Modeling aspects of speech processing in bats—behavioral and neurophysiological studies

Karl-Heinz Esser [*]

Neuroethology of Sound Communication Laboratory (NSCL), Department of Neurobiology, University of Ulm, Albert-Einstein-Allee 11, 89069 Ulm, Germany

Abstract

The goal of this article is to provide an integrated view of what we can learn about speech processing from animal studies. 'Integrated' refers here to the attempt to explore biologically important sounds from the level of vocal learning, to neural representation at various stations of the auditory pathway, to audio-motor integration and its influence on sound production. The methodologies required include developmental and behavioral studies as well as electrophysiological characterization of neuronal activity and electrical stimulation of (pre-)frontal/premotor areas. Three unique animal models, otherwise found scattered in the published literature, are brought together here: (1) the lesser spear-nosed bat (*Phyllostomus discolor*) as an animal model for audio-vocal learning; (2) the mustached bat (*Pteronotus parnellii*) as an animal for modeling syntax processing at the auditory cortical single-unit level; (3) the short-tailed fruit bat (*Carollia perspicillata*) as an animal model elucidating cortical audio-motor integration.
© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Audio-vocal learning; Vocal learning; Acoustic communication; Communication calls; Hand-rearing; Syntax processing; Auditory cortex; Audio-motor integration; Frontal cortex; Prefrontal cortex; Electrophysiology; Echolocation

1. Introduction

This report intends to give an integrated view of what we can learn about speech processing from animal studies. 'Integrated' refers here to the attempt to explore biologically important sounds from the level of vocal learning, to neural representation at various stations of the auditory pathway, to audio-motor integration and its influence on sound production (Fig. 1). Although the scheme in Fig. 1 shows the silhouette of the brain of a bat, it is virtually identical to what is known as the "speech chain" in humans (Denes and Pinson, 1993). In the human case, the behavioral level corresponds to the production of speech sounds (i.e., biologically important sounds) either by another person or the listener himself. Arriving at the auditory periphery, speech sound waves are converted into action potentials that travel along the eigth nerve to the central auditory system. With respect to speech perception and production, the ascending auditory pathway has two major functions. In one case, it enables us to process messages transmitted by other speakers. It can also be used for continuously comparing the quality of sounds we produce with the sound qualities we intended to produce (Denes and Pinson, 1993). Such audio-motor integration takes place when an

[*] Tel.: +49-731-50-22635/22631; fax: +49-731-50-22629.
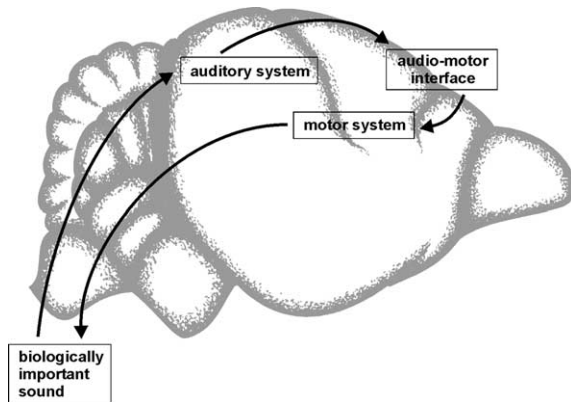   *E-mail address:* kalle.esser@biologie.uni-ulm.de (K.-H. Esser).

Fig. 1. A general framework for thinking about behavior guided by auditory information. Processing of sounds in the auditory pathways of the brain is the prerequisite for mammals, including humans, to perceive, analyse, and identify biologically important sounds such as conspecific vocalizations. Vocal utterances are released by activation of the motor system of the brain. The diagram shows how sound processing and output generation may be coupled via the audio-motor interface in the frontal cortex (Eiermann and Esser, 2000 and below [4]).

individual controls his own sound production by auditory feedback, but also when a heard signal (e.g. conspecific vocalizations) elicits a behavioral reaction in the receiver (Esser and Walkowiak, 2001). In humans, an area in the frontal lobe in the dominant hemisphere (i.e., Broca's area and associated regions) is involved in both sentence comprehension and articulation (Dronkers et al., 2000). Thus, this area is commonly regarded as a key area in cortical audio-motor integration. Activation of such prefrontal/premotor areas prepares a motor output via the motor system of the brain. In the above speaker–listener situation, the motor output includes ventilation of the lungs and movements of the vocal cords, the tongue, and the lips which all are jointly involved in the production of speech. Modern brain imaging techniques allow the exploration of sensory (e.g. auditory cortical) and (pre-)motor (e.g. frontal cortical) representations of speech at the level of large neuronal populations, networks, and systems in living, active brains (Saper et al., 2000). However, we still rely on single-unit animal studies to gain insight into details of brain activation in the processing of complex, species-specific vocalizations.

In the following, I introduce three unique bat models that contributed to our understanding of how biologically significant acoustic information is processed centrally and how it is coupled to motor behavior. First, a model for audio-vocal learning (i.e., the modification of vocal output by auditory experience) provided by the lesser-spear nosed bat (*Phyllostomus discolor*). Second, an animal for modeling syntax processing at the auditory cortical level (i.e., the mustached bat, *Pteronotus parnellii*). Third, an animal model elucidating cortical audio-motor integration as evidenced in the short-tailed fruit bat (*Carollia perspicillata*).

2. Audio-vocal learning

A corner-stone of both human evolution and ontogenetic development is the emergence of speech. To understand how vocal learning is influenced by the human infant's linguistic environment, speech scientists study auditory perception rather than manipulate the baby's acoustic environment on a long-term scale. In passerine (or song) birds, altering the amount and quality of song exposure has produced major insights into innate perceptual abilities and selective vocal learning (for review see Nottebohm, 1999; Doupe and Kuhl, 1999). In contrast to songbirds, the ontogeny of vocalization in non-human mammals is thought to be largely under genetic control (Marler and Mitani, 1988). Does this mean that we have no animal models in which to study mammalian audio-vocal learning? Fortunately, we do have a number of such models including humpback whales (*Megaptera novaeangliae*; Payne and Payne, 1985), bottle-nosed dolphins (*Tursiops truncatus*; Caldwell and Caldwell, 1972; Richards et al., 1984; McCowan and Reiss, 1997), and spear-nosed bats (*Phyllostomus hastatus*; Boughman, 1998; *Phyllostomus discolor*; Esser and Schmidt, 1989; Esser, 1994; for work on primates see Snowdon and Hausberger, 1997). Among the two species of phyllostomid bats, *P. discolor* has been studied in most detail (e.g. hand-rearing of pups in the presence of a call tutor; Esser, 1994) and will be reviewed here.

Juvenile *P. discolor* emit isolation calls when detached from their mothers' nipples and in response to the maternal directive call (Esser and Schmidt, 1989). Maternal directive calls are characterized by an individually distinct, sinusoidal frequency modulation pattern. During ontogeny, pups progressively adapt the structure of their isolation calls (Fig. 2) to their mother's vocal signature. Between about 50 and 100 days of age, infant calls become virtually indistinguishable from the particular maternal reference signal. Apart from a limited *intra*-individual and a comparatively high *inter*-individual variation of major maternal call features (e.g. modulation frequency, number of modulation cycles per call), the ability of the bats' auditory system to resolve those differences in communication call structure can be regarded as a prerequisite for audio-vocal learning in this species (Esser, 1998). Conclusive evidence for this ability, that is, to acoustically discriminate between individuals, was obtained in series of carefully designed psychoacoustic experiments
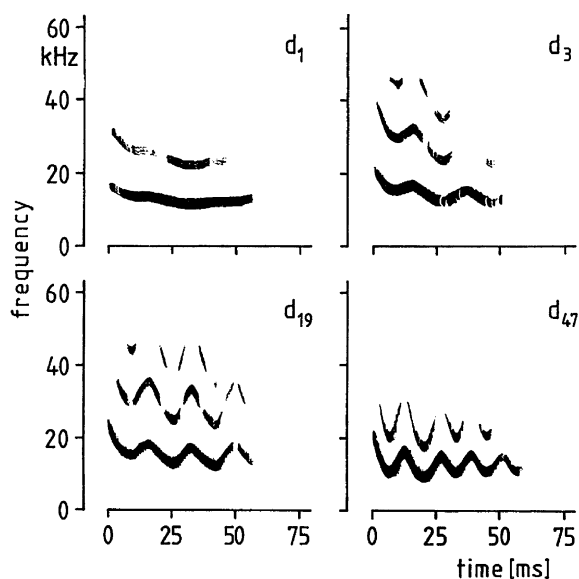

Fig. 2. Development of the sinusoidal frequency-modulation pattern in calls of pups of *P. discolor*. $d_1 - d_{47}$: sonagrams of isolation calls recorded from "pup X" at the respective ages. Note that, as early as by day 47, isolation calls were statistically indistinguishable from its own mother's directive calls (not shown) in terms of both number of frequency minima and maxima per call and modulation frequency. (From: Esser and Schmidt, 1989.)

(e.g. Esser and Kiefer, 1996; Esser and Lud, 1997; Esser, 1998). Typically, hand-rearing (or foster rearing) of animals is regarded as *the* method of choice to distinguish between innate and learned behaviors (e.g. song development in birds; Marler, 1999). Hence, to unequivocally demonstrate audio-vocal learning in *P. discolor*, two groups of bats were hand-reared in social isolation. One group (i.e., the control) was raised in the complete absence of conspecific vocalizations, whereas a digitally-stored maternal directive call was interactively presented to the pups of a second, experimental group prior to each feeding (six feeding sessions per day; 250 play-backs per session). The bats in the play-back group adapted their isolation call structure to the external acoustic reference signal (Fig. 3B). Conversely, the unstimulated control pups developed a pattern independent of that of the mother's call (Fig. 3A; Esser, 1994). On the basis of these studies, *P. discolor*, a bat which can be easily bred in the laboratory, offers unique possibilities for future studies on mammalian audio-vocal learning. Such studies might include the influence of real-time auditory feedback on vocalization (compare Müller et al., 2000) or experience-induced plasticity in both the auditory system and the motor-control system for vocalization.

3. Syntax processing by auditory cortical neurons

Human communication is governed by sequencing rules—speech sounds cannot occur in any order, but only in specified orders. The same is true of words. We will call all such sequencing constraints by the general term syntax. However, rule systems for the sequencing of species–specific vocalizations have also been found in animals, such as birds (Balaban, 1988; Marler and Peters, 1988), non-human primates (Tembrock, 1977; Cleveland and Snowdon, 1982) and the mustached bat, *Pteronotus parnellii* (Kanwal et al., 1994). Therefore, I will take a broader view by defining syntax as any system of rules that allows one to predict the sequencing of communication signals (Snowdon, 1982). The representation of syntactic structure in the human brain cannot be readily
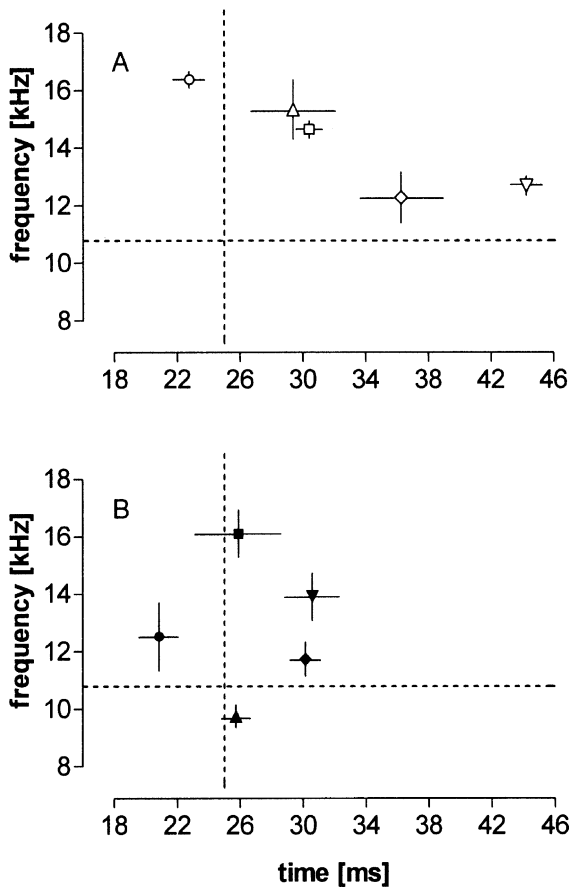
Fig. 3. Comparison of the time intervals ($\bar{x} \pm s$, (ms), x-axis) between the first and the third frequency minimum of isolation calls of 50-day-old pups (control (top) vs. play-back group (bottom)). This time interval (t) was chosen as a reciprocal measure of modulation frequency ($f_{mod} = 2/t$) since shared by the calls of all animals studied. Individual means (n = 10 animals) are represented by different symbols. In addition, the frequency values of the third minimum are plotted on the y-axis ($\bar{x} \pm s$, (kHz)). The corresponding values (ms, kHz) of the invariable, computer-generated reference signal are indicated by the dashed lines. The adaptation of isolation calls to the temporal structure of the play-back signal is restricted to the group of acoustically stimulated animals (bottom graph). (Redrawn from: Esser, 1994.)

studied at the single-unit level. The mustached bat, *P. parnellii*, is a promising animal for studies on syntax processing at the single-neuron level of the auditory cortex for two reasons (Esser et al., 1997). First, the auditory cortex of this species is probably the best understood of all mammals (O'Neill, 1995). Second, mustached bats frequently com-

bine, according to syntactical rules, otherwise independently emitted simple syllables to higher-order vocal constructs, so-called composites (Kanwal et al., 1994). These rules are reflected by the strong constraints on the use of simple syllables as components of heterosyllabic composites. For example, this bat could, in theory, produce 342 disyllabic composites. However less than 15 have been found so far (Kanwal et al., 1994). Below, I will focus on time-domain processing of composite communication calls (Esser et al., 1997) in a previously defined area (the frequency-modulated (FM)–FM area) of the mustached bat auditory cortex (O'Neill and Suga, 1979).

A sequence of 10 digitally-stored natural composites (Fig. 4) was used in these experiments. All calls were played back at a rate of 1/s and at sound pressure levels of 50, 60, 70, 80, and 90 dB. Initially, a "best composite" was determined for each neuron by summing the number of elicited action potentials over the five stimulus levels. All further tests were performed by using the best composite at its best amplitude. As shown in Fig. 4, each composite (2–11; number 1 was reserved for a no-stimulus control) consisted of two parts or "syllables" (connecting points of syllables indicated by arrows in Fig. 4). As a major measurement, we determined the neuron's response ratio in percent. In accordance with previous studies (e.g. Fitzpatrick et al., 1993), "facilitation" was indicated if the response to the entire composite was >120% of the sum of responses to the heterosyllabic parts. Conversely, "suppression" was defined as a response to the original <80% of the sum of responses to the individual syllables.

More than 50% of the 107 neurons studied where highly sensitive to spectrotemporal features emerging *de novo* from the combination of syllables to composite communication calls as indicated by response ratios of 122–1108% (facilitation) and 35–77% (suppression). For example, the neuron shown in Fig. 5B–D responded vigorously with facilitation (Fig. 5B; response ratio = 304%) to the composite dRFM-cDFM (Fig. 5A), but responded poorly to the individual syllable components (Fig. 5C and D). In contrast, in another neuron (Fig. 5F–H), the initial syllable (sHFM) of composite sHFM-fSFM (Fig. 5E) inhibited the
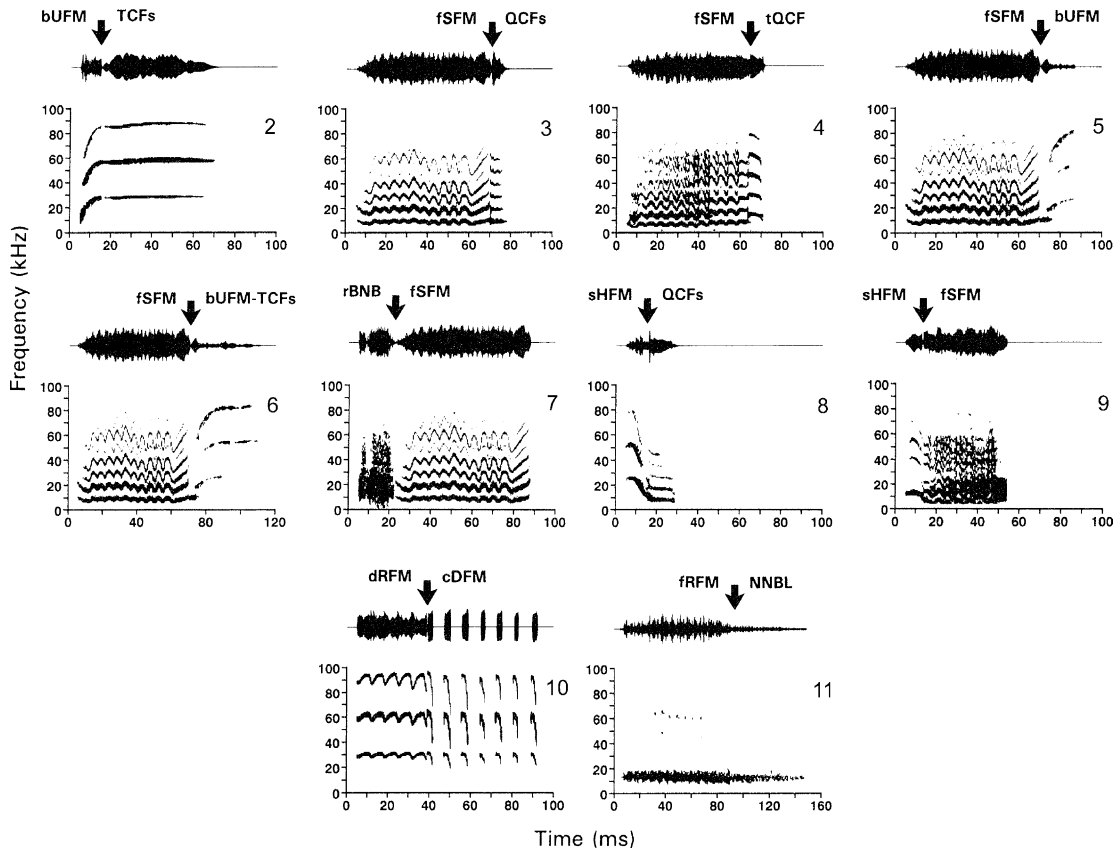
Fig. 4. Oscillograms (top) and sonagrams (bottom) of composite communication calls (nos. 2–11) of the mustached bat, *P. parnellii* (Kanwal et al., 1994). Composites are made up of two (all except no. 6) or three (no. 6) distinct components (syllables) that the bats combine without an intervening silent interval (connecting points of syllables indicated by arrows). The composites (e.g. bUFM-TCFs, no. 2) are named according to the simple syllables from which they are composed (for complete nomenclature see Kanwal et al., 1994). (From: Esser et al., 1997.)

spike response to the following syllable (fSFM, Fig. 5F) by 56%, to which the unit responded robustly when presented alone (Fig. 5H). The importance of the temporal structure of the composite calls in eliciting neuronal responses became most evident when playing the whole stimulus in reverse, introducing a silent period between the syllables, or reversing the order of syllables within a composite (Esser et al., 1997). Except for two of 21 neurons studied, responses to the reversed composite call were always reduced as compared with responses to the corresponding original composite. For instance, the neuron shown in Fig. 6 demonstrated temporal facilitation of its off-response (Fig. 6B) to the composite fSFM-bUFM (Fig. 6A), but almost failed to respond to the in-

dividual components when presented alone (Fig. 6C and D) or even to the entire composite call if played in reverse (Fig. 6E). Moreover, the other time-domain manipulations of the composite structure mentioned above, typically resulted in the loss of a facilitated response or even in the complete loss of a response. For example, introducing a silent period $\geqslant 0.5$ ms between both syllables of the composite fSFM-bUFM (Fig. 6A) resulted in a progressive decay or in a total loss of the response (Fig. 6F). Similarly, reversing the order of syllables resulted in the loss of a facilitated response (not shown).

The observations from single neurons reported here, provide multiple lines of evidence that syntax in communication calls is processed by neurons
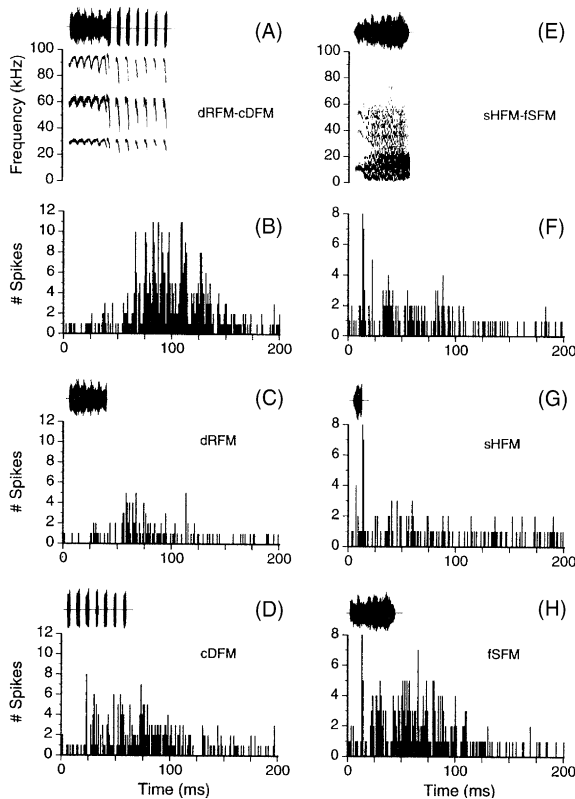
Fig. 5. Left column (A–D) temporal facilitation (response ratio = 304%): (A) Oscillogram and sonagram of composite no. 10 (dRFM-cDFM), the best composite for unit no. 1. (B) Peristimulus-time histogram (PSTH; bin width = 1 ms) shows unit's strong response to the original composite (dRFM-cDFM). (C) PSTH shows unit's weak response to the first syllable (dRFM). (D) PSTH shows unit's weak response to the second syllable (cDFM). Right column (E–H) temporal suppression (response ratio = 44%). (E) Oscillogram and sonagram of composite no. 9 (sHFM-fSFM), the best composite for unit no. 25. (F) PSTH shows unit's response to original composite (sHFM-fSFM). (G) PSTH shows unit's response to the first syllable (sHFM). (H) PSTH shows unit's strongest response was to the second syllable (fSFM). (From: Esser et al., 1997.)
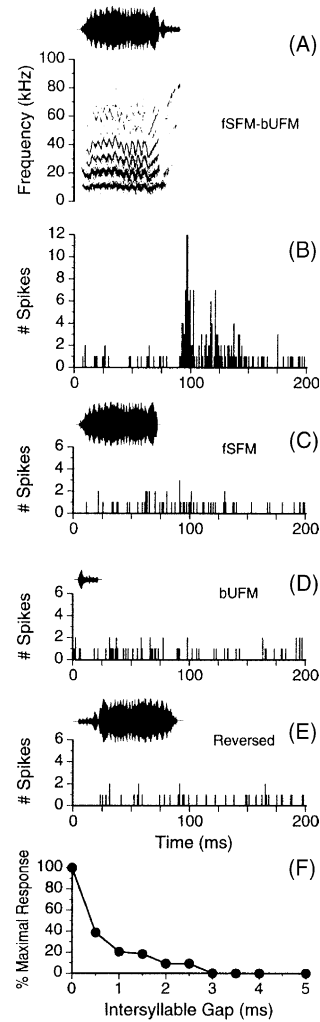

Fig. 6. (A) Oscillogram and sonagram of composite no. 5 (fSFM-bUFM), the best composite for unit no. 102. (B) PSTH shows unit's robust response to the original composite (fSFM-bUFM; response ratio = 1108%). (C) Unit almost failed to respond to the first syllable (fSFM). (D) Unit nearly failed to respond to the second syllable (bUFM). (E) No response to playback of the reversed composite. (F) Unit's response decreased dramatically and eventually ceased when a silent period ⩾ 0.5 ms was inserted between the two components of composite fSFM-bUFM. (From: Esser et al., 1997.)

(and hence represented) in the mammalian non-primary auditory cortex. The emergence of such combination-sensitive neurons at the auditory cortical level, tuned to particular syllable combinations in the time domain is in accordance with previous findings in this and other species. The auditory cortex generally seems to serve as a substrate for complex temporal processing, including temporally extended processing of brief acoustic

signals (Esser et al., 1997). The question whether or not combination-sensitive neurons are also involved in the perception of some parameters of human speech, such as syntax, cannot be answered with present day technologies in humans. The

discussed studies on bats, however, provide answers that arguably reflect general coding strategies that might also be true for humans.

4. Cortical audio-motor integration

Based on neuroanatomical findings in *P. parnellii* (Kobler et al., 1987; Casseday et al., 1989), it was hypothesized that an area in the bat frontal cortex is part of an audio-motor feedback loop. Although convergence of auditory inputs and motor outputs is essential for both echolocation and interindividual acoustic communication, combined studies of sound-evoked neuronal responses and premotor functions in the auditory-related region of the bat frontal cortex were missing. In the short-tailed fruit bat, *Carollia perspicillata*, we first localized a frontal auditory field (frontal auditory field (FAF); Eiermann and Esser, 2000; Fig. 7) by using both retrograde tracing and metabolic brain mapping techniques. We then studied the response properties of neurons from this area and the motor effects of electrical microstimulation at identical electrode

positions. The majority of neurons studied (46%) responded preferentially to paired, downward FM signals (so-called FM–FM combinations) and smaller numbers of neurons preferred tones (25%), single downward FMs (17%), upward FMs (1%), clicks (9%), or band-limited noise (1%). In these experiments, single downward FMs and FM–FM combinations were designed to represent individual harmonics of the species' biosonar call (pulse) and pulse-echo combinations respectively (Fig. 7, bottom left). Further, within each stimulus type, sounds were varied along several physical dimensions (e.g., frequency and sound pressure level) until being most effective for response generation in the neuron under consideration. Below, I will focus on responses of FAF neurons to biologically significant sounds (i.e., stimuli mimicking biosonar pulse-echo combinations).

Responses of neurons preferring FM–FM combinations ($n = 35$ units; PSTH example in Fig. 7) were found to be facilitated without exception. Neurons virtually failed to respond to unpaired downward FMs. Further, in contrast to auditory cortical FM–FM neurons in this (Esser and Eiermann, 2002) and other bat species (O'Neill, 1995),
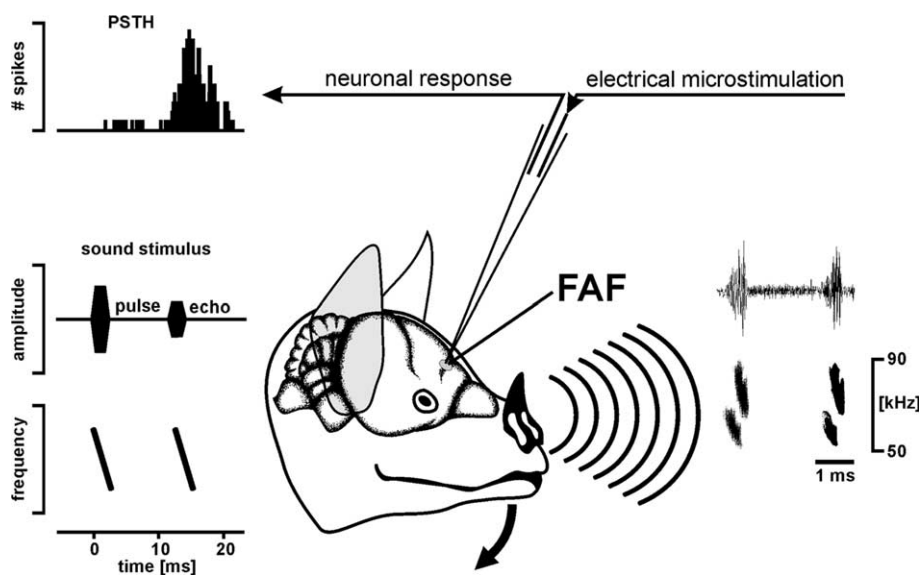

Fig. 7. The bat FAF. FAF neurons respond preferentially to sound stimuli mimicking individual harmonics of biosonar pulse-echo combinations. A PSTH example of such a response is given on top. Electrical microstimulation at identical electrode positions (re the neuronal recordings) leads to motor effects such as jaw movements and vocalizations (see oscillograms and sonagrams of biosonar-like downward FMs; bottom right). Further explanations, see text.

frontal cortical FM–FM neurons were rarely (<10%; 3 of 35 units) tuned to a particular pulse-echo delay that would correspond to a certain target distance in echolocation. The fact that these neurons were strictly combination sensitive for pulse-echo stimuli, but rarely exhibited delay-tuning suggests that FM–FM neurons in the bat frontal cortex are likely to function as novelty detectors by indicating the presence of a sound reflecting object irrespective of its distance. A second probable function of the *Carollia* frontal auditory field can be derived from the FM-sweep preferences observed. A particular sub-type of FM-responsive neurons, corresponding to about one sixth of units studied, preferred FMs starting at frequencies between 130 and 140 kHz, fairly close to the high-frequency roll-off of the species' behavioral audiogram (Bitter et al., 2001). Since the use of higher harmonics is generally thought to improve the resolution of bat echolocation, such neurons are likely to participate in object recognition. Assuming that both novelty detection and object identification are biological functions of the frontal auditory field, this area, in accordance with its hodology (Kobler et al., 1987; Casseday et al., 1989), seems further predestined to trigger changes in the bats' echolocation-related motor behavior. This prediction has been confirmed experimentally in a recent series of microstimulation experiments (Eiermann and Esser, unpublished). Briefly, motor effects arising from the *Carollia* FAF include both movements of the lower jaw (indicated by the curved arrow in Fig. 7) and the nasal emission of biosonar-like calls (Fig. 7, right). Interestingly, both movements of the mandible and vocalizations have been elicited in humans by intra-operative electrical stimulation of Broca's area (Binkofski, 2000 and personal communication). Hence, striking functional analogies between the FAF and Broca's area are suggested. In addition to echolocation pulses, the short-tailed fruit bat also possesses a complex repertoire of communication calls (e.g. Porter, 1979). Based on a digital library of this species' communication calls (Straub and Esser, 2000) and on techniques for a real-time manipulation of vocal-auditory feedback (e.g. Lohmann et al., 2001), our *Carollia* model offers a unique chance to study cortical audio-motor integration with respect to more complex, and therefore more speech-like, vocalizations at the single-neuron level.

5. Conclusions and future directions

Comparative studies have demonstrated great similarities among mammals (including man and bats) in the structure and function of auditory pathways (e.g. ubiquitous presence of all major auditory nuclei, common principles of tonotopic organization; Webster, 1992). One might therefore argue that both simple and complex auditory stimuli are often processed in a similar way, independent of the species under consideration. Even perceptual features that have previously been thought to be speech-specific, such as categorical perception, perceptual constancy despite variability in many acoustic dimensions, perception of formant structure in multi-tone complexes, and phoneme perception, have gained acceptance as general preadaptations for the analysis and recognition of communication sounds in mammals, including humans (e.g. Ehret, 1992; Ehret and Riecke, 2002). Consequently, it seems reasonable to suppose that our understanding of human speech processing can substantially benefit from acquiring knowledge of how animals process conspecific vocalizations at the neuronal level.

As reviewed above, my co-workers and I successfully modelled various aspects of speech processing in "the bat". However, for several reasons, the goal of exploring biologically important sounds from the level of vocal learning, to neural representation at various stations of the auditory pathway, to audio-motor integration and its influence on sound production could not be realized by using a single microchiropteran species. First, with respect to communication calls, the modification of vocal output by auditory experience, so far, has unequivocally been shown only in the genus *Phyllostomus*. Second, as opposed to *P. parnellii*, the syntactical rules for the generation of higher order vocal constructs such as heterosyllabic composites are not yet available for any phyllostomid bat. On the other hand, unlike *P. discolor*, mustached bats do not breed in a labo-

ratory environment and hence are largely un-available for studies on vocal development. Third, neurophysiological studies of the auditory and the motor control systems for vocalization within the same species are not possible without a continuous animal supply. Hence, for the cortical audio-motor integration project, *C. perspicillata* was mainly chosen on the basis of its availability (for details see Esser and Eiermann, 1999).

In the audio-motor integration approach, using replicas of biosonar signals, can be looked at only as a first step. To be able to study cortical processing with respect to more complex and therefore more speech-like vocalizations (compare 4.), we recently established a digital communication-call library for each of the two phyllostomid species bred in this laboratory (*Carollia perspicillata*; Straub and Esser, 2000; *Phyllostomus discolor*; Pistohl and Esser, 1998). Since, in *P. discolor*, in addition to audio-vocal learning and a non-human primate-like repertoire of communication calls (about 20 structurally well-defined call types; Pistohl and Esser, 1998), also a frontal auditory field could be recently identified (Esser, unpublished), this species has the obvious potential to become a premier animal model in the field of mammalian vocal communication.

Acknowledgements

References

Balaban, E., 1988. Cultural and genetic variation in swamp sparrows (*Melospiza georgiana*). II. Behavioral salience of geographic song variants. Behaviour 105, 292–322.

Binkofski, F., 2000. Perceptive and executive functions of Broca's region: evidence from neuroimaging. Eur. J. Neurosci. 12 (Suppl. 11), 189.

Bitter, K.S., Heffner, R.S., Koay, G., Heffner, H.E., 2001. Behavioral audiogram of the short-tailed fruit bat, *Carollia perspicillata*. Assoc. Res. Otolaryngol. Abs. 24, 63–64.

Boughman, J.W., 1998. Vocal learning by greater spear-nosed bats. Proc. R. Soc. Lond. B 265, 227–233.

Caldwell, M.C., Caldwell, D.K., 1972. Vocal mimicry in the whistle mode by an Atlantic bottlenosed dolphin. Cetology 9, 1–8.

Casseday, J.H., Kobler, J.B., Isbey, S.F., Covey, E., 1989. Central acoustic tract in an echolocating bat: an extralemniscal auditory pathway to the thalamus. J. Comp. Neurol. 287, 247–259.

Cleveland, J., Snowdon, C.T., 1982. The complex vocal repertoire of the adult cotton-top tamarin (*Saguinus oedipus oedipus*). Z. Tierpsychol. 58, 231–270.

Denes, P.B., Pinson, E.N., 1993. The speech chain: the physics and biology of spoken language. W.H. Freeman and Company, New York, NY.

Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. 22, 567–631.

Dronkers, N.F., Pinker, S., Damasio, A., 2000. Language and the aphasias. In: Kandel, E.R., Schwartz, J.H., Jessell, T.M. (Eds.), Principles of Neural Science. McGraw-Hill, Maidenherd.

Ehret, G., 1992. Preadaptations in the auditory system of mammals for phoneme perception. In: Schouten, M.E.H. (Ed.),The Auditory Processing of Speech: from Sounds to Words. Mouton de Gruyter.

Ehret, G., Riecke, S., 2002. Mice and humans perceive multiharmonic communication sounds in the same way. Proc. Natl. Acad. Sci. USA 99, 479–482.

Eiermann, A., Esser, K.-H., 2000. Auditory responses from the frontal cortex in the short-tailed fruit bat *Carollia perspicillata*. NeuroReport 11, 421–425.

Esser, K.-H., 1994. Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. NeuroReport 5, 1718–1720.

Esser, K.-H., 1998. Psychoacoustic studies in neotropical bats. In: Nielzén, S., Olsson, O. (Eds.), Clinical Psychoacoustics. Lund University Press.

Esser, K.-H., Eiermann, A., 1999. Tonotopic organization and parcellation of auditory cortex in the FM-bat *Carollia perspicillata*. Eur. J. Neurosci. 11, 3669–3682.

Esser, K.-H., Eiermann, A., 2002. Processing of frequency-modulated sounds in the *Carollia* auditory and frontal cortex. In: Thomas, J., Moss, C., Vater, M. (Eds.), Echolocation in Bats and Dolphins. University of Chicago Press, Chicago.

Esser, K.-H., Kiefer, R., 1996. Detection of frequency modulation in the FM-bat *Phyllostomus discolor*. J. Comp. Physiol. A 178, 787–796.

Esser, K.-H., 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, K.-H., Schmidt, U., 1989. Mother–infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae)—evidence for acoustic learning. Ethology 82, 156–168.

Esser, K.-H., Walkowiak, W., 2001. Audiomotor integration in vertebrates. In: Elsner, N., Kreutzberg, G.W. (Eds.), The Neurosciences at the Turn of the Century. Proceedings of the 4th Meeting of the German Neuroscience Society, Vol. I. Thieme.

Esser, K.-H., Condon, C.J., Suga, N., Kanwal, J.S., 1997. Syntax processing by auditory cortical neurons in the FM–FM area of the mustached bat *Pteronotus parnellii*. Proc. Natl. Acad. Sci. USA 94, 14019–14024.

Fitzpatrick, D.C., Kanwal, J.S., Butman, J.A., Suga, N., 1993. Combination-sensitive neurons in the primary auditory cortex of the mustached bat. J. Neurosci. 13, 931–940.

Kanwal, J.S., Matsumura, S., Ohlemiller, K., Suga, N., 1994. Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. J. Acoust. Soc. Amer. 96, 1229–1254.

Kobler, J.B., Isbey, S.F., Casseday, J.H., 1987. Auditory pathways to the frontal cortex of the mustache bat, *Pteronotus parnellii*. Science 236, 824–826.

Lohmann, P., Müsseler, J., Esser, K.-H., 2001. Vocal effects of delayed auditory feedback in the lesser spear-nosed bat *Phyllostomus discolor*. In: Abstracts of the 6th International Congress of Neuroethology, Bonn, Germany.

Marler, P., 1999. On innateness: are sparrow songs "learned" or "innate"? In: Hauser, M.D., Konishi, M. (Eds.), The Design of Animal Communication, A Bradford Book.

Marler, P., Mitani, J., 1988. Vocal communication in primates and birds: parallels and contrasts. In: Todt, D., Goedeking, P., Symmes, D. (Eds.), Primate Vocal Communication. Springer, Berlin.

Marler, P., Peters, S., 1988. The role of song phonology and syntax in vocal learning preferences in the song sparrow, *Melospiza melodia*. Ethology 77, 125–149.

McCowan, B., Reiss, D., 1997. Vocal learning in captive bottlenose dolphins: a comparison with humans and non-human animals. In: Snowdon, C.T., Hausberger, M. (Eds.), Social Influences on Vocal Development. Cambridge University Press, Cambridge.

Müller, K., Aschersleben, G., Esser, K.-H., Müsseler, J., 2000. Effects of delayed auditory feedback on speech: just a problem of displaced rhythm? In: Desain, P., Windsor, L. (Eds.), Rhythm Perception and Production, Swets & Zeitlinger, Lisse.

Nottebohm, F., 1999. The anatomy and timing of vocal learning in birds. In: Hauser, M.D., Konishi, M. (Eds.), The Design of Animal Communication. A Bradford Book.

O'Neill, W.E., 1995. The bat auditory cortex. In: Popper, A.N., Fay, R.R. (Eds.), Hearing by Bats. Springer, Berlin.

O'Neill, W.E., Suga, N., 1979. Target range-sensitive neurons in the auditory cortex of the mustache bat. Science 203, 69–73.

Payne, K., Payne, R., 1985. Large scale changes over 19 years in songs of humpback whales in Bermuda. Z. Tierpsychol. 68, 89–114.

Pistohl, D., Esser, K.-H., 1998. The vocal repertoire of the leser spear-nosed bat (*Phyllostomus discolor*). Zoology 101 (Suppl. I), 84.

Porter, F.L., 1979. Social behavior in the leaf-nosed bat, *Carollia perspicillata*. II. Social communication. Z. Tierpsychol. 50, 1–8.

Richards, D.G., Wolz, J.P., Herman, L.M., 1984. Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, (*Tursiops truncatus*). J. Comp. Psychol. 98, 10–28.

Saper, C.B., Iversen, S., Frackowiak, R., 2000. Integration of sensory and motor function: the association areas of the cerebral cortex and the cognitive capabilities of the brain. In: Kandel, E.R., Schwartz, J.H., Jessell, T.M. (Eds.), Principles of Neural Science. McGraw-Hill, Maidenherd.

Snowdon, C.T., 1982. Linguistic and psycholinguistic approaches to primate communication. In: Snowdon, C.T., Brown, C.H., Petersen, M.R. (Eds.), Primate Communication. Cambridge University Press, Cambridge.

Snowdon, C.T., Hausberger, M., 1997. Social Influences on Vocal Development. Cambridge University Press, Cambridge.

Straub, O., Esser, K.-H., 2000. The vocal repertoire of the short-tailed fruit bat (*Carollia perspicillata*)—generating template sounds for playback experiments. Zoology 103 (Suppl. III), 36.

Tembrock, G., 1977. Tierstimmenforschung. Wittenberg Lutherstadt, A. Ziemsen Verlag.

Webster, D.B., 1992. An overview of mammalian auditory pathways with an emphasis on humans. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), The Mammalian Auditory Pathway: Neuroanatomy. Springer, Berlin.