Spatiotemporal contrast enhancement and feature extraction in the bat auditory midbrain and cortex

Susanne Hoffmann,1 Alexander Warmbold,1 Lutz Wiegrebe,1 and Uwe Firzlaff1,2
1Division of Neurobiology, Department of Biology II, Ludwig-Maximilians-University Munich, Planegg-Martinsried, Germany; and 2Technische Universität München, Lehrstuhl für Zoologie, Freising-Weihenstephan, Germany

Submitted 28 March 2013; accepted in final form 17 June 2013

Hoffmann S, Warmbold A, Wiegrebe L, Firzlaff U. Spatiotemporal contrast enhancement and feature extraction in the bat auditory midbrain and cortex. J Neurophysiol 110: 1257–1268, 2013. First published June 19, 2013; doi:10.1152/jn.00226.2013.—Navigating on the wing in complete darkness is a challenging task for echolocating bats. It requires the detailed analysis of spatial and temporal information gained through echolocation. Thus neural encoding of spatiotemporal echo information is a major function in the bat auditory system. In this study we presented echoes in virtual acoustic space and used a reverse-correlation technique to investigate the spatiotemporal response characteristics of units in the inferior colliculus (IC) and the auditory cortex (AC) of the bat Phyllostomus discolor. Spatiotemporal response maps (STRMs) of IC units revealed an organization of suppressive and excitatory regions that provided pronounced contrast enhancement along both the time and azimuth axes. Most IC units showed either spatially centralized short-latency excitation spatiotemporal imbedded in excitation. This complementary arrangement of excitation and suppression was very rarely seen in AC units. In contrast, STRMs in the AC revealed much less suppression, sharper spatiotemporal tuning, and often a special spatiotemporal arrangement of two excitatory regions. Temporal separation of excitatory regions ranged up to 25 ms and was thus in the range of temporal delays occurring in target ranging in bats in a natural situation. Our data indicate that spatiotemporal processing of echo information in the bat auditory midbrain and cortex serves very different purposes: Whereas the spatiotemporal contrast enhancement provided by the IC contributes to echo-feature extraction, the AC reflects the result of this processing in terms of a high selectivity and task-oriented recombination of the extracted features.

echolocation; microchiroptera; spatial hearing; Phyllostomus discolor

THE ANALYSIS OF SPATIOTEMPORAL acoustic information is crucial for most animals. Numerous studies have addressed this issue in terms of motion-direction sensitivity, auditory looming perception, or auditory scene analysis (e.g., Ahissar et al. 1992; Bregman 1990; Ghazanfar et al. 2002).

Spatiotemporal auditory processing is also especially important for echolocating bats, navigating at high speed through a dark, densely structured environment. The auditory extraction of both temporal and spatial characteristics of echoes referenced to the emitted calls is essential for detecting prey and avoiding obstacles. This sensory challenge has guided a number of studies, which have revealed precise temporal processing of echo delay and stimulus duration as the hallmark of the bat auditory system (Ehrlich et al. 1997; Mittmann and Wenstrup 1995; Olsen and Suga 1991; O’Neill and Suga 1979). In addition, there is strong evidence for the high spatial directionality of the bat auditory system, which is mainly mediated by the interaction of the large sound-transformation structures (ears and nose leaves) and the high frequencies used by bats (e.g., Muller 2010; Vanderelst et al. 2010).

Spatial tuning of auditory neurons in bats has been described in terms of sensitivity to interaural intensity differences (IID; e.g., Covey et al. 1991; Park and Pollak 1993; Wenstrup et al. 1986) and/or interaural envelope time differences (ITD; Borina et al. 2011; Lohuis and Fuzessery 2000). With free-field stimulation, spatial tuning of auditory neurons was measured or reconstructed from binaural properties (e.g., Fuzessery and Pollak 1985; Grinnell and Grinnell 1965; Makous and O’Neill 1986), and could often be directly related to the transfer functions of the head and outer ears (HRTF; Jen et al. 1989; Sun and Jen 1987).

Very few studies have looked at the interdependence of temporal and spatial parameters and how they are reflected in neural responses. It was shown that coding of pulse repetition rate as well as coding of pulse-echo delay is influenced by the spatial position of stimulus presentation (Hoffmann et al. 2010; Suga et al. 1990; Valentine and Moss 1997; Wu and Jen 1996).

The experimental paradigms of these studies, however, do not allow for an unbiased sampling of the very large parameter space opened by the range of call-echo delays and spatial positions that a bat encounters when echolocating in a densely structured environment. In the present study we attempted to overcome these limitations by constructing spatiotemporal response maps (STRMs) of units in the inferior colliculus (IC) and the auditory cortex (AC) of the bat Phyllostomus discolor in response to a long-duration, random spatiotemporal sequence of echoes of the sonar emission of this species. From these sequences and the corresponding neural responses, STRMs were constructed using a reverse-correlation technique. Of specific interest for the current study is how spatiotemporal response properties are refined between the auditory midbrain and cortex. The STRMs recorded in the IC reveal conspicuous spatiotemporal configurations, which cannot be predicted from classical IID tuning curves. The complementary arrangement of excitation and suppression provides pronounced contrast enhancement along both the time and azimuth axes. In AC, most STRMs reveal a very different, complex structure of sharply tuned, spatially and temporally separated excitatory and suppressive regions. This suggests a role of these neurons in the processing of biosonar signals in behaviorally relevant tasks such as target ranging and/or spatial target localization.
MATERIALS AND METHODS

Surgery. All experiments complied with the principles of laboratory animal care and were conducted under the regulations of the current version of the German Law on Animal Protection (approval 209.1/211-2531-68/03 and 55.2-1-54-2531-128-08 Regierungsverwaltung Oberbayern). All bats (P. discolor) originated from a breeding colony in the Department of Biology II of the Ludwig-Maximilians-University in Munich. For experiments, animals were kept separated from other bats under seminatural conditions (12:12-h day-night cycle, 65–70% relative humidity, 28°C) with free access to food and water. The principle surgical procedure has been described in detail elsewhere (Schuller et al. 1986). To alleviate postoperative pain, an analgesic (0.2 mg/kg body wt meloxicam; Metacam; Boehringer-Ingelheim, Ingelheim, Germany) was administered after the surgery. After the surgery, the principle surgical procedure has been described in detail elsewhere (Schuller et al. 1986). To alleviate postoperative pain, an analgesic (0.2 mg/kg body wt meloxicam; Metacam; Boehringer-Ingelheim, Ingelheim, Germany) was administered after the surgery. The bats were anesthetized using a combination of medetomidine (Domitor; Novartis, Mississauga, ON, Canada), midazolam (Dormicum; Hoffmann-La Roche, Mississauga, ON, Canada), and fentanyl (Fentanyl-Janssen; Janssen-Cilag, Neuss, Germany; 0.4, 4, and 0.04 μg/g body wt, respectively). The skin overlying the skull was opened along the midline, and the skull surface was freed from tissue. A small metal tube was fixed to the skull using a microglass composite to fixate the animal to a stereotaxic device, and the accurate skull position in stereotaxic coordinates was determined as described in detail elsewhere (Schuller et al. 1986). To alleviate postoperative pain, an analgesic (0.2 mg/kg body wt meloxicam; Metacam; Boehringer-Ingelheim, Ingelheim, Germany) was administered after the surgery.

Fig. 2. Cut-out of the continuous stimulus sequence used for recordings of spatiotemporal response maps (STRMs). Top: time signal and spectrum for channel 1 (left ear). Bottom: time signal and spectrum for channel 2 (right ear).

This echo sequence was then digital-to-analog (DA) converted (sampling rate 192 kHz; RME-Audio Fireface 400, Synthax, Haimhausen, Germany) and presented binaurally to the bats via custom-made earphones with a flat frequency response ±3 dB between 10 and 100 kHz (Schuller 1997) for an extended period until enough spikes to generate an STRM were recorded (see below), typically between 5 and 10 min. The presentation level of the echo sequence was adjusted so that the root mean square (RMS) of the sequence was 20–40 dB above the threshold at a unit’s best frequency (BF; see below).

Experiments were conducted in a sound-attenuated chamber. Extracellular recordings were made with conventional glass-insulated tungsten electrodes (impedance 1–2 MΩ; Alpha Omega, Ubstadt-Weiher, Germany) in anesthetized bats (medetomidine, midazolam, and fentanyl; 0.4, 4, and 0.04 μg/g body wt, respectively). During recordings, anesthesia was maintained by injecting half of the initial dose of the anesthetic every 2 h. Note that responses recorded from cortical units under this anesthesia regime reflected the behavioral performance of P. discolor well (Firzlaff et al. 2006). Recording sessions could last up to 4 h per day and were repeated 4 days a week. Electrode penetrations in the AC were run obliquely to the brain surface with different mediolateral and rostrocaudal angles. Penetrations in the IC were run perpendicularly to the surface of the IC with different mediolateral and rostrocaudal angles. Because it was not always possible to clearly discriminate the activity of a single neuron, the term “unit” will be generally used in the following to describe the collective activity of one to three neurons recorded at a recording site.

While the continuous stimulus sequence (see above) was presented, spikes were amplified, filtered [RA16, RX5; Tucker Davis Technologies (TDT), Gainesville, FL], and simultaneously recorded as a continuous trace with the same Fireface 400 audio interface [analog-to-digital (AD) rate 192 kHz] used for presentation of the stimulus sequence and were then stored on a personal computer. In addition, the stimulus was re-recorded simultaneously on a separate channel. Thus the temporal relationship of the spikes and the stimulus preceding each spike was exactly preserved for later analysis. To characterize the basic response properties of each unit, frequency-response areas were determined by presenting pure tones of 20-ms duration (0.5-ms rise and fall time). The pure tones were presented with 10 repetitions monaurally to the contralateral ear with randomly changing combinations of frequency and sound pressure level [DA converter: RX6 (TDT); sampling rate 260 kHz; attenuator: PA5 (TDT)] with a repetition rate of 1.3 Hz. Action potentials were recorded using an AD converter (RA16, RX5; sampling rate 25 kHz) and Brainware (TDT). The BF of a unit was then determined as the
frequency where a response could be elicited at the lowest presentation level. In addition, the binaural response properties were determined in a subset of units at the unit’s BF using the averaged binaural intensity method (Irving et al. 1996). Pure tones (duration 20 ms) were presented binaurally with increasing intensity at the contralateral ear and decreasing intensity at the ipsilateral ear, and vice versa. The IID was changed from −40 to +40 dB in steps of 10 dB. The level of the signals was chosen so that at 0 dB IID, both ears were stimulated at 20 dB above the unit’s threshold at BF. Thereafter, the same stimuli were presented monaurally to the contralateral and ipsilateral ear, respectively. Stimuli were presented with 10 repetitions and a repetition rate of 1.3 Hz.

In a subset of units, the static spatial receptive field was mapped by presenting single echolocation calls convolved with the same IRs from 25 azimuth positions as described above for the recording of the STRMs. The single echoes were binaurally presented with 20 repetitions in random order with a repetition rate of 1.3 Hz.

After the completion of an experiment, electrolytic lesions were made to the brain to reconstruct the position of recording sites from subsequent histological processing in standardized stereotaxic coordinates of a brain atlas of P. discolor (Nixdorf A, Fenzl T, Schwellnus B, unpublished observations). Animals were then euthanized by an intraperitoneally applied lethal dose of pentobarbital sodium (0.16 mg/g body wt) and subsequently transperused.

**Construction of STRMs.** For each spike the spatial stimulus positions included in the last 50 ms before spike occurrence were indicated in a 25 × 50 spike matrix by assigning the value 1 to the corresponding space-time bins (Fig. 3, A–C). The mean value for each bin was then calculated from the sum of all spike matrices (Fig. 3D).

To evaluate the level of background activity uncorrelated to the stimulus, random matrices were constructed from randomly drawn 50-ms segments of the stimulus sequence (Fig. 3E). For each neuron the number of randomly drawn segments was equal to the number of spike-triggered segments. In the final STRM, excitatory (positive) and suppressive (negative) deviations from this background activity were considered to be significant on the basis of a pixelwise t-test (Fig. 3F).

Formation of excitatory and suppressive areas was based on a clustering of significant ($P < 0.05$) positive or negative t-values, respectively. Based on these areas, STRMs were classified into five classes (see RESULTS). The spatiotemporal position of selected areas was determined by finding the ellipse that best fitted the spatiotemporal distribution of significant t-values (via 2-dimensional cross-correlation) and determining its centroid (not shown).

Because the construction of STRMs is based on a spike-triggered average, STRMs were generated only for units from which at least 400 spikes could be evoked during the time of stimulus presentation. Spike numbers ranged between 415 and 10,868 (mean: 3,362 ± 2,594) for IC units and between 402 and 7,991 (mean: 1,930 ± 1,456) for AC units.

**Evaluation of STRMs: predictability of pulse-pair responses.** Evaluation of STRMs was done with conventional pulse-pair stimuli. In these experiments spikes were amplified, filtered, and AD converted using TDT equipment (RA16, RX5, sampling rate 25 kHz) and recorded with Brainware (TDT). Pulse pairs were presented with 20 repetitions at a repetition rate of 1.3 Hz.

Each pulse consisted of an echolocation call of P. discolor convolved with the same HRIRs as described above. The spatiotemporal presentation order of combinations of two of these pulses was adjusted accordingly to the spatiotemporal order of excitatory areas in the STRM of Complex units (see below): Each of the two pulses was presented from a spatial position located in the center of each of the two excitatory areas, respectively. Temporal delays separating the two pulses ranged from 1 to 15 ms (1-ms steps), thus including the range of the temporal separation of the two excitatory areas in the STRM under test. Units with areas temporally separated more that 15 ms were not included in the procedure. Afterward, the spatiotemporal presentation order of the two pulses was reversed. It can be expected that pulse-pairs that match the spatiotemporal arrangement of excitatory areas elicit the strongest response. The presentation order of the delay steps was always random. In addition, each of the two pulses was presented alone. In “simple” units (see below), the two pulses were presented from the same spatial position and only the delay was varied.

To evaluate the predictive power of STRMs of “complex” units (see below), we tested if the response elicited by pulse pairs was significantly stronger ($t$-test, $P < 0.05$) than the sum of the response elicited by each pulse alone, a criterion indicating facilitative interactions (e.g., Olsen and Suga 1991).

### RESULTS

Recordings were derived from 104 units in the IC of 3 bats and from 295 units in the AC of 7 bats. In the AC, most units were recorded from layers III to V. Ninety-four percent (98/104) of units in the IC showed robust responses to the echo sequence and displayed significant excitatory or suppressive areas in the STRM, whereas this was found in only 52% (153/295) of cortical units.

Examples of STRMs recorded in the IC are shown in Fig. 4A. The STRMs are clearly classifiable on the basis of the spatiotemporal distribution of excitatory and suppressive areas. Thirty-six of 98 IC units (37%) revealed a short-latency excitatory area followed by or sometimes even spatiotemporally imbedded in a suppressive area and were classified as E/S units (Fig. 4A, first column). In contrast, 28 of 98 units (29%) revealed a complementary pattern, with a short-latency suppressive area followed by or spatiotemporally imbedded in an excitatory area and were classified as S/E units (Fig. 4A, second column).

The AC reveals fundamentally different STRMs; an arrangement of excitation and suppression as found in the IC was encountered only very rarely. Only 5 of 153 units (3%) were classified as E/S, and only 2 units (1%) were classified as S/E. Instead, the largest class of cortical units (58/153, 38%) revealed complex STRMs with two distinct excitatory areas separated in space and time, which could sometimes be separated by an area of suppression (complex units; Fig. 4B, third column). Complex units were only rarely encountered in the IC (9/98, 9%). The second largest class of units in the AC showed only one sharply tuned excitatory area. If suppression was present, it was weak and did not surround the excitation (simple units, 57/153, 37%; Fig. 4B, fourth column). Also, simple units were rarely encountered in the IC (6/98, 6%).

In both the IC and AC, some STRMs were recorded that did not fit into one of the four classes described above. These STRMs showed, for example, more than two excitatory or suppressive areas in various combinations, or they showed no clear spatial tuning at all. To avoid a confusing classification scheme, they were combined and labeled “unclassifiable” (IC: 19/98, 19%; AC: 31/153, 20%).

The possibility exists that the combination of two excitatory areas seen in complex units might be simply due to the combination of two simple units simultaneously recorded in a multiunit cluster. Therefore, the ratio of multiple and single neurons included in the class of complex units was analyzed. Single neurons were defined on the basis of interspike interval histograms. The criterion was that only 2% or fewer of the interspike intervals of a unit were shorter than 1 ms. Forty percent (23/58) of all units in the AC classified as complex were single neurons. An example of a single neuron is shown in Fig. 5, A–D, showing clearly the
pattern of two spatiotemporally separated excitatory areas characteristic of complex-type units in the AC.

To summarize, the principal characteristic of spatiotemporal tuning in the IC was a conspicuous complementary pattern of excitation and suppression, whereas in the AC, spatiotemporally sharp tuning and/or double tuning was found.

Relationship between BF and STRM type. We analyzed the BF of units belonging to the major STRM classes in the AC and the IC. BFs in simple units in the AC and E/S units in the IC were higher (median: 65 kHz in both types; interquartiles: 23 and 23.5 kHz, respectively) than in AC complex units and IC S/E units (median: 61.5 and 58.5 kHz, respectively; interquartiles: 20 and 9 kHz, respectively). A Kruskal-Wallis test (Matlab, with correction for multiple testing) revealed that this difference was statistically significant ($P < 0.05$) between AC simple units and IC S/E units.

Temporal relationship of excitatory areas and arrangement of complex units on the cortical surface. The temporal separation between the two excitatory areas in the STRMs of AC complex units was in the range of 2.0–25.5 ms (median: 9.0 ms). This range covers the delays occurring between pulse

---

Fig. 3. Schematic of the STRM generation procedure. A: electrode signal. B and C: $25 \times 25$ spike matrices with time of occurrence of stimuli relative to spikes (time re spikes; 1-ms bins) indicated on the abscissa and the corresponding stimulus position in azimuth indicated on the ordinate. D: spike matrix calculated from the mean number of spikes in each space-time bin. E: random matrix. F: STRM. Significant pixels ($P < 0.05$) correspond to a $t$-value of $\pm 1.9$. Excitation in the STRM is shown in light levels of gray and suppression in dark levels of gray. Positive azimuth positions belong to the contralateral hemisphere. For further details see MATERIALS AND METHODS.
and echo that a bat would perceive during target approach in a natural situation (e.g., Schnitzler et al. 1987). Hagemann et al. (2010) have shown that in a sibling frequency-modulated (FM) bat species (*Carollia perspicillata*), delay-tuned neurons are chronotopically arranged in the dorsal auditory cortex. Figure 6A shows the location of recording sites in the AC of *P. discolor*. Seventy-nine percent (46/58) of the complex units were located in the dorsal fields of the AC of *P. discolor*, mainly in the posterior dorsal field (PDF). Only 21% (12/58) of the complex units were located in the posterior ventral field (PVF). However, as can be seen in Fig. 6B, even these units were located close to the ventral border of the PDF. Figure 6C shows that the temporal separation (“delay”) between the excitatory areas in AC complex units is chronotopically arranged: the separation increases significantly along the rostrocaudal axis.
The trend seen in the pooled data is also visible in the data from individual bats (Fig. 7, A–F), which show a consistent covariation of rostrocaudal position and delay. Only in one animal (Fig. 7E) is the trend only weak. From one animal (Fig. 7F) data were recorded only from caudal positions; thus no trend can be seen; however, all complex units recorded from this animal exhibit a long delay and thus fit well into the overall trend. Note that because of the low number of data points from individual bats, correlation was not tested for significance. However, although the number of complex units recorded from single bats is sometimes low, the data show a clear trend toward a topographic representation of delay in the dorsal cortex of the individual bats.

**Relationship between STRMs and classical binaural response properties.** For all IC units, a classical binaural characterization based on the responses of BF-tone IIDs was determined. The comparison of the STRM classification with transient broadband echoes and the BF-tone IID classification in IC units is shown in Fig. 8. Figure 8A shows the exemplary STRMs of an IC E/S and an IC S/E unit together with the corresponding IID functions obtained by classical binaural BF-tone stimulation. Although the STRMs reveal fundamentally different excitatory and suppressive areas, the IID functions are very similar to each other. Moreover, there is no correlation between the binaural classification (excitatory/inhibitory, excitatory/excitatory, or excitatory/nonresponsive) and the STRM classification: both E/S and S/E IC units are equally distributed among the binaural classes (Fig. 8B).

In the AC, the comparison of IID functions with the STRM is further complicated by the complex temporal features of the STRMs: the most prominent class of AC units has complex STRMs with two excitatory areas, with the first one typically near midline and the second one typically on the contralateral side. This complex spatiotemporal arrangement of excitatory and often suppressive areas cannot be faithfully captured by a BF-tone IID function. In summary, the description of binaural response properties in terms of BF-tone IID functions is not a good predictor for the spatial tuning of both IC and AC units with dynamical, time-variant stimulation.

**Evaluation of STRMs.** In a subset of 17 simple and 14 complex cortical units, the predictive power of STRMs was tested with isolated pulses or pulse pairs. Exemplary evaluations of the STRM of two cortical units are shown in Fig. 9. The STRM shown in Fig. 9A defines the unit as simple. The spatial receptive field, recorded with isolated pulses, is shown in Fig. 9B. The data show that the spatial tuning is in good agreement. However, the stationary receptive field does not reveal the weak suppression following the single excitatory area in the STRM. The effect of this suppression is revealed by pulse-pair stimulation, where both pulses are presented from the same azimuth but the delay between the pulses is varied from 1 to 15 ms (Fig. 9C). The response to the second pulse is increasingly reduced with decreasing delay. The STRMs of 13 of the 17 simple units (76%) provided qualitatively correct predictions of responses to isolated pulses or pulse pairs. If
areas of suppression were visible in the STRM, temporal suppressive interactions could also be observed with the two-pulse paradigm.

An example of a complex cortical unit is shown in Fig. 9, D–F. The unit has two excitatory areas with latencies of ca. 14 and 9 ms at spatial positions of 0° and +22.5°, respectively. Single pulses from 0° and +22.5° evoke no or only weak responses. However, a combination of the two pulses with a delay of 5–7 ms evokes strong responses (Fig. 9, E and F). The STRMs of 4 of the 14 complex units (29%) provided correct predictions of responses to pulse pairs in that facilitative interaction between the two pulses were observed.

**DISCUSSION**

This study was designed to investigate the spatiotemporal response properties of units in the auditory midbrain and cortex of an echolocating bat. For the first time with bats, a reverse-correlation technique was used to construct STRMs, allowing for unbiased sampling of a large parameter space of spatiotemporal stimulus combinations. Our results revealed a fundamentally new spatiotemporal organization of excitatory and suppressive regions in the STRM of units in the bat IC. The spatiotemporal organization of both E/S and S/E IC units shown in our results could not be predicted from conventional IID functions.

This shows that the auditory system of bats has more complex processing capacities than has been demonstrated so far. Conventional stimuli with simple temporal or spatial properties often used in electrophysiological studies might not always be adequate to uncover these capacities. The use of long sequences of random spatiotemporally distributed stimuli in our study has revealed at least some of these sophisticated processing properties, which are especially important in the context of long sequences of reflected echoes a bat perceives when flying in a densely structured environment.

When the STRMs are compared to results from studies using more conventional stimuli, the stimulation rate used for STRM recordings might be of importance. For example, studies by Wu and Jen (1996) or Hoffmann et al. (2010) showed that stimulus repetition rate or stimulus intervals have a strong influence on spatial tuning of cortical neurons in bats. The stimulus density of 625 s⁻¹ in our study might seem to be very high. However, given the short stimulus duration of 3 ms, the overall stimulation density per second is comparable to the stimulus density used for spectrotemporal receptive field recordings by deCharms et al. (1998). In addition, call rates within so-called strobe groups emitted by P. discolor during flight are up to 90 Hz (Rother and Schmidt 1982). It is thus not unrealistic to expect a high number of echoes to be reflected toward the bat from multiple directions when the bat flies in a densely structured environment. The stimulus rate of the continuous echo sequences in our study might have influenced the neural responses, but at the same time the use of echo sequences might be a step toward a more naturalistic stimulation for probing the auditory system of bats.

In the present study, we only used stimuli that varied in the horizontal plane to record STRMs. However, Jenison et al. (2001) demonstrated that spatiotemporal receptive fields in the AC of cats also had, of course, a vertical component. In a different set of experiments (not included in the present data set), we used virtual echoes that varied in elevation but not in...
the horizontal dimension. These experiments showed that 41% of neurons in the AC of *P. discolor* had STRMs with a response pattern comparable to those described for the complex units in this report (Hoffmann et al. 2013). Whereas the central excitatory area was always located around 0° elevation, the second area of excitation was often below or above 0° elevation. This basically confirms our findings in the present study, i.e., the existence of the complex units in the AC of *P. discolor*, but also makes clear that the number of complex units might be underestimated by using spatial echo stimuli only in the horizontal plane.

**Feature extraction and recombination: the roles of the bat IC and AC.** Our results fit well into current concepts of ascending auditory neural processing (Nelken 2004) where the IC is regarded as the highest stage of auditory feature extraction. The conspicuous organization of suppression and excitation in E/S and S/E units might serve to enhance the spatiotemporal contrast and thus increase the spatiotemporal selectivity of neural responses. The extracted features can then be recombined in the AC in a behaviorally relevant manner.

An impressive example of the concept of feature extraction in the IC and reorganization of response properties in the AC is the neural processing of target range in constant-frequency (CF) bats (for review see Wenstrup and Portfors 2011). In the IC of the mustached bat, delay tuning, as the neural mechanism for target range detection, is created by spectrotemporal integration. The integrative interactions can be inhibitory or facilitative. Delay tuning properties are then reorganized to create ordered maps of pulse-echo delay in the AC.

The cortical complex units described in the current study may serve to encode target range: STRMs of complex units in the AC exhibited more than one excitatory area separated in time and space. We hypothesize that the excitatory area with the longer latency is sensitive to the emitted call and that the excitatory area with the shorter latency is sensitive to the echo. This is supported by the fact that the long-latency excitatory area is in most cases located nearer to the midline (e.g., Fig. 5A), in accordance with that, an emitted call would equally stimulate both ears. Thus a complex unit responds strongly to call-echo combinations that match not only the temporal but also the spatial arrangement of its excitatory fields. Indeed, control recordings with pulse-pair stimulation, as a first approximation of call-echo combinations (e.g., O’Neill and Suga 1982), reveal that the spatiotemporally arranged excitatory areas of some complex units have predictive power for pulse-pair stimuli. The results suggest that pulse-echo delay sensitivity can also have a strong directional component, which might be important for encoding object position or to track objects during complex flight maneuvers while echo reflection angles are changing.

We are fully aware that in “classical” delay-tuned neurons like those found in *Pteronotus* (e.g., O’Neill and Suga 1982), facilitation is evoked by call-echo pairs where the echo is attenuated with respect to the call. These level differences were not included in our paradigms for STRM recordings and pulse-pair stimulation. However, in delay-tuned neurons in the phyllostomid FM-bat *Carollia perspicillata*, facilitation could well be evoked with call-echo pairs presented with the same

---

**Fig. 9. Predictive power of STRMs evaluated by pulse-pair stimulation.** A: STRM of an AC simple unit. The black ellipse marks the excitatory area of the unit, and the cross marks the spatial position from which the pulse pairs are presented. B: raster plot showing the static spatial receptive field of the unit shown in A. Note that in A the time axis is inverted. C: raster plot showing the response to single pulses and pulse pairs. The top row shows the spontaneous activity of the unit. 1st and 2nd, spatial position in azimuth of the first and the second pulses of the pair. Black vertical lines represent the different onsets of presentation of the first pulse and the constant onset of presentation of the second pulse. D: STRM of an AC complex unit. The black ellipses mark the excitatory areas of the unit, and the crosses mark the spatial positions from which the pulses are presented. E: raster plot showing the response to single pulses and pulse pairs for the complex unit shown in D. F: quantification of responses elicited by single pulses and pulse pairs. The filled bar represents the response evoked by both single pulses; asterisks mark responses significantly different from this response.
level (Hagemann et al. 2010), although best facilitation was indeed generally evoked with echo-level differences of about –18 dB with respect to the call. The lack of level differences in our stimulation paradigms thus does not per se hamper the interpretation of STRMs with two excitatory areas as being involved in target ranging.

This interpretation is further supported by the range of temporal separations of excitatory areas: These were between 2 and 25.5 ms, corresponding to a target range between 34 and 434 cm. This is in the same order of delays occurring in target ranging for bats in a natural situation (e.g., Schnitzler et al. 1987) and also in the same range of delays found in delay-tuned neurons in the AC of Carollia (Hagemann et al. 2010).

In C. perspicillata, delay tuned neurons were located in dorsal regions of the AC (Hagemann et al. 2010). In our study, most complex units were also located in a posterior dorsal region of the AC of P. discolor. This region has been described as a distinct cortical field, the PDF, on the basis of neuroanatomic and neurophysiological features (Hoffmann et al. 2008).

In Carollia, Hagemann et al. (2010) showed that in this cortical region, delay-tuned neurons were orderly arranged to form a chronotopic map of target range. For the cortical complex units in our study, the temporal separations between the two excitatory areas were also arranged topographically in the PDF of P. discolor (cf. Fig. 6C). The topographical arrangement closely resembled that of best delays in delay-tuned neurons in Pteronotus parmelli and Carollia in that the temporal separation increased along the rostrocaudal axis (Hagemann et al. 2010; O’Neill and Suga 1982).

If one looks at the cortical organization of delay-tuned neurons in bats, it is conspicuous that two different types exist. In the FM bats Eptesicus and Myotis, delay-tuned neurons are mainly found in the primary auditory cortex and do not form a chronotopic map (Dear et al. 1993; Sullivan 1982). In contrast, a chronotopic map was found in dorsal areas of the AC in the CF bats Rhinolophus and Pteronotus (Schuller et al. 1991; Suga and O’Neill 1979) as well as in the FM bat Carollia (Hagemann et al. 2010). Our data suggest that a similar map also exists in Phyllostomus. There might be several reasons for these organizational differences. The similarity in the cortical representation of delay tuning in Pteronotus, Carollia, and Phyllostomus might simply reflect the close evolutionary relationship of the families Mormoopidae and Phyllostomidae (Jones and Teeling 2006), as already suggested by Hagemann et al. (2010). However, Rhinolophus belongs to the just remotely related Old World family of Rhinolophidae. Therefore, additional functional aspects might also influence the cortical organization of delay-tuned neurons in bats. It has been suggested that both foraging type and sonar call design might influence the necessity of cortical map formation (Altes 1989; O’Neill 1995).

Delay tuning in the IC of FM bats? In contrast to CF bats, where delay-tuned neurons have already been found at the level of the IC (Mittmann and Wenstrup 1995), the IC in FM bats has not been systematically investigated concerning delay tuning. In FM bats, studies have so far reported delay-tuned neurons only at the level of the intercollicularis nucleus (Feng et al. 1978) or intercollicular nucleus (Dear and Suga 1995). In the current study, complex units with two spatiotemporally separated excitatory areas, which might represent delay-tuned neurons, were rare in the IC.

For delay tuning in CF and FM bats, different underlying mechanisms have been described: in CF bats, the combination of different frequency channels is crucial (Wenstrup and Portfors 2011), whereas in FM bats, delay tuning is based on homoharmonic interactions and does not require the combination of inputs from different frequency channels (Feng 2011; Hagemann et al. 2010). In CF bats, complex interactions between different inhibitory circuits have been described (Sanchez et al. 2008; Wenstrup and Portfors 2011), whereas in FM bats, a paradoxical latency shift has been discussed as a possible underlying mechanism (Feng 2011; Ma and Suga 2008). Thus the different mechanisms underlying the creation of delay-sensitive responses of neurons in CF and FM bats might also be reflected in a different neural representation of target range in the IC.

Predictive power of the cortical STRMs. The experimental design in this study was explicitly chosen to understand spatiotemporal coding of echo features in the bat auditory midbrain and cortex. Thus the very general evaluation procedure used with classic spectrotemporal or spatiotemporal receptive fields was not applied. Instead, we chose an evaluation procedure that was tailored to the scientific question, namely, how do units respond to pulse-echo pairs with different spatial separation and temporal separation? Therefore, we used the classical pulse-echo stimulation for evaluation instead of testing the STRMs with long sequences of high-density echoes as for the acquisition of the STRMs.

Our data show that although the stimulation paradigm for the STRM acquisition is quite different from that for the evaluation, the STRMs have overall good predictive power for a unit’s response to pulse-echo pairs. In 76% of cortical simple units and in 29% of cortical complex units, the STRM provided a good estimate for responses to transient acoustic stimulation. Spike timing and its relationship to the spatiotemporal arrangement of echoes for pulse pairs seem to be well captured in the STRM. For complex units, the predictability might have been increased if level differences between the two pulses had been used during stimulation, as required by classical delay-tuned neurons (Hagemann et al. 2010; O’Neill and Suga 1982). For the complex units, the partial success of the pulse-pair evaluation further strengthens the view that these units are tuned to echo delay.

Mechanisms determining STRM features in the IC and AC. STRMs in the IC of P. discolor showed spatial features that are in many respects inconsistent with classical IID functions (cf. Fig. 8). The neuronal basis underlying these unusual features might be similar to what has been described by Li et al. (2010). The authors found that in neurons in the IC of bats, ipsilateral signals evoked subthreshold excitatory postsynaptic potentials (EPSPs), even though binaural signals with the same ipsilateral intensity generated spike suppression. The authors suggested that these ipsilaterally evoked EPSPs could influence the responsiveness of IC neurons to dynamic signals such as moving sound sources or multiple sounds that occur when echolocating bats fly in complex environments. The continuous stimulus sequence used to record STRMs in our study resembles an acoustic scene generated by echoes returning with various temporal intervals from various spatial positions in a densely structured environment. Therefore, it is well possible that the special spatiotemporal response characteristics of E/S or S/E
units in the IC of *P. discolor* in our study are created by mechanisms as described by Li et al. (2010).

In the STRM of simple units in the AC, areas of excitation were generally spatially more focused compared with the excitation in complex units. The same can often be observed by the comparison of E/S and S/E units in the IC. As a trend, both AC simple units and IC E/S units showed higher BF s as responses to pure-tone stimuli. It is most likely that directional properties of the pinna are reflected in STRM properties. Whereas the directionality of hearing in *P. discolor* is quite focused for frequencies above 65 kHz, the directionality patterns are broader and show irregularities in azimuth and elevation for frequencies around 55 kHz (Firzlaff and Schuller 2003). Complex units in the AC and E/S units in the IC had median BFs just in the 55- to 62-kHz range. The influence of pinna directionality on spatial tuning of auditory neurons has also been shown in several other studies (e.g., Jen et al. 1989; Sun and Jen 1987). The temporal characteristics (i.e., response latency and response duration) might also be influenced by the directionality of hearing: because the echolocation calls of *P. discolor* contain much energy in the high-frequency range, E/S units are stimulated with high sound pressure level, which causes short spike latency and secures high temporal precision. At lower frequencies (around 55 kHz), less energy is contained in the echolocation calls. Excitatory areas in S/E units are therefore also temporally less focused, with longer latencies and imprecise spike timing.

However, the overall structure of the different STRM types, e.g., the temporal order of excitation and suppression, must be determined by other factors than mere pinna directionality, probably by combination of input from neurons located “downstream” or mechanisms as discussed above.  

**Influence of anesthesia.** All experiments in this study were conducted with anesthetized bats. A possible influence of the anesthetics on results presented in this report will be briefly discussed.

The anesthetic used in this study was a combination of medetomidine, midazolam, and fentanyl (see MATERIALS AND METHODS). Medetomidine did not produce an observable effect on neural activity in the IC and the primary auditory cortex of the Mongolian gerbil (Ter-Mikaelian et al. 2007). Immunoreactivity for endogenous opioids such as enkephalin is generally low or absent in the primary auditory cortex and the medial geniculate body, but abundant in other nuclei of the auditory pathway (Aguilar et al. 2004; Robertson and Mulders 2000). The main effect of enkephalin seems to be inhibition of the cochlear neural output via the descending olivocochlear bundle (Burki et al. 1993). Thus the opioid fentanyl used in our study might have had a nonspecific overall inhibitory effect on neural activity to acoustic stimulation.  

Midazolam, like other benzodiazepines, enhances GABA<sub>A</sub>-mediated inhibition and might be the main source of possible anesthesia effects in this study. Park and Pollak (1994) showed that azimuthal receptive fields in the IC of bats were influenced by GABAergic inhibition, i.e., blocking inhibition often decreased azimuthal sensitivity. Temporal processing was influenced by GABAergic disinhibition in that the phase-locking ability of bat IC neurons was increased (Lu et al. 1998). In addition, anesthesia with pentobarbital sodium (which also enhances GABA<sub>A</sub>-mediated inhibition) reduced stimulus-induced activity in neurons in the rat AC (Gaese and Ostwald 2001). The use of midazolam in our study might therefore have led to a decrease in overall neural activity and a sharpening of spatial tuning but a decrease in temporal coding precision. However, previous studies have shown that neural responses measured in bats anesthetized with medetomidine, midazolam, and fentanyl reflected the behavioral performance of the bats well (Firzlaff et al. 2006, 2007). The influence of the anesthesia should therefore only be moderate.

**Spatiotemporal processing in other animals and humans.** The processing of spatiotemporally changing acoustic information has been intensively studied in the context of acoustic motion detection. Neurons sensitive to the direction of moving sound sources have been found in several species, including monkeys (Ahissar et al. 1992), cats (Stumpf et al. 1992), rats (Doan and Saunders 1999), and barn owls (Wagner and Takahashi 1992). Except for motion in the horizontal plane, motion of approaching objects is also of special importance for the survival of an organism and has been investigated in both psychophysical and electrophysiological experiments (Ghazanfar et al. 2002; Maier and Ghazanfar 2007). In addition, imaging studies in humans have revealed the existence of cortical areas specifically involved in motion processing (e.g., Baumgart et al. 1999; Griffiths et al. 1996).

Jenison et al. (2001) measured spatiotemporal receptive fields in the AC of anesthetized cats. They showed that spatial tuning in AC was, in about 14% of their recorded units, inseparable in space and time. This means that spatial tuning changed over time, a feature that was interpreted by the authors in terms of motion direction selectivity. The complex STRMs described in the present study reveal very pronounced and noncontinuous changes of spatial tuning over time. The very different shapes of the spatiotemporal tuning of cortical neurons in cats and bats provide further evidence for the above-mentioned hypothesis that feature recombination in the AC is driven by species-specific requirements. For an echolocating bat, the analysis of pulse-echo delay appears to govern the spatiotemporal tuning in the AC, whereas for a cat, continuous, spatiotemporally arranged excitatory areas may serve the encoding of the motion of an external sound source. This leads to the interesting hypothesis that if we had recorded the STRM in the current (but) study not with sequences of spatially distributed ultrasonic echoes, but with lower-frequency (<30 kHz) noise bursts as in Jenison et al. (2001), STRMs might have revealed very different properties serving passive-acoustic pursuit (Fuzessery et al. 1993). This remains to be tested in future experiments.

In summary, the current data show that spatiotemporal tuning, revealed with a dense, unbiased stimulation technique and a reverse-correlation analysis, is fundamentally different in the bat auditory midbrain and cortex. The conspicuous configuration of STRMs in the IC may serve as a spatiotemporal contrast enhancement and thus represents an advanced stage of echo-acoustic feature extraction. STRMs in the AC typically reflect the result of a task-specific recombination of echo-acoustic features.

**ACKNOWLEDGMENTS**

We thank Harald Luksch and Michael Pecka for comments on an earlier draft of this article.

*J Neurophysiol* • doi:10.1152/jn.00226.2013 • www.jn.org
DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

S.H., A.W., and U.F. performed experiments; S.H., A.W., L.W., and U.F. analyzed data; S.H. and U.F. prepared figures; S.H., L.W., and U.F. drafted manuscript; S.H., A.W., L.W., and U.F. approved final version of manuscript; L.W. and U.F. conception and design of research; L.W. and U.F. interpreted results of experiments; U.F. edited and revised manuscript.

REFERENCES


J Neurophysiol • doi:10.1152/jn.00226.2013 • www.jn.org


J Neurophysiol • doi:10.1152/jn.00226.2013 • www.jn.org