by a given individual. Therefore, it remained unclear whether the measured activity reflected the fact that the number of bats attracted by a given stimulus differed, or that a given individual responded differently to stimuli carrying different affect cues. In contrast, we studied how a given individual evaluated frequency and rhythm parameters, and analyzed the relative importance of these cues for classification.

Conclusion

Our results substantiate that *M. lyra* is able to perceive and evaluate synthetic replicas of its social calls on the basis of parameters that also transmit paralinguistic prosodic information in human language. These necessary prerequisites for a communication of identity and affect via paralinguistic prosodic cues have thus developed far before human language.

Acknowledgments The behavioural experiments in this study did not require an animal experimentation approval. Permission to keep and breed *M. lyra* in fulfilment of Sec. 11 Abs. 1, S. 1, No. 1 Tiersch G has been given to the Institute fuer Zoologie der Stiftung Tierärztliche Hochschule Hannover by the Ordnungsamt, Landeshauptstadt Hannover, Gewerbe- und Veterinaerabteilung, dated 24 March 2003. We would like to thank the Department of Wildlife Conservation, Colombo, Sri Lanka for granting the export license for the bats and our cooperation partner Dr. W. B. Yapa, University of Colombo, Sri Lanka, for logistic support. We thank Dr. J. Pillat for programming the stimulus presentation software, B. Seidl and M. Doerrie for analysing the natural contact calls and S. Schierbaum for generating the synthetic syllables, which we combined for the present stimuli.

References

- Anward J (2002) Other voices, other sources. In: Linell P, Aronsson K (eds) Jagen och rösterna: Goffman, Viveka och samtalet/Selves

- and voices: Goffman, Viveka, and dialogue. University of Linköping, Linköping, pp 127–148
- Balcombe JP (1990) Vocal recognition of pups by mother Mexican free-tailed bats, *Tadarida brasiliensis mexicana*. *Anim Behav* 39:960–966
- Banase R, Scherer KR (1996) Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol* 70(3):614–636
- Barclay RMR, Thomas DW (1979) Copulation call of *Myotis lucifugus*: a discrete situation-specific communication signal. *J Mammal* 60(3):632–634
- Barclay RMR, Fenton MB, Thomas DW (1979) Social behavior of the little brown bat, *Myotis lucifugus*. II. Vocal communication. *Behav Ecol Sociobiol* 6:137–146
- Barrett J, Paus T (2002) Affect-induced changes in speech production. *Exp Brain Res* 146:531–537
- Bastian A, Schmidt S (2008) Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J Acoust Soc Am* 124(1):598–608
- Behr O, von Helversen O (2004) Bat serenades—complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behav Ecol Sociobiol* 56:106–115
- Belin P, Fecteau S, Bédard C (2004) Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci* 8(3):129–136
- Blumstein DT, Armitage KB (1997) Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Anim Behav* 53:143–171
- Blumstein DT, Daniel JC (2004) Yellow-bellied marmots discriminate between the alarm calls of individuals and are more responsive to calls from juveniles. *Anim Behav* 68:1257–1265
- Bohn KM, Wilkinson GS, Moss CF (2007) Discrimination of infant isolation calls by female greater spear-nosed bats, *Phyllostomus hastatus*. *Anim Behav* 73:423–432
- Boughman JW (1997) Greater spear-nosed bats give group-distinctive calls. *Behav Ecol Sociobiol* 40:61–70
- Boughman JW, Wilkinson GS (1998) Greater spear-nosed bats discriminate group mates by vocalizations. *Anim Behav* 55:1717–1732
- Brown PE (1976) Vocal communication in the pallid bat, *Antrozous pallidus*. *Z Tierpsychol* 41(1):34–54
- Bußmann H (1990) Lexikon der Sprachwissenschaft. Alfred Kröner Verlag, Stuttgart
- Charrier I, Mathevon N, Jouventin P (2002) How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. *J Exp Biol* 205:603–612
- Charrier I, Mathevon N, Jouventin P (2003) Vocal signature recognition of mothers by fur seal pups. *Anim Behav* 65:543–550
- Clement MJ, Gupta P, Dietz N, Kanwal JS (2006) Audiovocal communication and social behavior in mustached bats. In: Kanwal JS, Ehret G (eds) Behavior and neurodynamics for auditory communication. Cambridge University Press, Cambridge, pp 57–84
- Davidson SM, Wilkinson GS (2004) Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Anim Behav* 67:883–891
- Doerrie M (2001) Kontaktrufe von *Megaderma lyra* aus Südindien und Sri Lanka: Ein Populationsvergleich. Diploma Thesis, Tierärztliche Hochschule, Hannover
- Doerrie M, Schmidt S, Suba M, Sripathi K (2001) Contact calls of the bat, *Megaderma lyra*: a comparison between an Indian and a Sri Lankan Population. *Zoology* 104(Suppl IV):5
- Eizirik E, Murphy WJ, O'Brien SJ (2001) Molecular dating and biogeography of the early placental mammal radiation. *J Hered* 92:212–219
- Esser KH, Lud B (1997) Discrimination of sinusoidally frequency-modulated sound signals mimicking species-specific communication calls in the FM-bat *Phyllostomus discolor*. *J Comp Physiol A* 180:513–522
- Esser KH, Schmidt U (1989) Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae)—evidence for acoustic learning. *Ethology* 82:156–168
- Fant G (1960) Acoustic theory of speech production with calculations based on X-ray studies of Russian articulations. Description and analysis of contemporary standard Russian. Mouton, The Hague
- Fay RR (1988) Hearing in vertebrates: a psychophysics databook. Hill and Fay, Winnetka
- Fichtel C, Hammerschmidt K (2002) Responses of redfronted lemurs to experimentally modified alarm calls: evidence for urgency-based changes in call structure. *Ethology* 108:763–777
- Fichtel C, Hammerschmidt K (2003) Responses of squirrel monkeys to their experimentally modified mobbing calls. *J Acoust Soc Am* 113(5):2927–2932
- Fischer J, Hammerschmidt K, Cheney DL, Seyfarth RM (2001) Acoustic features of female chacma baboon barks. *Ethology* 107:33–54
- Fitch WT (2000) The evolution of speech: a comparative review. *Trends Cogn Sci* 4(7):258–267
- Fitch WT, Hauser MD (2002) Unpacking “Honesty”: vertebrate vocal production and the evolution of acoustic signals. In: Simmons AM, Fay RR, Popper AN (eds) Acoustic communication. Springer handbook of auditory research. Springer, New York, pp 65–137
- Fitch WT, Kelley JP (2000) Perception of vocal tract resonances by whooping cranes *Grus americana*. *Ethology* 106:559–574
- Frommolt KH, Goltsman ME, Macdonald DW (2003) Barking foxes, *Alopex lagopus*: field experiments in individual recognition in a territorial mammal. *Anim Behav* 65:509–518
- Geiss S, Schrader L (1996) Temporal and structural features of infant calls in relation to caregiving behaviour in common marmosets, *Callithrix j. jacchus*. *Behav Process* 38:183–191
- Gelfand DL, McCracken GF (1986) Individual variation in the isolation calls of Mexican free-tailed bat pups (*Tadarida brasiliensis mexicana*). *Anim Behav* 34:1078–1086
- Goymann W, Leippert D, Hofer H (1999) Parturition, parental behaviour, and pup development in Indian false vampire bats, *Megaderma lyra*. *Z Säugetierkunde* 64:321–331
- Hare JF (1998) Juvenil Richardson’s ground squirrels, *Spermophilus richardsonii*, discriminate among individual alarm callers. *Anim Behav* 55:451–460
- Johnstone T, Scherer KR (2000) Vocal communication of emotion. In: Lewis M, Haviland J (eds) The handbook of emotion. Guilford, New York, pp 220–235
- Kaplan JN, Winship-Ball A, Sim L (1978) Maternal discrimination of infant vocalizations in squirrel monkeys. *Primates* 19(1):187–193
- Koller S (1967) Graphische Tafeln zur Beurteilung statistischer Zahlen. Steinkopff, Darmstadt
- Koopman KF (1994) Chiroptera: vol 1: systematik/systematics. Walter de Gruyter, Berlin
- Krumbholz K, Schmidt S (1999) Perception of complex tones and its analogy to echo spectral analysis in the bat, *Megaderma lyra*. *J Acoust Soc Am* 105(2):898–911
- Kulzer E (2005) Chiroptera: vol 3: biologie/biology. Walter de Gruyter, Berlin
- Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. *Nature* 392:917–920
- Lakshminarayanan K, Ben Shalom D, van Wassenhove V, Orbelo D, Houde J, Poeppel D (2003) The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain Lang* 84:250–263
- Lavner Y, Gath I, Rosenhouse J (2000) The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels. *Speech Commun* 30:9–26

- Leippert D (1994) Social behaviour on the wing in the false vampire, *Megaderma lyra*. *Ethology* 98:111–127
- Leippert D, Goymann W, Hofer H, Marimuthu G, Balasingh J (2000) Roost-mate communication in adult Indian false vampire bats (*Megaderma lyra*): an indication of individuality in temporal and spectral pattern. *Anim Cogn* 3:99–106
- Manser MB (2001) The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proc Biol Sci* 268:2315–2324
- Manser MB, Bell MB, Fletcher LB (2001) The information that receivers extract from alarm calls in suricates. *Proc Biol Sci* 268:2485–2491
- Monticelli PF, Tokumaru RS, Andes C (2004) Isolation induced changes in Guinea Pig *Cavia porcellus* pup distress whistles. *Ann Acad Bras Cienc* 76(2):368–372
- Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ (2001) Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614–616
- Neuweiler G (2000) *The biology of bats*. Oxford University Press, New York
- Nitsche V (1993) Gap detection in the bat, *Tadarida brasiliensis*. In: Elsner N, Heisenberg M (eds) *Proceedings of 21st Göttingen Neurobiol conference*. Georg Thieme, Stuttgart, p 261
- Owren MJ, Seyfarth RM, Cheney DL (1997) The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): implications for production processes and functions. *J Acoust Soc Am* 101(5):2951–2963
- Pfalzer G, Kusch J (2003) Structure and variability of bat social calls: implications for specificity and individual recognition. *J Zool Lond* 261:21–33
- Ramus F, Hauser MD, Miller C, Morris D, Mehler J (2000) Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288:349–351
- Randall JA, Rogovin KA (2002) Variation in and meaning of alarm calls in a social desert rodent *Rhombomys opimus*. *Ethology* 108:513–527
- Reby D, Hewison M, Izquierd M, Pépin D (2001) Red deer (*Cervus elaphus*) hinds discriminate between the roars of their current harem-holder stag and those of neighbouring stags. *Ethology* 107:951–959
- Rendall D (2003) Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J Acoust Soc Am* 113(6):3390–3402
- Rickheit G, Herrmann T, Deutsch W (2003) *Psycholinguistik. Ein internationales Handbuch./Psycholinguistic. An international handbook*. Walter de Gruyter, Berlin
- Russ JM, Jones G, Racey PA (2005) Responses of soprano pipistrelles, *Pipistrellus pygmaeus*, to their experimentally modified distress calls. *Anim Behav* 70:397–404
- Schehka S, Esser KH, Zimmermann E (2007) Acoustical expression of arousal in conflict situations in tree shrews (*Tupaia belangeri*). *J Comp Physiol A* 193:845–852
- Scherer KR (1974) Voice quality analysis of American and German speakers. *J Psycholinguist Res* 3(3):281–298
- Scherer KR (1986) Vocal affect expression: a review and a model for future research. *Psychol Bull* 99(2):143–165
- Scherer KR (1989) Vocal correlates of emotional arousal and affective disturbance. In: Wagner H, Manstead A (eds) *Handbook of social psychophysiology*. John Ltd, pp 165–197
- Scherrer JA, Wilkinson GS (1993) Evening bat isolation calls provide evidence for heritable signatures. *Anim Behav* 46:847–860
- Schmidt S (2000) Echolocation behaviour and comparative psychoacoustics in bats. In: Manley GA, Oeckinghaus H, Fastl H, Klump G, Kössl M (eds) *Auditory worlds: sensory analysis and perception in animals and man*. Wiley, Weinheim, pp 226–245
- Schmidt S (2005) Understanding wildlife: behavioural ecology and acoustic communication in a Sri Lankan bat, *Megaderma lyra*. *Sci Sri Lanka* 6:17–19
- Schuchmann M (2006) Spatial echo suppression and echo-acoustic object normalization in echolocating bats. Dissertation, Ludwig-Maximilian-Universität, München
- Searby A, Jouventin P (2003) Mother-lamb acoustic recognition in sheep: a frequency coding. *Proc Biol Sci* 270:1765–1771
- Shizawa Y, Nakamichi M, Hinobayashi T, Minami T (2005) Playback experiment to test maternal responses of Japanese macaques (*Macaca fuscata*) to their own infant's call when the infants were four to six months old. *Behav Process* 68:41–46
- Sousa-Lima RS, Paglia AP, da Fonseca GAB (2002) Signature information and individual recognition in the isolation calls of Amazonian manatees, *Trichechus inunguis* (Mammalia: Sirenia). *Anim Behav* 63:301–310
- Sproul C, Palleroni A, Hauser MD (2006) Cottontop tamarin, *Saguinus oedipus*, alarm calls contain sufficient information for recognition of individual identity. *Anim Behav* 72:1379–1385
- Tincoff R, Hauser MD, Tsao F, Spaepen G, Ramus F, Mehler J (2005) The role of speech rhythm in language discrimination: further tests with a non-human primate. *Dev Sci* 8(1):26–35
- Toro JM, Trobalon JB, Sebastián-Gallés N (2003) The use of prosodic cues in language discrimination tasks by rats. *Anim Cogn* 6:131–136
- Ververidis D, Kotropoulos C (2006) Emotional speech recognition: resources, features, and methods. *Speech Commun* 48:1162–1181
- Weary DM, Fraser D (1995) Calling by domestic piglets: reliable signals of need? *Anim Behav* 50:1047–1055
- Weary DM, Lawson GL, Thompson BK (1996) Sows show stronger responses to isolation calls of piglets associated with greater levels of piglet need. *Anim Behav* 52:1247–1253
- Wiegrebbe L, Schmidt S (1996) Integration of pulse trains in the bat, *Megaderma lyra*. In: Elsner N, Schnitzler HU (eds) *Proceedings of the 24th Göttingen Neurobiology conference 1996*. Georg Thieme, Stuttgart, p 234
- Wilkinson GS, Boughman JW (1998) Social calls coordinate foraging in greater spear-nosed bats. *Anim Behav* 55(2):337–350
- Williams CE, Stevens KN (1972) Emotions and speech: some acoustical correlates. *J Acoust Soc Am* 52(4B):1238–1250
- Zwicker E, Fastl H (1990) *Psychoacoustics: facts and models*. Springer, Berlin